

Supply Chain Management: Supplier Financing Schemes and Inventory Strategies

Min Wang

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
under the Executive Committee
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2012

©2012
Min Wang
All Rights Reserved

ABSTRACT

Supply Chain Management: Supplier Financing Schemes and Inventory Strategies

Min Wang

This dissertation addresses a few fundamental questions on the interface between supplier financing schemes and inventory management. Traditionally, retailers finance their inventories through an independent financing institution or by drawing from their own cash reserves, without any supplier involvement (*Independent Financing*). However, suppliers may reduce their buyers' costs and stimulate sales and associated revenues and profits, by either (i) adopting the financing function themselves (*Trade Credit*), or (ii) subsidizing the inventory costs (*Inventory Subsidies*). In the first part (Chapter 2) we analyze and compare the equilibrium performance of supply chains under these three basic financing schemes. The objective is to compare the equilibrium profits of the individual chain members, the aggregate supply chain profits, the equilibrium wholesale price, the expected sales volumes and the average inventory levels under the three financing options, and thus provide important insights for the selection and implementation of supply chain financing mechanisms. Several of the financing schemes introduce a new type of inventory control problem for the retailers in response to terms specified by their suppliers. In Chapter 3 we therefore consider the inventory management problem of a firm which incurs inventory carrying costs with a *general* shelf age dependent structure and, even more generally, that of a firm with shelf age and delay dependent inventory and backlogging costs. Beyond identifying the structure of optimal replenishment strategies and corresponding algorithms to compute them, it is often important to understand how changes in various primitives of the inventory model impact on the optimal policy parameters and performance measures. In spite of a voluminous lit-

erature over more than fifty years, very little is known about this area. In Chapter 4, we therefore study monotonicity properties of stochastic inventory systems governed by an (r, q) or (r, nq) policy and apply the results in our general theorems both to standard inventory models and to those with general shelf age and delay dependent inventory costs.

Table of Contents

1	Introduction	1
1.1	General Introduction	1
1.2	Inventory Subsidy versus Supplier Trade Credit in Decentralized Supply Chains	3
1.3	Inventory Models with Shelf Age and Delay Dependent Inventory Costs	10
1.4	Monotonicity Properties of Stochastic Inventory Systems	13
2	Inventory Subsidy versus Supplier Trade Credit in Decentralized Supply Chains	19
2.1	Literature Review	20
2.2	Model	23
2.2.1	Independent Financing (IF)	25
2.2.2	Inventory Subsidies (IS)	26
2.2.3	Trade Credit (TC)	26
2.2.4	A General Model	27
2.3	The Stackelberg Game under a Given ECCR	28
2.4	The Remaining Two Stackelberg Games	31
2.4.1	The Stackelberg Game under Given Wholesale Price	31
2.4.2	Comparative Statics and the Full Stackelberg Game	33
2.5	Summary of the Three Games	38
2.6	Comparison of Different Financing Mechanisms	39
2.6.1	Comparing IS and TC	39

2.6.2	Comparing the Two Supplier Financing Mechanisms with IF	40
2.6.3	Numerical Study: Additional Comparison Results	42
2.7	Generalizations with Default Risk	43
2.7.1	Model with Default Risks for the Supplier	45
2.7.2	Model with Default Risks for the Supplier and the Bank	47
2.8	Conclusions	48
3	Inventory Models with Shelf Age and Delay Dependent Inventory	
	Costs	51
3.1	Periodic Review Models	51
3.1.1	Full Backlogging	55
3.1.2	Systems with Lost Sales or Partial Backlogging	60
3.1.3	General Delay Dependent Backlogging Costs	60
3.2	Continuous Review Models	65
3.2.1	Renewal Demand Processes	65
3.2.2	Compound Renewal Processes	72
4	Monotonicity Properties of Stochastic Inventory Systems	79
4.1	Literature Review	80
4.2	Model and Preliminaries	82
4.3	Monotonicity of r^* and R^*	89
4.4	Monotonicity of q^*	95
4.4.1	The Continuous Model	95
4.4.2	The Discrete Model	100
4.5	Monotonicity of the Optimal Cost Value	103
4.6	Conclusions and Future Work	105
A	Appendices for Chapter 2	107
A.1	Proofs for Sections 1-7 of Chapter 2	107
A.2	Algorithm of the Full Stackelberg Game with Participation Constraint	128
A.3	Proof for Section 8 of Chapter 2	130

A.4 Comparisons under a Perfect Coordination Scheme	138
B Proofs for Chapter 3	140
B.1 Proofs	140
C Proofs for Chapter 4	144
C.1 Proofs	144

List of Figures

2.1	Monotonicity of the optimal ECCR and base-stock level	36
4.1	q^* is roughly decreasing with respect to backlogging cost rate b . . .	102

List of Tables

1.1	Comparisons among the financing schemes	9
4.1	Monotonicity results of q^* under some common demand distributions	101

Acknowledgments

First of all, I want to thank my advisor Professor Awi Federgruen. I can hardly overstate my gratitude to him. Without his constant support and encouragement, I would have been lost and probably have given up my pursuit of academic career. His wisdom, passion as well as hardworking would have a lifelong impact on me. He is a true role model for me in my future career as a researcher as well as advisor. I consider it a great honor and privilege to have him as my PhD supervisor and mentor.

My deep thanks go to my advisor at Case Western Reserve University, Professor Matthew Sobel. He gave me important intellectual guidance and unselfish support during my study at the Weatherhead School of Management. I am especially indebted to him for his encouragement and assistance for my application to other PhD programs in my second year at Weatherhead.

I want to give thanks to my dissertation committee members: Professors Omar Besbes, Fangruo Chen, Guillermo Gallego, and Van-Anh Truong. They have taken valuable time from their busy schedule reviewing my thesis, listening to my presentation and providing insightful feedbacks and suggestions.

I would like to thank the many people who have taught and impacted me at Columbia: Professors Nelson Fraiman, Donald Goldfarb, Tim Huh, Alp Muharremoglu, Garrett van Ryzin, Karl Sigman, Gabriel Weintraub, Ward Whitt, David Yao, and Assaf Zeevi. Their knowledge and teaching made study at Columbia a stimulating and rewarding experience.

I am grateful to my fellow students: Roger Lederman, Shiqian Ma, Guodong Pang, and Ruxian Wang. I cherish the friendship with them and enjoy their company during our years at Columbia.

I want to express special thanks to my Christian friends at the Church in New York City, in particular, Wei Chen, Daorong Feng, Li Li, Zhenhong Lv, Telin Sun, and Shengkun Zhang. Their prayers and fellowship had supported me tremendously. They carried me to the light from God when I was weak and fell into darkness.

To my parents and in-laws, Mingzhong Wang and Delan Chu, Shaodong Wang and Cuijuan Yue, my dear husband, Bing, and my daughters, Rebecca and Rachel

Chapter 1

Introduction

1.1 General Introduction

This dissertation addresses a few fundamental questions on the interface between supplier financing schemes and inventory management. In the first part (Chapter 2) we analyze and compare the equilibrium performance of supply chains under three basic financing schemes, described below. The objective is to compare the equilibrium profits of the individual chain members, the aggregate supply chain profits, the equilibrium wholesale price, the expected sales volumes and the average inventory levels under the three financing options, and thus provide important insights for the selection and implementation of supply chain financing mechanisms. Several of the financing schemes introduce a new type of inventory control problem for the retailers in response to terms specified by their suppliers. In Chapter 3 we therefore consider the inventory management problem of a firm which incurs inventory carrying costs with a *general* shelf age dependent structure and, even more generally, that of a firm with shelf age and delay dependent inventory and backlogging costs. Beyond identifying the structure of optimal replenishment strategies and corresponding algorithms to compute them, it is often important to understand how changes in various primitives of the inventory model impact on the optimal policy parameters and performance measures. In spite of a voluminous literature over more than fifty years, very little is known about this area. In Chapter 4, we therefore study mono-

tonicity properties of stochastic inventory systems governed by an (r, q) or (r, nq) policy and apply the results in our general theorems both to standard inventory models and to those with general shelf age and delay dependent inventory costs.

Traditionally, retailers finance their inventories through an independent financing institution or by drawing from their own cash reserves, without any supplier involvement (*Independent Financing*). However, suppliers may reduce their buyers' costs and stimulate sales and associated revenues and profits, by either (i) adopting the financing function themselves (*Trade Credit*), or (ii) subsidizing the inventory costs (*Inventory Subsidies*). In Chapter 2 we characterize the equilibrium performance of a supply chain consisting of a supplier and a retailer under the above three fundamental financing options. We assume that the terms of trade are specified by the supplier, so that the performance of the supply chain is given by the equilibrium of a Stackelberg game with the supplier selecting wholesale prices and/or inventory subsidies or interest charges. (We also address an alternative perspective where these terms of trade are selected to achieve perfect coordination in the decentralized supply chain.) Our main objective is to derive rankings of various performance measures of interests, in particular the expected profit of the individual chain members, the supply chain wide profit, the wholesale price, the expected sales volume and the average inventory level.

In Chapter 3 we consider, in a variety of periodic and continuous review models, the inventory management problem of a firm with shelf age and delay dependent inventory costs. We show how any model with a general *shelf age* dependent holding cost structure may be transformed into an equivalent model in which all expected inventory costs are *level-dependent*. We develop our equivalency result, first, for periodic review models with full backlogging of stockouts. This equivalency result permits us to characterize the optimal procurement strategy in various settings and to adopt known algorithms to compute such strategies. For models in which all or part of stockouts are lost, we show that the addition of any shelf age dependent cost structure does not complicate the structure of the model beyond what is required under the simplest, i.e., linear holding costs. We elaborate a similar equivalency

result for general delay dependent backlogging cost structures; *this* equivalency requires either a restriction on the actions sets or on the shape of the backlogging cost rate functions. We proceed to show that our results carry over to continuous review models, with demands generated by compound renewal processes; the continuous review models with shelf age and delay dependent carrying and backlogging costs are shown to be equivalent to periodic review models with convex level dependent inventory cost functions.

In Chapter 4, we consider inventory systems which are governed by an (r, q) or (r, nq) policy. We derive general conditions for monotonicity of the optimal cost value and the three optimal policy parameters, i.e., the optimal reorder level, order quantity and order-up-to level, as a function of the various model primitives, be it cost parameters or complete cost rate functions or characteristics of the demand and leadtime processes. These results are obtained as corollaries from a few general theorems, with separate treatment given to the case where the policy parameters are continuous variables and that where they need to be restricted to integer values. The results are applied both to standard inventory models and to those with general shelf age and delay dependent inventory costs.

1.2 Inventory Subsidy versus Supplier Trade Credit in Decentralized Supply Chains

It is well known and broadly documented that in the United States and Europe, companies depend heavily on supplier financing mechanisms for their working capital, which consists primarily of inventories. For example, Petersen and Rajan (1997), quoting Rajan and Zingales (1997), observed for the United States that trade credit financing is the single largest source of external short-term financing. For European markets, this phenomenon has been documented by Wilson and Summers (2002) and Giannetti et al. (2008). If trade credit financing is the dominant source of credit in first world countries, it is very likely to be more dominant in emerging economies with a less developed banking industry and capital markets. In addition,

reliance on trade credit financing, by necessity, increases in economic environments where bank credit is severely curtailed.

Under trade credit, a supplier adopts the complete financing function, traditionally assumed by a third-party financial institution - hereafter referred to as the bank - or by the customer herself drawing from her own cash reserves. Inventory subsidies represent a third alternative: here, the financing function continues to be assumed by a bank or the customer herself, but the supplier agrees to cover part of the financing and/or physical inventory costs. This practice prevails, for example, in the automobile industry when manufacturers pay the dealer so-called “holdbacks”, i.e., a given amount for each month a car remains in the dealer’s inventory. (The holdback amount may be varied as a function of the amount of time the car has been on the dealer’s lot.) Other industries where suppliers provide inventory subsidies to retailers and distributors include the book and music industries as well as personal computers, apparel and shoes. See Narayanan et al. (2005) and Nagarajan and Rajagopalan (2008) for a more detailed discussion of the prevalence of inventory subsidies.

Many have been intrigued why supplier financing is as prevalent as it is. After all, the credit function would seem to be a core competency of the banking world. The economics literature offers a variety of explanations: first, there are the aforementioned lending capacity limits resulting from internally or externally imposed capital ratio requirements. Mian and Smith (1992) argue that suppliers may be in a better position than banks to monitor what activities credit loans are used for. Additional explanations can be found in Biais and Gollier (1997), Jain (2001), Cuñat (2007), Burkart and Ellingsen (2004), Frank and Maksimovic (2005), Nadiri (1969) and Wilner (2000).

However, the above explanations ignore a primary function of various supplier financing mechanisms, namely, to reduce the customer’s risks and to share these risks in the most advantageous way possible, thereby stimulating sales and associated revenues and profits.

While trade credit financing has primarily been analyzed by economists, it ap-

pears that inventory subsidies have been studied, almost exclusively, in the operations management literature; see, for example, Anupindi and Bassok (1999), Cachon and Zipkin (1999), Narayanan et al. (2005) and Nagarajan and Rajagopalan (2008). These papers have demonstrated that inventory subsidies may be an advantageous way, for the supply chain as a whole, to reduce the customers' risks and to stimulate their purchases and sales.

The objective of this paper is to characterize the equilibrium performance of a supply chain consisting of a supplier and a (single) buyer, hereafter referred to as the retailer, under the following three fundamental financing options:

(I) *Independent Financing* (IF): this reflects the traditional business model where the retailer finances her inventories through a bank or by drawing from her own cash reserves, without any supplier involvement.

(II) *Inventory Subsidies* (IS): same, except that the supplier offers to cover a specific part of the capital costs associated with the retailer's inventories.

(III) *Trade Credit* (TC): here the supplier adopts the financing role otherwise assumed by a bank or the retailer herself, as in (I) and (II).

In particular under TC arrangements, interest charges may accrue at a rate which depends on the amount of time a unit has been in stock. For example, the credit terms may include an interest free grace period, as in "30 (60, 90) days net". Similarly, inventory subsidies may be dependent on the "shelf age" of the items. The aforementioned "holdbacks" in the automobile industry is a case in point: typically the holdback is only paid for a limited period of time - for example a quarter - during which a car remains in a dealer's lot. To provide a fundamental framework to compare the equilibrium performance of the supply chain under the three financing options, we confine ourselves to the case where interest accrues at a *constant* rate.

The main objective of this paper is to compare equilibrium profits of the individual chain members, supply chain profits, the equilibrium wholesale price, the expected sales volumes and the average inventory levels under the three financing options.

To this end, we develop a model unifying the three mechanisms, which is based on the following assumptions: we consider a supply chain with a single supplier providing a single item to a retailer who sells the item to consumers at a given retail price. We consider a periodic review, infinite horizon model where consecutive demands are iid with a known distribution. Demands that can not be satisfied from existing stock are lost. The supplier incurs variable procurement costs at a given cost rate. Two types of inventory carrying costs are incurred for the retailer's inventories: physical storage and maintenance costs and financing costs which depend on the specific financing mechanism adopted by the supply chain. Both the supplier and the retailer face a per dollar financing cost α_s and α_r , per unit of time, when using the bank as a financier of its inventories or when financing the inventories from internal funds. The capital cost rates α_s and α_r are typically significantly different from each other, even when both chain members use bank loans to finance their working capital, see §2.2. In our base model, we assume bankruptcy risks are negligible. However, in §2.7, we discuss two generalizations of the base model where the retailer may default.

We assume that the terms of trade are specified by the supplier to maximize his expected profit under the corresponding optimal procurement policy of the retailer. This perspective gives rise to so-called Stackelberg games with the supplier (the retailer) as the leader (the follower). It reflects many, if not most, supply chain settings and explains why this is assumed in most supply chain models. Other perspectives do arise as when a perfect coordination mechanism is adopted, with the aggregate first-best profits split in accordance with a given allocation rule, such as a Nash bargaining solution, reflecting the relative bargaining powers of the chain members. We pursue the latter perspective in Appendix A.4.

We now summarize our main results. We distinguish among three Stackelberg games, depending upon whether only the wholesale price, or only the financing terms (i.e., the trade credit interest rate, under TC, or the subsidy for the capital cost rate, under IS) are selected endogenously, or whether the supplier starts out selecting *both*. (We refer to the latter game as the full Stackelberg game.) We fully characterize

the optimal strategies of the supply chain members in these three games. (See §2.5 for a summary of the many intuitive and counterintuitive structural properties of the equilibria in the various Stackelberg games, as well as comparative statics results.) The characterization of the equilibrium in the full Stackelberg game requires that the demand distribution satisfies a slightly stronger variant of the Increasing Failure Rate (IFR) property. We show that this variant of the IFR condition is satisfied by many families of distributions, in particular, all uniform, exponential and Normal distributions.

We proceed with a systematic comparison of the various above mentioned equilibrium performance measures across the different financing mechanisms. We show, in full generality, that the supplier is better off under the equilibrium TC arrangement as opposed to IS, if and only if his cost of funds (α_s) is lower than that of the retailer (α_r). As to the remaining comparisons, we confine ourselves to the above three classes of demand distributions (and a few, very minor parameter conditions). Here, we show that the retailer's and the supplier's preference for the IS versus TC contract are perfectly aligned. In other words, the retailer's optimal profit level is higher under TC as opposed to IS if and only if her cost of funds is higher than that of the supplier. The same, simple, necessary and sufficient condition reveals whether the wholesale price is lower, and the expected sales volume and average inventory level are higher under IS, or whether the opposite rankings prevail.

Assume next that the supply chain initially operates under IF where the retailer arranges her own financing internally or from a third-party bank without any supplier's subsidies. If the supplier maintains the wholesale price that applies under IF, both supply chain members benefit by switching to an IS arrangement. Maintaining the same wholesale price, it is usually, although not always, beneficial for both supply chain members to switch from IF to a TC arrangement as well. (We show that if the supplier's cost of capital is lower than that of the retailer, this is indeed guaranteed, in fact with greater benefits accruing to both chain members than under the IS arrangement.)

Clearly, if the supplier is willing and able to deviate from the wholesale price he

would charge under IF, and adopt an IS agreement with a wholesale price-inventory subsidy combination which optimizes his profits in the IS Stackelberg game, this results in additional profit improvements for him (beyond those achieved when maintaining the old wholesale price). However the same need not apply to the retailer's profit. Indeed, we have conducted an extensive numerical study which consistently reveals that the resulting profits for the retailer are *lower* than those she enjoys under IF. The overall conclusion therefore is that the adoption of an appropriately designed IS or TC agreement always benefits the supplier, but to entice the retailer, the specific terms need to be specified to ensure that her resulting profits are maintained or improved as well. This gives rise to Stackelberg games with a participation constraint, which we characterize in Appendix A.2. The above numerical study identifies several other rankings between equilibrium performance under IF, versus those under TC or IS.

Table 1.1 summarizes our comparison results. Π^* , Π_r^* and Π_s^* denote the expected aggregate profits and that of the supplier and the retailer, respectively, while w^* and s^* denote the equilibrium wholesale price and the expected sales volume per period. Superscripts indicate which of the three financing mechanisms the measure refers to. All inequalities with a numbered footnote are proven in the paper, the footnote indicating where precisely. The remaining rankings are based on the above extensive numerical study reported.

The above structural results for the Stackelberg games carry over to the two generalized models considered in this paper, where the retailer *may* default. In this case some or all of the creditors receive only part of the amount due to them, the so-called recovery rate. In the first generalized model, we assume only the supplier faces this default risk when engaging in a TC arrangement while the bank has immunized itself from this risk; in the second generalization, both the supplier and the bank face default risks. As to the comparison results across these financing mechanisms, they carry over to the first generalized model in the sense that many rankings can be established by comparing α_r with an index that depends on α_s , the recovery rate and parameters describing the default and reorganization process. In the second

	$\alpha_s \leq \alpha_r$	$\alpha_s > \alpha_r$
Supplier's Profit	$\Pi_s^{*IF} \leq^1 \Pi_s^{*IS} \leq^2 \Pi_s^{*TC}$	$\Pi_s^{*IF}, \Pi_s^{*TC} \leq^3 \Pi_s^{*IS}$
Retailer's Profit	$\Pi_r^{*IS} \leq^4 \Pi_r^{*TC} \leq^a \Pi_r^{*IF}$	$\Pi_r^{*TC} \leq^5 \Pi_r^{*IS} \leq \Pi_r^{*IF}$
Aggregate Profit	$\Pi^{*IF} \leq \Pi^{*IS} \leq^6 \Pi^{*TC}$	$\Pi^{*IF}, \Pi^{*TC} \leq^7 \Pi^{*IS}$
Wholesale Price	$w^{*IF} \leq w^{*TC} \leq^8 w^{*IS}$	$w^{*IF} \leq w^{*IS} \leq^9 w^{*TC}$
Expected Sales	$s^{*IF} \leq s^{*IS} \leq^{10} s^{*TC}$	$s^{*IF}, s^{*TC} \leq^{11} s^{*IS}$

Table 1.1: Comparisons among the financing schemes

1-2: See Proposition 2.4 (i) and Theorem 2.5 (b) for a proof, respectively.

3: See Proposition 2.4 (i) and Theorem 2.5 (b) for a proof. 4-5: See Theorem 2.6 (c) for a proof.

6: It follows from inequalities in footnotes 2 and 4

7: $\Pi^{*TC} \leq \Pi^{*IS}$ follows from inequalities in footnotes 3 and 5.

a: It holds numerically in all but nine instances. The nine exceptions are instances where the supplier's cost of capital is substantially lower than that of the retailer and the variable profit margin $(p - c)/c = 0.11$ is small. The small profit margin severely limits the supplier's ability to raise the price while his significant capital cost rate advantage allows for major profit improvement compared to IF.

8-9: See Theorem 2.6 (e) for a proof. 10-11: $s^{*TC} \leq s^{*IS}$ follows from Theorem 2.6 (d).

model with default risks, these comparisons need to be made numerically.

1.3 Inventory Models with Shelf Age and Delay Dependent Inventory Costs

One of the main objectives of any inventory planning model is to analyze the trade-off between competing risks of overage and underage. This requires an adequate representation of the carrying costs associated with all inventories, as well as the cost and revenue consequences of shortages. Early contributors, e.g., the pioneering textbooks by Hadley and Whitin (1963) and Naddor (1966), discussed possible paradigms to represent the carrying and shortage costs.

One standard paradigm is to assume that carrying costs can be assessed, either continuously or periodically, as a (possibly non-linear) function of the prevailing total inventory, irrespective of its age composition. Similarly, shortage costs are assumed to accrue as a (again, possibly non-linear) function of the total shortfall or backlog, irrespective of the amount of time the backlogged demand units have remained unfilled. We refer to this type of carrying and shortage cost structures as *level-dependent* inventory costs. After the above mentioned early discussions in Hadley and Whitin (1963) and Naddor (1966), *this* paradigm has been adopted in virtually every inventory model.

There are, however, many settings where carrying costs need to be differentiated on the basis of the inventory's shelf-age composition. First, inventories are often financed by trade credit arrangements, where the supplier allows for a payment deferral of delivered orders, but charges progressively larger interest rates as the payment delay increases. For example, the supplier frequently offers an initial interest-free period (e.g., 30 days) after which interest accumulates. Moreover, interest rates often increase as a function of the item's shelf age. These trade credit schemes have been considered in Gupta and Wang (2009) as well as Chapter 2. We refer to the latter for a discussion of how prevalent this practice is. Another setting with shelf age dependent inventory cost rates arises when the supplier subsidizes

part of the inventory cost. For example, in the automobile industry, manufacturers pay the dealer so-called “holdbacks”, i.e., a given amount for each month a car remains in the dealer’s inventory, up to a given time limit (see, e.g., Nagarajan and Rajagopalan (2008)). The resulting inventory cost rate for any stocked item is, again, an increasing function of the item’s shelf age.

Even when inventory costs grow as a *linear* function of the loan term or the amount of time the purchased units stay in inventory, *time varying* purchase prices or interest rates necessitate disaggregating inventory levels according to the time at which the units were purchased, i.e., in accordance with the items’ shelf age. As an example, in the dynamic lot sizing literature, Federgruen and Lee (1990) modeled holding costs as proportional to the items’ purchasing price, which varies with their purchase period. As a consequence, holding costs depend on the items’ shelf age. Even more general shelf-age dependencies are assumed in Levi et al. (2011) and its generalization, i.e., so-called metric holding costs, in Stauffer et al. (2011). Finally, beyond capital costs, inventories often incur maintenance related expenses; these, too, vary as a function of the items’ shelf age.

Similar to shelf age dependent holding costs, backlogging costs may also depend on the amount of time by which delivery of a demand unit is delayed. This may reflect the structure of contractually agreed upon penalties for late delivery or, in case of *implicit* backlogging costs, the fact that customers become less or more impatient over time. This type of backlogging costs has been studied by Chen and Zheng (1993), Rosling (1999, 2002) and Huh et al. (2010).

In this paper, we show how periodic and continuous review models with a general *shelf age* dependent holding cost structure may be transformed into an equivalent “standard” model in which all expected inventory costs are *level-dependent*. These equivalency results allow for the rapid identification of the structure of an optimal policy, in various models. It also allows for the immediate adoption of algorithms to *compute* optimal policies. Moreover, in periodic review models, all shelf age dependent inventory holding cost components are transformed into linear holding costs, however, with a specific modified random leadtime distribution.

We develop our equivalency result, first, for models with full backlogging of stockouts. These models allow for a one-dimensional state representation via the so-called *inventory position*, i.e., on-hand inventory + outstanding orders - backlogs. This equivalency result permits us to characterize the optimal procurement strategy in various settings. For example, assuming demands are independent with exogenously given distributions, a simple time-dependent (s, S) policy, acting on the inventory position is optimal under fixed-plus-linear order costs. In the absence of fixed delivery costs, this structure further simplifies to that of a base-stock policy. In the special case where all model parameters are stationary, we show that the base-stock levels increase as we progress to the end of the planning horizon; moreover these levels can be determined myopically by computing, for each period, the minimum of a period specific, closed-form convex function. When each of the demand distributions depends on the buyer's retail price, the optimal combined inventory and pricing strategy is a so-called base-stock/list price policy, assuming no fixed ordering costs prevail, and leadtimes are negligible. In the presence of such fixed costs, and assuming the stochastic demand functions have additive noise terms, the optimal combined strategy is of the so-called (s, S, p) structure: the procurement part of the combined strategy continues to be of the (s, S) type. Other variants of this model and of the associated optimal strategies are discussed as well.

We generalize our equivalency results for models in which all or part of stockouts are lost. Here, the state of the system needs to be described with a multi-dimensional inventory vector: more specifically, under a positive leadtime $L > 0$, it is well known that the state of the system needs to be represented by an $(L + 1)$ - state vector, keeping track of the inventory on hand and all outstanding orders from the last L periods, separately. Here, we show that the addition of any shelf age dependent cost structure does not complicate the structure of the model beyond what is required under the simplest, i.e., *linear* holding costs.

A different transformation, due to Huh et al. (2010), allows for the treatment of general delay dependent backlogging cost structures, but only under an assumption guaranteeing either that no demand unit is delayed by more than the leadtime plus

one periods, or that the incremental backlogging cost rate is constant for delays in excess thereof. The first condition is equivalent to assuming that the inventory position after ordering is always non-negative. We review this transformation in §3.1.3, and prove various optimality results that can be obtained under general delay dependent cost structures.

In §3.2, we show how general shelf age dependent holding and delay dependent backlogging costs can be handled in *continuous* review models. Starting with the case of renewal demand processes, we show that, for these, an equivalent model with a convex, inventory level dependent cost structure can be obtained, in full generality, i.e., without *any* policy restrictions. This equivalency result is based on a very different so-called “single unit decomposition approach”, in the spirit of those introduced by Axsäter(1990, 1993) and Muharremoglu and Tsitsiklis (2008). The equivalency result allows us to conclude, for example, that in a system with fixed-plus-linear ordering costs, an (r, q) - policy is optimal, and the long run average cost as a function of r and q , is of a structure enabling the use of the algorithm in Federgruen and Zheng (1992) to identify the *optimal* parameters. Finally we characterize how various model primitives such as the leadtime distribution, the shape of the marginal shelf age dependent cost function and that of the delay dependent backlogging cost function impact the optimal policy parameters r^* and $R^* \equiv r^* + q^*$. We also show (in §3.2.2) that the results for periodic inventory systems carry over to continuous review models with general *compound* renewal demand processes. We show that under minor assumptions, the model is equivalent to a periodic review model with convex inventory level dependent carrying and backlogging costs. Under fixed-plus-linear costs, this implies, for example, that an (s, S) policy is optimal.

1.4 Monotonicity Properties of Stochastic Inventory Systems

In the past fifty years, a voluminous literature has arisen on inventory models. In many elementary models, we are able to prove that the optimal procurement

strategy has a relatively simple structure characterized by a few policy parameters. Moreover, for several of those models, we have identified efficient algorithms, able to compute the optimal combination of policy parameters.

Nevertheless, most of these models fail to be used widely by practitioners or to be taught in Operations Management classes or textbooks, with the exception of Economic Order Quantity (EOQ)- and newsvendor type models. The continued popularity of the latter two classes of inventory models can not be attributed to the applicability of the underlying model assumptions, which, in fact, are very restrictive and fail to fit many of the settings where they are routinely applied. Instead, their continued popularity is based on the fact that they allow for closed form expressions of the optimal policy parameters, thus providing easy and immediate insights into how various model primitives (cost parameters, demand processes, leadtimes etc.) impact on the above policy parameters. As articulated by Geoffrion (1976), the main purpose of models is to provide “insights, not numbers”. At the most basic level, the model user wishes to understand whether optimal policy parameters and associated performance measures increase or decrease as a function of the various model primitives.

In this paper, we derive general conditions under which monotonicity of the optimal parameters and associated key performance measures, with respect to *general* model primitives, can be established within a (single-item) inventory system governed by an optimal (r, q) or (r, nq) policy. Under an (r, q) policy, the system is monitored continuously and a replenishment order of a fixed size q is placed whenever the inventory position drops to the level r . When the demand process experiences jumps of an arbitrary magnitude, it is sensible to apply an (r, nq) policy, with the order quantity specified as the minimum multiple of q required to bring the inventory position back above the reorder level r . (r, q) or (r, nq) policies are also frequently used in serial systems, see, e.g., Shang and Zhou (2009, 2010). In contrast to the above EOQ related deterministic models, the optimal (r, q) or (r, nq) policy parameters can not be obtained in closed-form, but need to be computed algorithmically, even under the simplest demand processes, i.e., Brownian motions or

Poisson processes. Prior literature, reviewed in the next section, has studied the impact of changes of a few *specific* model primitives, in particular, the leadtime and the leadtime distributions.

Many key performance measures are directly related to the optimal policy parameters r^* , $R^* \equiv r^* + q^*$ and q^* . Operations managers are concerned with the maximum inventory (position), the average inventory level and the minimum inventory, the latter being related to the so-called safety stock concept. Logistics managers focus on the average order size or order frequency, its reciprocal. Suppliers often prefer a regular order pattern associated with high order frequency to allow for a smooth production/distribution schedule. Financial analysts and macroeconomists pay particular attention to the sales/inventory ratio, also referred to as the inventory turnover.

Beyond providing general insights into inventory systems governed by (r, q) or (r, nq) policies, the above monotonicity properties have additional benefits: *first*, many of the parameters or distributions in the model are difficult to forecast and the model user needs to understand in which direction an under- (or over-)estimate biases the optimal policy parameters. *Second*, the monotonicity properties can be exploited when the model needs to be solved repeatedly for many parameter values. This situation arises either because of uncertainty about a parameter or because service level constraints are added to the model which, when handled via Lagrangian relaxation, requires the repeated optimization of a traditional aggregate cost function for many multiplier values or combinations thereof. (Such service level constraints include constraints on the *fill rate*, i.e., the fraction of demand that can be filled immediately without backlogging, or the *ready rate*, i.e., the fraction of time the system has stock, or the expected amount of time a backlogged demand has to wait before being filled.) If it is known that an increase in a parameter or Lagrange multiplier from a value μ^0 to μ^1 results in an increase or (decrease) of q^* , say, this fact can be exploited, for example, when using the algorithm in Federgruen and Zheng (1992): when re-optimizing the model for $\mu = \mu^1$, one may, then, start with $q = q^*(\mu^0)$ and increase(decrease) q in the outer loop in the algorithm, see the

Appendix.

Depending upon whether the sample paths of the leadtime demand process are continuous or step functions, the long-run average cost is of the form:

$$c(r, q|\theta) = \frac{\lambda K + \int_r^{r+q} G(y|\theta) dy}{q}, \quad (1.1)$$

or

$$c(r, q|\theta) = \frac{\lambda K + \sum_{y=r+1}^{r+q} G(y|\theta)}{q}. \quad (1.2)$$

In both (1.1) and (1.2), λ and K represent the long-run average demand rate and the fixed cost incurred for every order batch of size q respectively. All other model primitives $\theta \in \Theta$ impact the long-run average cost exclusively via the so-called instantaneous expected cost function $G(y|\theta)$. When the long-run average cost of an (r, q) or (r, nq) policy is given by (1.1)[(1.2)], we refer to the model as the continuous [discrete] model. Since the representations in (1.1) and (1.2) are common under (r, q) or (r, nq) policies, we henceforth confine ourselves to the former, without loss of generality.

The fixed cost K impacts only the first term in the numerator of (1.1) and (1.2). Zheng (1992) already showed that the optimal reorder level r^* is decreasing while the optimal order size q^* and the optimal order-up-to level $R^* \equiv r^* + q^*$ are increasing in this parameter¹. In contrast, the average demand rate λ impacts both terms in the numerator of the long-run average cost function and the net monotonicity effect on the optimal policy parameters is therefore, sometimes, ambiguous². We establish our monotonicity properties with respect to all other general model primitives $\theta \in \Theta$, merely requiring that the space Θ be endowed with a partial order \preceq . As such, θ may be a cost parameter, or a parameter of the demand or leadtime distribution. Alternatively, θ may represent the distribution of a random variable or a complete stochastic process, or a cost rate function.

¹Zheng (1992) confines himself to the continuous model; A similar treatment of the discrete model can easily be obtained based on the algorithm in Federgruen and Zheng (1992).

²See, however, Corollary 4.1(f)-(g) and Corollary 4.2(e).

Our first main result is that the optimal reorder level r^* and the optimal order-up-to level R^* are decreasing (increasing) in θ whenever the function $G(y|\theta)$ is supermodular (submodular) in (y, θ) , that is, any of the difference functions $G(y_2|\theta) - G(y_1|\theta)$, with $y_1 < y_2$, is increasing (decreasing) in θ . Thus, the monotonicity patterns of r^* and R^* are identical in the continuous model (1.1) and the discrete model (1.2) and the general conditions under which they are obtained are identical as well.

As to the remaining policy parameter q^* , i.e., the optimal order quantity, here the monotonicity patterns that can be expected, differ themselves, between the continuous model (1.1) and the discrete model (1.2). In the continuous model, we show that q^* can often be guaranteed to be monotone in various model parameters. In the discrete model, occasional unit increases (decreases) between stretches where $q^*(\theta)$ is decreasing (increasing) can not be excluded. This gives rise to a new monotonicity property which we refer to as *rough monotonicity*: an integer valued function is roughly decreasing (increasing) if the step function does not exhibit any pair of consecutive increases (decreases). We show that *pure* monotonicity of q^* in the continuous model and *rough* monotonicity in the discrete model, with respect to any model parameter, can be guaranteed if the supermodularity or submodularity property of $G(\cdot|\theta)$ function is combined with a single additional structural property of this instantaneous expected cost function. While the conditions in the continuous and discrete model are very similar, the required analysis is fundamentally different.

Finally, we identify a broad sufficient condition for monotonicity of the optimal cost value; to our knowledge, this condition encompasses all known applications as well as several new ones.

The most frequently used model in which the long-run average cost of an (r, q) or (r, nq) policy is given by (1.1) or (1.2), has the following assumptions: the item is obtained at a given price per unit; inventory costs are accrued at a rate which is a convex increasing function of the inventory level; stockouts are backlogged where backlogging costs are, again, accrued at a rate which is a convex increasing function of the backlog size; leadtimes are generated by a so-called exogenous and sequential process, ensuring that consecutive orders do not cross and the leadtimes

are independent of the demand process. We refer to this as the *standard inventory model*.

For these standard inventory models, our general results imply, in particular, that r^* and R^* are decreasing in the item's purchase price, assuming that the inventory carrying cost rate function increases monotonically with the purchase price. Similarly, r^* and R^* are decreasing in other parameters on which the marginal inventory carrying cost rate function depends monotonically, for example, the physical maintenance and warehousing cost per unit of inventory, or more generally, when the marginal holding cost rate function is replaced by a pointwise larger one. In contrast, r^* and R^* are both *increasing* when the marginal backlogging cost rate function is replaced by a pointwise larger one. As a final application for the standard inventory model, compare two leadtime demand processes such that the leadtime demand distribution under the first process is stochastically smaller than that under the second process. (Dominance of the steady-state leadtime demand distribution may arise because of a change of the demand process, a stochastic enlargement of the leadtime distribution, or both.) We show that r^* and R^* are always smaller under the first process compared to the latter. As far as q^* is concerned, our general results imply, for example, monotonicity with respect to the purchase price and holding cost rates, assuming that the leadtime demand distribution is log-concave or log-convex, a property shared by most classes of distributions. Similarly, q^* is monotone in the backlog cost rate if the complementary cumulative distribution of the leadtime demand distribution is log-concave or log-convex. In the case of normal leadtime demands, q^* is monotone in their mean and standard deviation. Similarly, if the demand process is a Brownian motion and leadtimes are fixed, q^* is increasing in the drift and volatility of the Brownian motion and in the leadtime. Sufficient conditions for (rough) monotonicity can often be stated in terms of broadly applicable properties of the cdf of the leadtime demand distribution such as log-concavity.

Chapter 2

Inventory Subsidy versus Supplier Trade Credit in Decentralized Supply Chains

We refer to Section 1.2 for an introduction of this chapter. In Section 1.2 we introduced the following three fundamental financing options:

(I) *Independent Financing* (IF): this reflects the traditional business model where the retailer finances her inventories through a bank or by drawing from her own cash reserves, without any supplier involvement.

(II) *Inventory Subsidies* (IS): same, except that the supplier offers to cover a specific part of the capital costs associated with the retailer's inventories.

(III) *Trade Credit* (TC): here the supplier adopts the financing role otherwise assumed by a bank or the retailer herself, as in (I) and (II).

This chapter is organized as follows. §2.1 reviews related literature. In §2.2 we model the supply chain under the IF, IS and TC financing schemes. We show that all three models can be synthesized into a single unified model. For this unified model, §2.3 characterizes the equilibrium behavior of the first Stackelberg game with an exogenously given inventory subsidy or trade credit interest charge. §2.4 achieves the same for the remaining two Stackelberg games, i.e., the game with an

exogenously given wholesale price and the full Stackelberg game, and is followed by a brief §2.5 in which the various structural results for the three games are summarized. §2.6 derives the above mentioned comparison results among the three financing mechanisms. In §2.7 we develop the two generalized models with default risks. The final §2.8 summarizes our findings and discusses other variants of our models.

2.1 Literature Review

Several papers in the operations management literature have analyzed the interaction between suppliers and retailers facing demand risks, under one of the above mentioned payment schemes. Most of the literature confines itself to a single supplier servicing a single retailer. The specific payment terms are either assumed to be selected by the supplier so as to maximize his profits, or by a third party (coordinator) so as to maximize chain wide profits. The former perspective gives rise to a Stackelberg game with the supplier as the leader, while in the latter the central question is whether a perfect coordination mechanism exists and if so how these parameters are to be selected.

Wholesale price only contract

The first Stackelberg game model in this general area is Lariviere and Porteus (2001), considering a single period model under a simple constant wholesale price scheme without any additional incentives, as in IF. The authors show that the supplier's equilibrium profit function under the retailer's best response is unimodal in the wholesale price as long as the demand distribution satisfies a generalization of the IFR property, which the authors refer to as IGFR (Increasing General Failure Rate). See Lariviere (1999) and Cachon (2003) for detailed reviews of this model. Numerical studies in these papers show that the chain wide profits under the Stackelberg solution are 20-30% below those obtained in the centralized solution. However, perfect coordination can only be achieved when the wholesale price equals the supplier's variable cost rate as shown by Pasternack (1985), resulting in an unsatisfactory arrangement where the supplier's profits are reduced to zero.

Wholesale price and inventory subsidy contract

Anupindi and Bassok (1999) and Cachon and Zipkin (1999) appear to be the first to consider the inventory cost subsidy as a mechanism by the supplier to assume part of inventory risks. As in IS, the former assume that the retailer(s) obtain independent financing for their purchases, while the supplier assumes part of the resulting financing/inventory costs. Considering a periodic review infinite horizon model, Anupindi and Bassok (1999) deals, among others, with settings where all retail demand is satisfied from a single stocking point. The authors analyze both the Stackelberg game that arises when the wholesale price is given and the supplier selects his subsidy of the inventory cost rate as well as the Stackelberg game that prevails when the wholesale price is chosen, but only in the special case of inventory cost subsidies. For the latter Stackelberg game, Anupindi and Bassok (1999) show that under Normal demands, an approximation for the equilibrium supplier's profit function is concave. For the former game, unimodality is verified numerically. We prove that the *exact* supplier's equilibrium profit function is unimodal in both Stackelberg games for a broad class of demand distributions which include the Normals as a special case, and for arbitrary choices of the exogenously specified contract terms. In addition, we characterize the (full) Stackelberg game which arises when the supplier controls both his wholesale price and his inventory subsidy. The single period model in Zhou and Groenevelt (2008) may be viewed as a variant of the Stackelberg game under a given inventory subsidy, indeed, the extreme case where the supplier assumes all the inventory costs. These authors also incorporate the possibility of a retailer going bankrupt when his loss exceeds a certain threshold. Another departure from the literature is the assumption that the retailer's capital cost rate is *endogenously* selected by the bank along with a maximum percentage of the purchase order which he is willing to finance. The bank selects these financing terms in advance of the supplier's wholesale price so as to break even in expectation.

Other papers have established that inventory subsidies may be used as an essential component of coordination mechanisms, again in a supply chain with a single supplier and a single retailer. Cachon and Zipkin (1999) establish this in an infinite

horizon model, assuming both chain members employ a base-stock policy. Similarly, Wang and Gerchak (2001) show, in a *single* sales period model, that a combined wholesale price/inventory cost rate subsidy contract achieves perfect coordination in the supply chain. (The demand distribution may depend on the stocking level.) The availability of an inventory cost rate subsidy as a lever allows for a continuum of coordination mechanisms beyond the single and unsatisfactory (wholesale price only) mechanism identified by Pasternack (1985). Narayanan et al. (2005) analyze the same combined wholesale price/inventory cost rate subsidy contracts to coordinate a supply chain consisting a supplier and two retailers. Finally, Nagarajan and Rajagopalan (2008) consider inventory subsidies in the context of Vendor Managed Inventories where an inventory cost rate subsidy is specified as part of the contractual agreement.

Wholesale price and trade credit contract

Kouvelis and Zhao (2009) and Yang and Birge (2011) have analyzed settings where the supplier himself acts as the financing institution, offering a trade credit option. In the former's Stackelberg game model, the supplier specifies a cash-on-delivery wholesale price and a capital cost rate for units paid at the end of the sales period. As in Zhou and Groenevelt (2008), the authors assume that the retailer's ability to pay upfront is constrained by her initial cash balance and they incorporate the possibility of bankruptcy at the end of period. This supplier financing scheme is compared with the one in which an outside bank finances the units bought on credit. As in Zhou and Groenevelt (2008), the interest rate is determined so that the bank breaks even in expectation. Yang and Birge (2011) consider a variant of this model, incorporating credit limitations for the supplier and a liquidation or distress cost in the event of bankruptcy.

Other related papers include Gupta and Wang (2009) who characterize the retailer's optimal procurement strategy under a TC payment scheme. In their paper, the supplier charges interest as a general nonlinear function of the amount of time elapsed between the delivery of goods and the payment. The authors show that under the standard linear holding and backlog cost the base-stock policy is optimal.

In Chapter 3, we provide a simple proof and extend the structural result for the optimal procurement policy to many other settings, for example, those with fixed procurement costs and/or price-dependent demand.

Finally, we refer to Buzacott and Zhang (2004) and references there for a limited stream of papers addressing the important topic of inventory management under credit restrictions in a single firm setting.

2.2 Model

We characterize the interaction between the supplier and the retailer in an infinite horizon periodic review system. We choose an infinite planning horizon so as to model an ongoing trade relationship between the supplier and the retailer involving many repeated procurement decisions for storable items. (The stationary infinite horizon model is a standard framework in the supply chain literature when representing repeating procurement decisions, see for example Cachon (2003) and Zipkin (2000).) While some inventory models involve parameters and distributions that fluctuate cyclically or, in dependence of a more general exogenous state variable, the treatment of such phenomena appears tangential to the questions raised in this paper and complicates its analysis and results needlessly. In other words, there is no reason to believe that the schemes' relative advantages and disadvantages differ in environments with stationary versus fluctuating parameters. See, however, §2.8 for a generalization of our model in which interest rates fluctuate stochastically.

The retailer faces a sequence of independent and identically distributed customer demands under an exogenously given retail price. She may place a purchase order with the supplier at the beginning of any period. Orders placed at the beginning of a period arrive in time to satisfy that period's demand. Unsatisfied demand results in lost sales. As a consequence, the retailer's sales volume, and hence, that of the supplier depend on the retailer's inventory replenishment strategy; the latter, in turn, depends on the structure of the selected supplier financing scheme. We distinguish between two types of inventory carrying costs: (i) physical storage and maintenance

costs, assumed to be proportional with each end-of-the-period inventory level, (ii) financing costs, the structure of which depends on the specific terms of the payment scheme in place. In the base model, we assume that the likelihood of the retailer defaulting on her payment is negligible. See, however, §2.7 for generalizations that allow for defaults.

Let D = the random demand observed in an arbitrary period, with a known cdf $F(y)$, continuously differentiable pdf $f(y)$, mean $\mu < \infty$ and standard deviation $\sigma < \infty$.

p = the per unit retail price.

c = the per unit supplier's variable procurement cost rate.

h_0 = the physical (storage and maintenance) cost per unit carried in inventory at the end of a period.

As mentioned in the Introduction, α_r denotes the capital cost rate incurred by the retailer under independent financing, that is when drawing from her own cash reserves or from a credit line offered by the bank. (The size of the cash reserves or the credit line is assumed to be ample.) In the former (latter) case, α_r represents the rate of return on its best alternative investment option (bank loan rate). If both capital sources are available, α_r denotes the lower of the two rates. α_s denotes the capital cost rate incurred by the supplier, determined analogously.

Even if α_r and α_s both represent bank loan rates, significant differences between these rates arise because of variety of factors. These include the firm's country or region, the industry it belongs to, the size of the credit line, the size of the firm (measured by its assets or gross revenues), the loan type, as well as its overall financial credit record and several financial ratios on its balance sheet and profit/loss statement¹. Several papers have estimated the relative importance of these factors

¹Bank rates for commercial loans and credit lines are determined as a spread with respect to a base rate, for example LIBOR (The London Interbank Offered Rate) or the U.S. prime rate. Firms are assigned one of a small number of possible risk ratings based on public or private bond ratings, the above mentioned financial ratios and its liquidity of the collateral provided (see, for instance, Koch (1995)). The dependency of bank rates on these characteristics is apparent in publicly available databases such as DealScan.

from large databases of bank loans, e.g., Berger and Udell (1990, 1995), Booth (1992), Petersen and Rajan (1994), Beim (1996) and Fernandez et al. (2008). While identifying many factors that explain major differences in bank loan rates, these studies conclude that the role of the borrower’s default risk is either insignificant or only of moderate importance².

In addition, all the above loan rate determinants represent characteristics of the firm’s past and global performance across all of its business units and product lines, as opposed to the specific product line considered here. Since in this paper we focus on firms with many markets and product lines, we treat, both under self- and bank-financing, the capital cost rates α_s and α_r as *exogenous* parameters (or exogenous stochastic processes, see §2.8), analogous to their treatment in almost all of the inventory and supply chain literature.

2.2.1 Independent Financing (IF)

Under IF, the retailer finances her inventories with a bank, or from her own cash reserves (self-financing), i.e., the supplier is paid immediately upon delivery of the procurement orders either by the bank or through self-financing. In the former case, IF is often implemented through factoring where the purchase invoice is owed to the bank, and the retailer draws from a credit line; in the latter case the purchasing invoice is paid from and owed to the firm’s cash reserves. The supplier charges a constant wholesale price w . It is optimal for the retailer to determine its procurement decisions in accordance with a base-stock policy, say with base-stock level y . (Under a base-stock policy, the inventory level is increased to the base-stock level in any period whose starting inventory is below that level.) Let

$s(y) = \mathbb{E} \min\{D, y\}$ = the expected retailer sales per period under a base-stock policy with base-stock level y ,

²Beim(1996), for example, states “Factors which are important in bond pricing, such as borrower risk and loan term, have only moderate importance in bank lending, while other factors such as borrower size and geographic location, lender identity, and pricing benchmark have unexpectedly high significance.”

$h(y) = \mathbb{E}(y - D)^+ =$ the expected end-of-the-period inventory level under a base-stock policy with base-stock level y .

$h(y)$ also denotes the expected number of units at the end of a period that have been received by the retailer but not yet paid for under the base-stock policy, so that $wh(y)$ represents the average per period amount of outstanding debt to the bank or the firm's own cash reserves, which is charged at a capital cost rate α_r , either by the bank or by drawing from the retailer's own cash reserves.

The profit functions of the retailer and the supplier are therefore given by:

$$\pi_r^{IF}(w, \beta, y) = (p - w)s(y) - (\alpha_r w + h_0)h(y), \quad \pi_s^{IF}(w, \beta, y) = (w - c)s(y)$$

2.2.2 Inventory Subsidies (IS)

Under IS, the retailer continues to finance her inventories with a bank, or from her own cash reserves. However, the supplier offers to subsidize the retailer's holding cost, specifically the capital cost component. Let β be the subsidy rate where $0 \leq \beta \leq \alpha_r$. Thus the effective capital cost rate of the retailer is $\alpha_r - \beta$. The profit functions of the retailer and the supplier are now given by:

$$\pi_r^{IS}(w, \beta, y) = (p - w)s(y) - [(\alpha_r - \beta)w + h_0]h(y), \quad \pi_s^{IS}(w, \beta, y) = (w - c)s(y) - \beta wh(y).$$

Note that IF is a special case of IS with $\beta = 0$.

2.2.3 Trade Credit (TC)

Under TC the supplier adopts the financing function, permitting the retailer to delay payments for purchased goods until the time of sale³. The supplier charges the retailer a given interest rate for each period during which payment for an item is outstanding. This per period interest rate may vary as a function of the item's shelf age⁴.

³In practice, the due date is often selected as a fixed calendar day, decoupled from the time of sale.

⁴For example, the supplier may charge a flat rate from the time of delivery, or he may offer an initial grace period of G periods without any interest charges, followed by a constant interest

To provide a fair comparison of IF and IS, we confine ourselves to the case where the supplier charges a flat interest rate $\alpha \leq \alpha_r$ irrespective of the item's shelf age. At the same time, we consider both the case where $\alpha \geq \alpha_s$ and the one where $\alpha < \alpha_s$, i.e., the supplier either incurs a net revenue or a net cost when extending trade credit to the retailer. As under IF and IS, it is easily verified that a base-stock policy continues to be optimal for the retailer⁵. Using y to denote the base-stock level, the expected amount of payables at the end of each period equals $wh(y)$, as before. This results in an expected per period cost $\alpha wh(y)$ to the retailer, and an equivalent revenue for the supplier. The supplier finances his working capital at a cost rate α_s , either by his bank or drawing from his cash reserves, therefore the net interest revenues for the supplier are given by $(\alpha - \alpha_s)wh(y)$. We conclude that the chain members' profit functions are now given by:

$$\pi_r^{TC}(w, \alpha, y) = (p-w)s(y) - (\alpha w + h_0)h(y), \quad \pi_s^{TC}(w, \alpha, y) = (w-c)s(y) + (\alpha - \alpha_s)wh(y).$$

2.2.4 A General Model

The IF, IS and TC models are special cases of the following general model.

$$\pi_r(w, \beta_g, y) = (p-w)s(y) - [(\bar{\alpha} - \beta_g)w + h_0]h(y) \quad (2.1)$$

$$\pi_s(w, \beta_g, y) = (w-c)s(y) - \beta_g wh(y) \quad (2.2)$$

Here β_g represents the supplier's effective capital cost rate (ECCR), a term of trade to be selected by him along with the wholesale price w . Moreover, the supplier's ECCR β_g is to be selected in a given interval $[\underline{\alpha}, \bar{\alpha}]$ with $\underline{\alpha} \leq \bar{\alpha}$. Note that $\bar{\alpha} - \beta_g$ may be interpreted as the retailer's ECCR.

rate α_1 , applied to any additional period. As a third example, the interest rate may be increased to a level $\alpha_2 > \alpha_1$ for units purchased more than G_2 periods ago. In general, as in Gupta and Wang (2009), let $\alpha(t)$ be the incremental interest rate charged for any unit with a shelf age of t periods, i.e., whose payment has been outstanding for t periods, and assume that $\alpha(\cdot)$ is an arbitrary non-decreasing function.

⁵This structural result continues to apply under general increasing shelf-age dependent interest rate function $\alpha(\cdot)$, see Gupta and Wang (2009) and Chapter 3. Chapter 3 also characterizes the optimal procurement policy under more general cost and revenue structures.

We obtain the IS model by setting $\beta_g = \beta$, $\underline{\alpha} = 0$ and $\bar{\alpha} = \alpha_r$. The IF model has the same parameters as IS except that $\beta_g = 0$. The TC model can be obtained by selecting $\beta_g = \alpha_s - \alpha$, $\underline{\alpha} = \alpha_s - \alpha_r$ and $\bar{\alpha} = \alpha_s$.

It is well known and easily verified that in the general model, the retailer's optimal base-stock level $y(w, \beta_g)$ in response to given trade terms (w, β_g) is given by a specific fractile of the demand distribution. More specifically, $y(w, \beta_g)$ is the fractile that satisfies

$$F(y) = 1 - \frac{(\bar{\alpha} - \beta_g)w + h_0}{p - w + (\bar{\alpha} - \beta_g)w + h_0}. \quad (2.3)$$

Note that $y(w, \beta_g)$ is decreasing in w and increasing in β_g .

Next in §2.3-2.4, we focus on the general model specified by the profit functions (2.1) and (2.2) and characterize the equilibrium behavior in three distinct Stackelberg games with the supplier as the leader and the retailer as the follower. In the first Stackelberg game the supplier selects the wholesale price under an exogenously given ECCR β_g ; In the second Stackelberg game the supplier selects his ECCR under a given wholesale price w , while in the third game the supplier controls both terms of trade, w and β_g .

2.3 The Stackelberg Game under a Given ECCR

Given the supplier's ECCR β_g , he chooses his optimal wholesale price on the interval $[0, p]$, taking into account the retailer's best response procurement strategy. (In some settings, the supplier may offer the product at a wholesale price below the variable cost rate while still realizing a profit on the basis of the finance charges.) This gives rise to the equilibrium profit function $\hat{\Pi}_s(w|\beta_g) \equiv \pi_s(w, \beta_g, y(w, \beta_g))$. Analysis of the Stackelberg game requires an understanding of the structure of the equilibrium profit function $\hat{\Pi}_s(\cdot|\beta_g)$. However, it is more convenient to express the supplier's equilibrium profit as a function of the retailer's base-stock level y , as opposed to his wholesale price. Since the retailer's optimal base-stock level $y(w, \beta_g)$ is strictly decreasing in w , we can write w as a function of y . Let $w(y)$ denote the

inverse function, i.e.,

$$w(y) \equiv \frac{p - (p + h_0)F(y)}{1 - (1 - \delta)F(y)}, \quad (2.4)$$

where $\delta = \bar{\alpha} - \beta_g \geq 0$. The supplier's equilibrium profit, as a function of the base-stock level y , is thus given by: $\Pi_s(y|\beta_g) = (w(y) - c)s(y) - \beta_g w(y)h(y)$. After determining the optimal targeted base stock level $y_{\beta_g}^*$, the corresponding wholesale price that induces this base stock level is immediate from (2.4).

The supplier's equilibrium profit as a function of y can be written as

$$\Pi_s(y|\beta_g) = w(y)(s(y) - \beta_g h(y)) - cs(y) = w(y)\xi(y) - cs(y). \quad (2.5)$$

where $\xi(y) \equiv s(y) - \beta_g h(y)$. In other words, if the supplier assumes part of the capital costs associated with the retailer's inventory, either in a TC or IS arrangement, this is equivalent to accepting a reduction of the expected sales volume per period $s(y)$ to the lower quantity $\xi(y)$, where the correction term is strictly proportional to the ECCR β_g . We therefore refer to $\xi(y)$ as the supplier's *effective expected sales volume*, in contrast to the *gross* expected sales quantity $s(y)$. In terms of the gross expected sales quantity, it is always in the supplier's interest to induce a higher base-stock level. Since $\xi'(y) = 1 - (1 + \beta_g)F(y)$, the effective expected sales volume increases with the base-stock level but *only* up to the $1/(1 + \beta_g)$ -th fractile of the demand distribution, i.e.,

$$\xi'(y) = 1 - (1 + \beta_g)F(y) > 0 \text{ if and only if } y < y_s \equiv F^{-1}\left(\frac{1}{1 + \beta_g}\right). \quad (2.6)$$

(In case $\beta_g = 0$, define $y_s \equiv \infty$.) Clearly, the supplier has no interest in targeting a base-stock level above this critical fractile y_s : for any $y > y_s$, either the effective expected sales volume itself is negative, resulting in negative profits, or the marginal profit $\Pi'(y|\beta_g) < 0$ since $w'(y) < 0$, $\xi'(y) < 0$ and $s'(y) > 0$. In other words, the critical fractile y_s is a natural bound for the targeted base-stock level.

Similarly, a second natural bound for the targeted base-stock level is given by the $1/(1 + h_0/p)$ -fractile, that is $F^{-1}(1/(1 + h_0/p))$. To verify this, note from (2.3) that $F(y)$ is a decreasing function of w , so that $F(y)$ is bounded from above by the expression for the right-hand side of (2.3), obtained when $w = 0$. Combining the

two upper bounds, we specify, without loss of generality, that

$$y_{max} \equiv \min \left\{ y_s, F^{-1} \left(\frac{1}{1 + h_0/p} \right) \right\} = F^{-1} \left(\frac{1}{1 + \max\{h_0/p, \beta_g\}} \right). \quad (2.7)$$

(As above, in case $\beta_g = h_0 = 0$, define $y_{max} \equiv \infty$.) We now show the supplier's equilibrium profit is quasi-concave both as a function of the base-stock level and as a function of the wholesale price.

Theorem 2.1 (Stackelberg game under given ECCR) *Assume the demand distribution is IFR. (a) The supplier's equilibrium profit $\Pi_s(\cdot|\beta_g)$ is quasi-concave on the full base-stock level range $[y_{min} = 0, y_{max}]$, achieving its maximum at a unique interior point $y_{\beta_g}^*$.*

(b) The supplier's equilibrium profit $\hat{\Pi}_s(\cdot|\beta_g)$, viewed as a function of w , is quasi-concave on the wholesale price range $[0, p]$, achieving its maximum at a unique interior point $w_{\beta_g}^ = w(y_{\beta_g}^*)$, and with $\lim_{w \uparrow p} \hat{\Pi}_s(w|\beta_g) = 0$.*

Anupindi and Bassok (1999) address this Stackelberg game in the special case of IF arrangements with $h_0 = 0$. The authors state that they are unable to determine if the supplier's equilibrium profit function is "convex, concave, or even unimodal in the wholesale price". This contrasts Lariviere and Porteus (2001), dealing with the *single period* Stackelberg game under IF, who show that the supplier's equilibrium profit function is quasi-concave under a generalization of the IFR condition. (Indeed, the structure of the profit functions in a single period model is essentially simpler.) Anupindi and Bassok (1999) obtain a characterization of the structure of the supplier's equilibrium profit function only for Normal demand distributions and under Nahmias (1993)' approximation of the inverse of the standard Normal cdf by a difference of power functions. Returning to single period models, Kouvelis and Zhao (2009) address a variant of the Lariviere and Porteus (2001) model with a bank loan rate endogenously determined so that the bank breaks even in expectation, considering the possibility of default. The authors show that in this case the Stackelberg game reduces to that of Lariviere and Porteus (2001) with the capital cost rate given by the bank's cost of funds. However, in the same model under supplier financing, the authors show that the supplier's equilibrium profit function

fails to be quasi-concave even under special cases of the class of IFR distribution, for example, when the failure rate is required to be convex as well as increasing; see Zhou and Groenevelt (2008) for a similar characterization in their model.

Finally, it is of interest to investigate how the optimal wholesale price depends on the (exogenously given) ECCR. Based on extensive numerical studies, we conjecture that the optimal wholesale price is increasing in the supplier's ECCR.

2.4 The Remaining Two Stackelberg Games

We first assume that the wholesale price is exogenously given and the supplier is able to select his ECCR value β_g . We characterize the equilibrium solution of this Stackelberg game in §2.4.1. In §2.4.2, we show how the equilibrium performance measures depend on the wholesale price and characterize the solution of the full Stackelberg game.

2.4.1 The Stackelberg Game under Given Wholesale Price

It is, again, useful to represent the supplier's equilibrium profits as a function of the targeted base-stock level (as opposed to the selected ECCR). It follows from (2.3) that

$$\beta_g(w, y) \equiv \bar{\alpha} - \left(\frac{p}{w} - 1\right) \left(\frac{1}{F(y)} - 1\right) + \frac{h_0}{w}. \quad (2.8)$$

Substituting (2.8) into (2.2), we obtain the desired representation of the supplier's equilibrium profits:

$$\Pi_s(y|w) = (w - c)s(y) - \left(w\bar{\alpha} - (p - w) \left(\frac{1}{F(y)} - 1\right) + h_0\right)h(y). \quad (2.9)$$

Since the ECCR must be selected in the interval $[\underline{\alpha}, \bar{\alpha}]$, the targeted base-stock level y satisfies the bounds, $\underline{y} \leq y \leq \bar{y}$, again specified as fractiles of the demand distribution:

$$\underline{y} \equiv F^{-1} \left(\frac{p - w}{p - w + w(\bar{\alpha} - \underline{\alpha}) + h_0} \right) \text{ and } \bar{y} \equiv F^{-1} \left(\frac{p - w}{p - w + h_0} \right). \quad (2.10)$$

We now show that the supplier's equilibrium profit function is quasi-concave on the relevant range $[\underline{y}, \bar{y}]$, similar to the corresponding function in the Stackelberg game

of the previous section. The result requires a variant of the IFR condition, which we refer to as the *Modified Increasing Failure Rate (MIFR)* property:

There exists a number L such that $\zeta(y) \equiv \frac{\int_0^y F(u)du}{F^2(y)} \frac{f(y)}{1-F(y)}$ is increasing for $y \geq L$.

The following Lemma identifies three important classes of distributions that satisfy the MIFR property for an appropriate value of L . (For some important distributions $\zeta(y)$ fails to be monotone on its complete support.)

Lemma 2.1 (a) *If a distribution is IFR with a non-increasing density function on the positive half line, it has the MIFR property with $L = 0$. Examples are the uniform and exponential distributions.*

(b) *All normal distributions have the IFR and MIFR property with $L = \mu - 1.8\sigma$.*

In summary, the MIFR condition is a variant of the traditional IFR property, satisfied on the *complete* support of exponential and uniform distributions. More generally, let

$$\bar{w} \equiv \sup\{w \leq p : \underline{y}(w) \geq L\}. \quad (2.11)$$

Since $\underline{y}(\cdot)$ is a decreasing function, the MIFR property is satisfied on the full feasible range $[\underline{y}(w), \bar{y}(w)]$ if and only if $w \leq \bar{w}$. We assume that \bar{w} is well defined, which, by the monotonicity property of $\underline{y}(\cdot)$, is equivalent to $L \leq F^{-1}(\frac{p}{p+h_0}) = \underline{y}(0)$. When $L = 0$ (, e.g., exponential and uniform distributions), $\bar{w} = p$. For Normals, $F(L) = F(\mu - 1.8\sigma) = 0.036$, so that, by (11), $\bar{w} = \frac{p(1-F(L))-F(L)h_0}{1-F(L)+F(L)(\bar{\alpha}-\underline{\alpha})} = \frac{p-0.037h_0}{1+0.037(\bar{\alpha}-\underline{\alpha})}$. In other words, either the MIFR property holds on the complete feasible range $[0, p]$, or it does so unless the retailer's profit margin is exceedingly low, in the Normal case resulting in a fill rate of 3.6%!

The marginal equilibrium profit function for the supplier is given by:

$$\begin{aligned} \Pi'_s(y|w) &= (w-c)(1-F(y)) - (w\bar{\alpha} + h_0 - (p-w)(\frac{1}{F(y)} - 1))F(y) - (p-w)\frac{h(y)f(y)}{F^2(y)} \\ &= (p-c)(1-F(y)) - (w\bar{\alpha} + h_0)F(y) - (p-w)\frac{h(y)f(y)}{F^2(y)} \end{aligned} \quad (2.12)$$

Let $y_0 = \inf\{L \leq y : \Pi'_s(y|w) \leq 0\}$ denote the smallest base-stock level for which the supplier's marginal equilibrium profit is negative. ($L \leq y_0 < \infty$ since

$\lim_{y \rightarrow \infty} \Pi'_s(y|w) = -(w\bar{\alpha} + h_0) - (p - w) \lim_{y \rightarrow \infty} h(y)f(y) \leq -(w\bar{\alpha} + h_0)$.) In Theorem 2.2, we show that the optimal base-stock level $y_w^* \equiv \operatorname{argmax}_{\underline{y} \leq y \leq \bar{y}} \Pi_s(y|w)$ and the associated optimal ECCR $\beta_g^*(w) \equiv \beta_g(w, y_w^*)$ are obtained by the projection of y_0 onto $[\underline{y}, \bar{y}]$, i.e., by determining the point in $[\underline{y}, \bar{y}]$ which is closest to y_0 .

Theorem 2.2 (Stackelberg game under given wholesale price) *Fix w and assume the demand distribution is MIFR on the feasible range $[\underline{y}, \bar{y}]$.*

(a) *The supplier's equilibrium profit function $\Pi_s(\cdot|w)$ is quasi-concave, with y_0 as its unique maximum on $[L, \infty)$.*

(b) *If $y_0 < \underline{y}$, $y_w^* = \underline{y}$ and $\beta_g^*(w) = \beta_g(w, \underline{y}) = \underline{\alpha}$.*

If $\underline{y} \leq y_0 \leq \bar{y}$, $y_w^ = y_0$ and $\underline{\alpha} < \beta_g^*(w) = \beta_g(w, y_w^*) < \bar{\alpha}$.*

If $y_0 > \bar{y}$, $y_w^ = \bar{y}$ and $\beta_g^*(w) = \beta_g(w, \bar{y}) = \bar{\alpha}$.*

(c) *The supplier's profit $\hat{\Pi}_s(\cdot|w)$, viewed as a function of β_g , is quasi-concave on $[\underline{\alpha}, \bar{\alpha}]$.*

As mentioned in §1.2, Anupindi and Bassok (1999) address this Stackelberg game as well, under the *assumption* that the supplier's equilibrium profit function $\hat{\Pi}_s(\cdot|w)$ is quasi-concave. (The authors verified numerically that the assumption holds for Normal demands when substituting their cdf with the above approximation in Nahmias (1993).) We prove that the *exact* function is quasi-concave under all Normal demand distributions, as well as under the much broader MIFR class. We are not aware of any other Stackelberg game models in which the supplier selects a credit rate or inventory subsidy, under a given wholesale price.

2.4.2 Comparative Statics and the Full Stackelberg Game

We now characterize how the optimal base-stock level, the associated ECCR and the supplier's profit value vary with the wholesale price. We start with a characterization of the impact the wholesale price has on the *unconstrained* optimal base-stock level $y_0(w) = \operatorname{argmax}_{y \geq L} \Pi_s(y|w)$. Rewrite (2.12) as:

$$\Pi'_s(y|w) = (p - c) - (p - c + p\bar{\alpha} + h_0)F(y) + (p - w)F(y)(\bar{\alpha} - \kappa(y)) \quad (2.13)$$

where $\kappa(y) \equiv \frac{h(y)f(y)}{F^3(y)}$. In general, $y_0(w)$ cannot be obtained in closed form except when the wholesale price $w = p$. It is easily verified from (2.13) that $y_0(p) = y^p$ where

$$y^p \equiv \max\left\{L, F^{-1}\left(\frac{p-c}{p-c+p\bar{\alpha}+h_0}\right)\right\}. \quad (2.14)$$

The following theorem shows that $y_0(\cdot)$ always varies monotonically with the wholesale price, approaching $y_0(p)$ as $w \uparrow p$. Interestingly, this *unconstrained* optimal base-stock level may be increasing in the wholesale price; indeed, as substantiated by Proposition 2.1 below, this is the prevalent case for most common demand distributions.

Theorem 2.3 *Assume the demand distribution is MIFR. (a) $y_0(\cdot)$ is a continuous function.*

(b) Exactly one of the following four scenarios arises:

(b-i) $L < y^p$, $\kappa(y^p) > \bar{\alpha}$ and the function $y_0(\cdot)$ increases on $[0, p]$ until reaching the level $y^p = y_0(p)$.

(b-ii) $L < y^p$, $\kappa(y^p) = \bar{\alpha}$ and $y_0(w) = y^p$ for all $w \in [0, p]$.

(b-iii) $L < y^p$, $\kappa(y^p) < \bar{\alpha}$ and $y_0(\cdot)$ decreases on $[0, p]$ until reaching the level $y^p = y_0(p)$.

(b-iv) $L = y^p$ and $y_0(\cdot)$ decreases on $[0, p]$ until it reaches the level $L = y^p$.

Observe that it is possible to determine which of the four patterns (b-i) - (b-iv) applies on the basis of a single comparison of the value $\kappa(y^p)$ and $\bar{\alpha}$ (, unless $L = y^p$ in which case pattern (b-iv) applies unequivocally).

While Theorem 2.3 allows for the unconstrained optimal base-stock level to be either increasing or decreasing in the wholesale price, the former case prevails for almost all MIFR distributions. This counterintuitive result is explained as follows. An *increase* in the wholesale price impacts the retailer's base-stock level/sales in two ways. As long as the inventory cost subsidy/interest rate reduction remains the same, the retailer's base-stock level/sales decreases as the wholesale price increases. We refer to this as a *direct negative effect*. On the other hand, an increase in the wholesale price induces an increase in the inventory cost subsidy/interest rate

reduction, which results in an *increase* in the retailer's base-stock level/sales. This is an *indirect positive effect*. Proposition 2.1 shows that for most MIFR distributions the indirect positive effect dominates the direct negative effect, i.e., an *increase* in the wholesale price results in an *increase* in the retailer's base-stock level.

Proposition 2.1 (a) *If the demand distribution is uniform, $y_0(\cdot)$ increases on $[0, p]$.*

(b) *If the demand distribution is exponential, $p \leq 4c$ and $\bar{\alpha} \leq 0.77$, $y_0(\cdot)$ increases on $[0, p]$.*

(c) *If the demand distribution is Normal, $y_0(\cdot)$ increases on $[0, p]$.*

We are now ready to characterize how the wholesale price impacts the supplier's equilibrium profits as well as the *constrained* base-stock level $y_w^* = \underset{\underline{y}(w) \leq y \leq \bar{y}(w)}{\operatorname{argmax}} \Pi_s(y|w)$ and his chosen ECCR. For any continuous function $H(w)$, we define w^0 to be a mode or a local maximum if $H(w^0) \geq H(w)$ on some interval including w^0 , with strict inequality to the left or right of w^0 .

Theorem 2.4 (Comparative Statics and Full Stackelberg Game) *Assume the demand distribution has both the IFR and MIFR properties on the feasible range $[\underline{y}(w), \bar{y}(w)]$ for $w \in [0, \bar{w}]$.*

The supplier's equilibrium profits

(a1) *The supplier's equilibrium profit function $\Pi_s^*(\cdot)$ is continuously differentiable.*

(a2) *The supplier's equilibrium profit function $\Pi_s^*(\cdot)$ is unimodal or bimodal where the two potential modes are $w_{\bar{\alpha}}^*$, $w_{\underline{\alpha}}^*$, or \bar{w} . Moreover, when $\bar{w} = p$, the two potential modes are $w_{\bar{\alpha}}^*$ and $w_{\underline{\alpha}}^*$.*

The optimal base-stock level and ECCR

(b1) *Assume $y_0(\cdot)$ is increasing, i.e., pattern (b-i) or (b-ii) in Theorem 2.3 applies.*

There exist two critical wholesale prices $0 < w_1 \leq w_2 \leq \bar{w}$, such that

(1) *The optimal base-stock level y_w^* equals \underline{y} when $w \leq w_1$, \bar{y} when $w < w_2$ and increases continuously from $\underline{y}(w_1)$ to $\bar{y}(w_2)$ when $w_1 < w \leq w_2$.*

(2) *The supplier's optimal effective capital rate β_g^* equals $\underline{\alpha}$ when $w \leq w_1$, $\bar{\alpha}$ when $w > w_2$ and increases continuously from $\underline{\alpha}$ to $\bar{\alpha}$ when $w_1 < w \leq w_2$.*

(b2) Assume $y_0(\cdot)$ is decreasing, i.e., pattern (b-iii) or (b-iv) in Theorem 2.3 applies. The optimal base stock level y_w^* is continuously decreasing on $[0, \bar{w}]$.

The full Stackelberg Game

(c) The optimal solution of the full Stackelberg game is $(w_{\bar{\alpha}}^*, \bar{\alpha})$, $(w_{\underline{\alpha}}^*, \underline{\alpha})$, or $(\bar{w}, \beta_g^*(\bar{w}))$. Moreover, when $\bar{w} = p$, it is either $(w_{\bar{\alpha}}^*, \bar{\alpha})$ or $(w_{\underline{\alpha}}^*, \underline{\alpha})$.

The monotonicity results in part (b1) for the prevalent case where $y_0(\cdot)$ is increasing are exhibited graphically in Figure 2.1 below.

Figure 2.1: Monotonicity of the optimal ECCR and base-stock level when $y_0(\cdot)$ is increasing.

Note: Demand follows a Normal distribution with $\mu = 100$ and $\mu/\sigma = 3$. $p = 130$, $c = 100$, $h = 5$, $\alpha_r = 0.1$, $\alpha_s = 0.03$.

In view of part (c), to solve the full Stackelberg game, it suffices to compute $w_{\bar{\alpha}}^*$ and $w_{\underline{\alpha}}^*$ as the unique maximum of the quasi-concave functions $\hat{\Pi}_s(w|\beta_g = \bar{\alpha})$ and $\hat{\Pi}_s(w|\beta_g = \underline{\alpha})$ respectively, as well as $\beta_g^*(\bar{w})$ as the unique maximum of the quasi-concave function $\hat{\Pi}_s(\beta_g|w = \bar{w})$ when $\bar{w} < p$.

Theorem 2.4 exhibits the possibility of the supplier's equilibrium profit function $\Pi_s^*(\cdot)$ being bimodal. Indeed, we have observed such instances, with a Normal demand distribution, however, only when the coefficient of variation is high ($\frac{\sigma}{\mu} \geq 1$) and the chain-wide profit margin $\frac{p-c}{c}$ is small (≤ 0.1). These instances combine a high degree of demand risk with low potential to achieve variable profit margins, thus enticing the supplier to offer a low wholesale price under which only a minimal amount of supplier financing can be justified, this as an alternative to the usual mode

$(w_{\bar{\alpha}}^*, \bar{\alpha})$ representing a maximal degree of supplier financing and a commensurately higher wholesale price. For the three classes of distributions discussed above (, see Proposition 2.1), the following proposition identifies generally satisfied conditions under which $\Pi_s^*(\cdot)$ is unimodal with a unique local maximum at $w = w_{\bar{\alpha}}^*$, or $w = \bar{w}$.

Proposition 2.2 *In the full Stackelberg game, the optimal solution is*

(a-UNI) $(w_{\bar{\alpha}}^*, \bar{\alpha})$ when the demand distribution is uniform.

(b-EXP) $(w_{\bar{\alpha}}^*, \bar{\alpha})$ when the demand distribution is exponential and the condition in Proposition 2.1(b) is satisfied, i.e., $\frac{p-c}{c} \leq 3$ and $\bar{\alpha} \leq 0.77$.

(c-NORM) either $(w_{\bar{\alpha}}^*, \bar{\alpha})$ or $(\bar{w}, \beta_g^*(\bar{w}))$ when the demand distribution is Normal with coefficient of variation $\frac{\sigma}{\mu} \leq 0.45$, the profit margin $\frac{p-c}{c} \geq 7.7\%$ and $p \max\{\alpha_r, \alpha_s\} + h_0 \leq p$.

The conclusions in Proposition 2.2 are, thus, guaranteed for uniform distributions. For exponential and Normals they hold almost invariably: single period interest rate almost never exceeds 77%, end-to-end profit margins are, invariably, above 7.7% and below 300%, while the single period total inventory carrying cost of an item is, typically, below its retail price. Normal distributions are generally not employed when the coefficient of variation is above 0.5; the upper bound in part (c) is 0.45, and, therefore, is hardly restrictive. We confine ourselves, henceforth, to the above three classes of problem instances addressed in Proposition 2.2 (, except where results hold unconditionally):

(C): The problem instance belongs to the class (UNI), (EXP) or (NORM).

As mentioned in §1.2 and shown in §2.6, we show that the relative magnitude of various performance measures of interest across different supplier financing mechanisms, IF, IS and TC, depend largely on the capital cost rates, α_r and α_s . We therefore complete *this* section with characterizations of how these capital cost rates impact on the full Stackelberg game equilibrium.

Proposition 2.3 (a) *Under TC, there exists a critical value $0 \leq \alpha_s^c \leq \infty$, such that, as α_s increases up to α_s^c , w^* increases with α_s while the retailer's trade credit*

interest rate $\alpha^* = 0$; as α_s increases beyond α_s^c , w^* is kept constant at the level \bar{w} and α^* is increasing in α_s .

(b) Under IS, there exist a critical value $0 \leq \alpha_r^c \leq \infty$, such that, as α_r increases up to α_r^c , the supplier provides the maximum subsidy $\beta^* = \alpha_r$ while increasing the wholesale price w^* ; as α_r increases beyond α_r^c , the supplier maintains the maximum wholesale price \bar{w} while increasing retailer's effective capital cost rate $(\alpha_r - \beta^*)$.

2.5 Summary of the Three Games

In the Stackelberg game that arises under a given supplier's ECCR β_g , we have shown that, as long as the demand distribution is IFR, the supplier's equilibrium profit varies as a quasi-concave function of the targeted base-stock level, or, as a quasi-concave function of the selected wholesale price. The optimal wholesale price is an interior point of the feasible price range.

Conversely, assume that the wholesale price is exogenously given and the supplier selects his ECCR. Under a slight variant of the IFR condition (MIFR), the supplier's equilibrium profit, again, first increases with the selected ECCR and, after reaching a peak, declines ever thereafter. How does the wholesale price impacts on the various performance measures? A key determinant is how, in the absence of any ECCR bounds, the *unconstrained* equilibrium base-stock level $y_0(\cdot)$ varies with the wholesale price. Here, we show that $y_0(\cdot)$ is either continuously increasing or decreasing, with a single closed form test determining which case prevails. Surprisingly, the former case is guaranteed under many classes of (MIFR) demand distributions (, e.g., Normals, exponentials and uniform distributions), and this is explained by the fact that the *direct* negative effect associated with a higher purchase price is dominated by the improved financing terms offered by the supplier in response to a higher wholesale price. When $y_0(\cdot)$ is increasing, the wholesale price interval can be partitioned into three parts: in the lower range, the optimal ECCR equals to the lowest possible value $\underline{\alpha}$; In the middle range it continuously increases to $\bar{\alpha}$ and stays there in the upper range. The equilibrium base-stock level and hence

the associated expected sales volume decrease with the wholesale price in the lower and upper range, but increase with the latter in the middle range. It is only in the far rarer case where $y_0(\cdot)$ is decreasing that the equilibrium base-stock level and the expected sales volume decrease monotonically with the wholesale price.

As far as the supplier's expected profit is concerned, it is, in general, a unimodal or bimodal function of w , where the potential mode(s) are $w_{\underline{\alpha}}^*$ or $w_{\overline{\alpha}}^*$ (corresponding to the ECCR equaling the lowest or highest possible value), or the highest feasible wholesale price \bar{w} . This also characterizes the solution of the full Stackelberg game. For the three classes of distributions discussed above, we show that under minor parameter conditions, only the price $w_{\overline{\alpha}}$ or \bar{w} solves the full Stackelberg game.

2.6 Comparison of Different Financing Mechanisms

We systematically compare the equilibrium results of various performance measures across the different financing mechanisms. In §2.6.1, we compare the IS and TC arrangements. In §2.6.2, we assume that the chain initially operates under IF, i.e., without any supplier financing, and investigate whether and when both chain members are better off when switching to TC or IS. Beyond the many theoretic rankings obtained in §2.6.1 and §2.6.2, reported in §2.6.3 is a numerical study identifying many other such rankings.

2.6.1 Comparing IS and TC

This subsection is devoted to the comparison of the supply chain performance under the two supplier financing schemes IS and TC. In the remainder of the section, we append a superscript “TC” or “IS” to any one of the parameters or performance measures. We first show the supplier's preference for the TC or the IS mechanism hinges entirely on a comparison of his capital cost rate with that of the retailer. Indeed, unlike the characterization of the equilibria in the various Stackelberg games, these comparison results apply under *any* demand distribution.

Theorem 2.5 (The supplier's profits) (a) For any wholesale price $w \in [0, p]$, the supplier prefers TC (IS) if $\alpha_s \leq (>) \alpha_r$.

(b) In the full Stackelberg game where the supplier chooses both the wholesale price and the trade credit interest/inventory subsidy, the supplier prefers TC (IS) if $\alpha_s \leq (>) \alpha_r$.

As far as ranking the retailer's profits and other performance measures is concerned, we show that, at least under condition (C) these rankings, too, depend solely on the relationship of the capital cost rates. Moreover, the retailer's preference is perfectly aligned with that of the supplier. Recall that the TC and IS models employ different parameters: $\underline{\alpha}^{IS} = 0$ and $\underline{\alpha}^{TC} = \alpha_s - \alpha_r$; $\bar{\alpha}^{IS} = \alpha_r$ and $\bar{\alpha}^{TC} = \alpha_s$. Note $\bar{\alpha}^{IS} - \underline{\alpha}^{IS} = \bar{\alpha}^{TC} - \underline{\alpha}^{TC} = \alpha_r$. Thus, by (2.10), \underline{y} and \bar{y} are the same under both TC and IS, and so is \bar{w} by (2.11).

Theorem 2.6 (The remaining performance measures) Assume condition (C).

For any given wholesale price $w \in [0, \bar{w}]$, if $\alpha_s \leq (>) \alpha_r$,

(A-1) The retailer prefers TC (IS);

(A-2) The optimal base-stock level and the expected sales volume are higher under TC (IS).

In the full Stackelberg game, if $\alpha_s \leq (>) \alpha_r$,

(B-1) The retailer prefers TC (IS).

(B-2) The optimal wholesale price is lower under TC (IS).

(B-3) The optimal base-stock level and the expected sales volume are higher under TC (IS).

2.6.2 Comparing the Two Supplier Financing Mechanisms with IF

Under any *given* wholesale price, both chain members are clearly better off going from IF to the Stackelberg equilibrium IS arrangement: under IS, the supplier has an additional term of trade to choose from and any selected subsidy rate results

in a point-wise larger profit function for the retailer. By Theorem 2.5 and 2.6(a), additional profit improvements are enjoyed by both when adopting the equilibrium TC arrangement, as long as $\alpha_s < \alpha_r$. We summarize:

Proposition 2.4 *For any $w \in [0, \bar{w}]$,*

- (i) *both chain members are better off under the Stackelberg equilibrium IS arrangement as opposed to IF;*
- (ii) *assuming $\alpha_s < \alpha_r$, the supplier gains even more benefit when switching to the Stackelberg equilibrium TC arrangement. The same is true for the retailer under condition (C).*

In other words, Proposition 2.4 shows that, under a *given* wholesale price, both chain members are better off under the Stackelberg equilibrium IS arrangement. This result holds for arbitrary demand distributions. When the supplier's cost of capital is lower than that of the retailer, the equilibrium TC arrangement results in even greater benefit for both chain members, albeit that the retailer's additional benefit has only been shown under Condition (C).

However, if $\alpha_s > \alpha_r$ and the choice is restricted to the IF and TC arrangements, *both* chain members may prefer IF. As an example, assume the wholesale price w is selected in the interval $[0, w_1^{IS}]$. Since $\alpha_s > \alpha_r$, by Theorems 2.5 and 2.6, we know that both chain members earn a lower profit under TC as opposed to IS under which arrangement their profits equal those under IF, since $\beta_g^{*IS}(w) = \underline{\alpha}^{IS} = 0$ for $w \in [0, w_1^{IS}]$.

Now suppose that, after switching to one of the two supplier financing mechanisms, the supplier could deviate from the wholesale price under IF and choose the one that yields the maximal profit for him, i.e., the optimal wholesale price in the full Stackelberg game, w^* . It follows immediately from the above Proposition that the supplier always prefers IS over IF, and, by Theorem 2.5, prefers TC even more, the latter as long as $\alpha_s < \alpha_r$. However, if the supplier may change the wholesale price, the retailer may not necessarily be better off under either of the supplier

financing schemes. Indeed, in our extensive numerical study, we found that the supplier's ability to select the wholesale price in an unrestricted manner almost always reduces the retailer's profit below its level under IF.

In order to induce the retailer to accept either IS or TC as a risk-sharing mechanism, it is thus more realistic to respecify the full Stackelberg game with a participation constraint that ensures the retailer's profit is no less than Π_r^{*IF} . Refer to Appendix A.2 for an efficient algorithm to solve *this* version of full Stackelberg game.

2.6.3 Numerical Study: Additional Comparison Results

Theorems 2.5 and 2.6 and Proposition 2.4 establish a large number of comparison results among important equilibrium performance measures in the full Stackelberg game under the IF-, IS- and TC- arrangements. We have focused on the following five performance measures: (i) the supplier's expected profit, (ii) the retailer's expected profit, (iii) the expected chain-wide profit, (iv) the equilibrium wholesale price and (v) the expected sales volume. Indeed, beyond the above mentioned theoretical results, a complete ranking of *almost all* of the five performance measures under the three financing schemes appears to prevail, depending only on whether the supplier's cost of capital is higher or lower than that of the retailer. These rankings have been identified and verified in a numerical study including 1080 problem instances and are reported in Table 1.1 in §1.2. In this numerical study, we use Normal and Exponential demand distributions. Note that, without loss of generality, one of the cost or revenue parameters can be chosen arbitrarily, for example, the unit price p . Similarly, under a Normal demand distribution with a given coefficient of variation σ/μ , both the retailer's and the supplier's equilibrium profit function are proportional with μ , so that, again without loss of generality, the latter can be chosen at an arbitrary level as well. For exponential demand distributions, the coefficient of variation is, of course, equal to one, but the equilibrium profit values fail to be proportional with the mean demand.

The various parameters are selected from the following lists:

$$p = 50, \quad h_0/p \in \{0.1\%, 1\%\}, \quad c/p \in \{50\%, 70\%, 90\%\},$$

$$\alpha_s \in \{0.5\%, 1\%, 3\%, 6\%, 9\%, 12\%\}, \quad \alpha_r - \alpha_s = \{-6\%, -3\%, -1\%, 1\%, 3\%, 6\%\}.$$

(Only instances with non-negative pairs (α_r, α_s) have been considered.)

For Normal demand distributions, the mean $\mu = 100$ and $\mu/\sigma \in \{2, 3, 4\}$. For Exponential demand distributions, the mean is selected from the list $\{50, 100, 150\}$.

We highlight a few of the perhaps unexpected rankings: the supplier financing schemes result in higher wholesale prices; nevertheless, they also result in a higher expected sales volume. This means that the improved credit terms offered under the supplier financing schemes provide a sales stimulus which exceeds the restrictive effect resulting from an increased wholesale price. Second, while the retailer's profit in the unconstrained full Stackelberg game is higher under IF as opposed to either of the supplier financing schemes (TC or IS), the *aggregate* profits in the chain are always *lower* under IF. This implies that both chain members can be made to benefit from a switch to one of the supplier financing schemes, as under the constrained full Stackelberg game described in §2.6.2 and Appendix A.2, with a participation constraint ensuring the retailer a profit level at least equal to her equilibrium profit under IF.

2.7 Generalizations with Default Risk

In this section, we discuss two generalizations of the base model that explicitly account for the possibility of the retailer defaulting. For the sake of brevity, we confine ourselves to the setting where, under IS, the retailer finances her inventory while loans from a third-party bank. The case of self-financing can be handled analogously. In the first default risk model, we assume that the supplier is exposed to default risks under a TC arrangement while the bank, under IS, has senior credit status or has secured its loans, alternatively, with adequate collateral. In the second default risk model, both the supplier and the bank are exposed to the same default risks.

The financial literature has developed two types of models for a firm's defaults

and reorganizations: *structural* models posit a stochastic process for the firm's aggregate cash flows, with defaults occurring when the wealth position of the firm falls below a critical threshold. This stochastic process is an exogenous primitive of the model, often of a standardized type, e.g., a diffusion process, which is independent of any of the firm's operational strategies. If the retailer's business activities were confined to the single item at the single market considered here, the equilibrium cash flow process resulting from the Stackelberg games could, in principle, be used as an endogenously determined alternative. However, most firms sell many products in multiple markets, and a structural model would continue to require an up-front assumption of a stochastic process for the firm's aggregate cash flows.

The second type of credit risk models is usually referred to as "*statistical*" or "*reduced form*" (RF). RF models assume that defaults are not directly based on the firm's cash flows. Instead, defaults occur at a given rate or intensity, either deterministically specified or randomly fluctuating as a function of an underlying state variable, for example the firm's ratings by standard credit agencies and general industry and economic indices. The state variable ω is usually assumed to fluctuate according to a Markov chain, see, e.g., chapters 5 and 6 in Lando (2004). In practice, RF models are used more frequently than their structural counterparts and typically generate more accurate predictions, see, for instance, Leland (2006, p9) .

We adopt a similar RF model in which relatively long periods of solvency are interrupted by defaults resulting in a reorganization prompted, in the U.S., for example, by a Chapter 11 filing. While representing the firm as alternating between intervals of solvency and reorganization, a realistic choice of model parameters should reflect the fact that for an A-rated firm, the average probability of a single default in a 10-year period, is no more than 1.65%, and for a Baa-rated firm 4.56%, see Leland (2009, p29). Thus even for a Baa-rated firm, the likelihood of two defaults occurring in a 20-year period is approximately 0.2%. The state of the firm is described by a triple (r, l, ω) where $r = 1(0)$ indicates that the firm is solvent (in reorganization), l denotes the number of periods since the beginning of the current solvency (reorganization) interval and ω the above mentioned world state variable.

The firm's state (r, l, ω) evolves according to some irreducible Markov chain with transitions from $S_1 \equiv \{(1, l, \omega)\}$ into $S_0 \equiv \{(0, l, \omega)\}$ representing defaults and those from S_0 to S_1 the completion of a reorganization. Let $\pi(1, l, \omega)$ denote the steady-state probability of being in state $(1, l, \omega)$ and $\theta(l, \omega)$ the likelihood of transitioning into S_0 , i.e, of a default, from this state. A default arises after a random interval of distress Δ in which the retailer's payments to some or all its creditors' are deferred. The distress interval is the last part of the solvency interval preceding the current default. During the reorganization, the creditors with less than senior status receive a random recovery rate ρ of the payables that are outstanding at its start. The joint distribution of (Δ, ρ) may depend on the firm's state immediately preceding the default.

2.7.1 Model with Default Risks for the Supplier

As mentioned, we first consider a model where, under TC, the supplier is exposed to default risks, but payments to the bank under IS are secured by proper collateral or because the bank retains senior creditor status. Thus, under IS and IF, the profits of the supplier and the retailer remain unaltered. As to TC, the optimal replenishment policy for the retailer is a state-dependent base-stock policy. However, we assume the retailer confines herself to simple base-stock policies only. We charge the expected losses due to partial recovery of outstanding payables in a distress interval to its last period(, after which the default occurs). If the default occurs following a period in which the firm is in state $(1, l, \omega)$, this expected loss equals $\mathbb{E}(\Delta(1 - \rho)|s = (1, l, \omega))$. This implies that the unconditional expected loss fraction of any payable amount to the supplier equals:

$$\gamma = \sum_{(l, \omega)} \pi(1, l, \omega) \theta(l, \omega) \mathbb{E}(\Delta(1 - \rho)|s = (1, l, \omega)). \quad (2.15)$$

(It is easily verified that $0 < \gamma < 1$. For A- or Baa-rated firms, the value of γ is less than 0.001: for a Baa-rated company, the one-year default probability is, on average, 0.14%, see Leland (2009, p29), while the median reorganization duration is approximately 1.5 years, see Denis and Rodgers (2007). Assume, further, that

the typical distress interval resulting in a default extends to 6 months, and the recovery rate $\rho = 0.5$. Employing a simple two state Markov chain to describe the firm's dynamics and fitting in transition probabilities to these statistics, we conclude that under these parameter assumptions, $\gamma = 0.0004$.) We obtain that the profit functions under TC are given by:

$$\begin{aligned}\pi_r^{TC}(w, \alpha, y) &= (p - w(1 - \gamma))s(y) - (\alpha w(1 - \gamma) + h_0)h(y), \\ \pi_s^{TC}(w, \alpha, y) &= (w(1 - \gamma) - c)s(y) + (\alpha(1 - \gamma) - \alpha_s)wh(y).\end{aligned}$$

Replacing the normal wholesale price w by the expected net wholesale price $w^n = w(1 - \gamma)$ and α_s by $\alpha_s^n = \alpha_s/(1 - \gamma)$, we derive that the TC model with default risks is *equivalent* to one without default risks and the supplier's capital cost rate α_s^n . Let w^{*n} and β_g^{*n} denote the equilibrium wholesale price and ECCR in the equivalent (no-default) model. The following theorem follows as a corollary of the results obtained for the base model.

Theorem 2.7 (Model with default risks for the supplier) *(a) All of the characterizations of the equilibrium behavior in the three Stackelberg games obtained for the base model in Theorems 2.1-2.4 and Propositions 2.1-2.3 apply both to the IS and TC contracts, under this default risk model. Moreover, the equilibria under TC are given by $(w_{\beta_g}^* | \beta_g = \beta_0) = (w_{\beta_g}^{*n}/(1 - \gamma) | \beta_g = \beta_0) [\beta_g^*(w = w_0) = \beta_g^{*n}(w = w_0(1 - \gamma))$ and $(w^*, \beta_g^*) = (w^{*n}/(1 - \gamma), \beta_g^*)$ in the first [second and full] Stackelberg games, respectively.*

(b) All of the comparison results in Theorems 2.5 and 2.6 and Proposition 2.4 continue to apply with relative rankings based on the comparison of $\alpha_s/(1 - \gamma)$ and α_r .

*(c) Under TC, there exists a critical value $0 \leq \gamma^c < \infty$ such that, as γ increases up to γ^c , even the net equilibrium price w^{*n} increases with γ , while the retailer's trade credit interest rate $\alpha^* = 0$; as γ increases beyond γ^c , $w^{*n} = w^*(1 - \gamma)$ is kept constant while α^* is increasing in α_s .*

Part (c) shows that as long as the loss fraction increases but remains below a critical level, the supplier chooses to increase his *net* wholesale price to compensate for the increasingly large default risks but maintains a zero trade credit interest rate. Beyond the critical loss fraction, the supplier maintains a constant *net* wholesale price but charges increasingly large trade credit interest rates instead.

2.7.2 Model with Default Risks for the Supplier and the Bank

If the bank is exposed to the same default risks as the supplier experiences under TC, this affects the profit function of the retailer under IF and IS. At the same time, the supplier being paid by the bank upon delivery is, under this arrangement, immune to the default risks unless the bank pays him only part upfront and the remainder if and when payment is received from the retailer. (In a factoring arrangement, for example, the percentage paid upfront is referred to as the *advance* a and the remainder as the reserve $1 - a$.) Recovery losses, now, reduce the net payment of the payables to the bank by a factor of $1 - \gamma$. Net payments to the supplier are reduced by a factor $\phi(\gamma) \leq 1$ with $\phi(0) = 1$; for example, in the factoring case with an advance percentage a , $\phi(\gamma) = a + (1 - \gamma)(1 - a)$. In response to the default risks, the bank may adjust the bank loan rate α_r as an increasing function of the loss fraction γ , see the discussion in Section 2.2. The profit functions are now given by:

$$\begin{aligned}\pi_r^{IS}(w, \beta, y) &= (p - w(1 - \gamma))s(y) - [(\alpha_r(\gamma)(1 - \gamma) - \beta)w + h_0]h(y), \\ \pi_s^{IS}(w, \beta, y) &= (\phi(\gamma)w - c)s(y) - \beta wh(y).\end{aligned}$$

In this case, the model under IF and IS is no longer equivalent to one without default risks and modified parameters; indeed, it may no longer be viewed as a special case of the general model described in (2.1) and (2.2). Nevertheless, we show in Appendix A.3:

Theorem 2.8 (Model with default risks for the supplier and the bank) *All the structural results for the three Stackelberg games, i.e., Theorems 2.1-2.4 and Propositions 2.1 and 2.2 continue to apply, both under IS and TC. Under TC, Proposition 2.3 continues to apply as well.*

As to the comparison results among the three financing schemes, they can no longer be proven on the basis of a simple comparison of two indices, (e.g., α_s and α_r or $\alpha_r/(1 - \gamma)$), but need to be evaluated numerically.

2.8 Conclusions

In this paper, we have characterized and compared the performance of a supply chain under three fundamental financing schemes: IF, IS and TC. The comparisons between these schemes have been enabled by showing that all of them can be embedded as special cases of a general model. Under each of these contracts, a wholesale price as well as a supplier's effective capital cost rate (ECCR) has to be selected. We have characterized the equilibrium behavior of the supply chain, assuming either the wholesale price is exogenously given and the ECCR is selected by the supplier in a Stackelberg setting, or, vice versa, the wholesale price is selected by the supplier under an exogenously given ECCR, as well as settings where both strategic parameters are determined endogenously (the full Stackelberg game). As an example, we have proven, see Theorem 2.4(c), that if the demand distribution is IFR and MIFR and the full wholesale price range is $[0, p]$ it is optimal in the full Stackelberg game for the supplier to select either the maximal or minimal feasible ECCR value. For uniform, exponential and Normal distributions, under minor parameter conditions, see Condition (C), it is, in fact, always optimal to engage in a *maximal* degree of supplier financing. We have shown that both the supplier and the retailer prefer TC over IS if and only if the supplier's cost of capital is lower than that of the retailer. This robust comparison result holds, in full generality, for the supplier, and for the retailer under the above Condition (C). Such robust results also prevail for the equilibrium wholesale price, the base-stock level and the expected sales volume. As a final highlight of our results, we have shown that a supplier financing scheme of either the IS or TC type, exists under which the supplier fares better than under IF. The same holds for the retailer, but only if the wholesale price is left unaltered. If the wholesale price can be varied by the supplier, the retailer invariably loses

when switching from IF to the optimal TC or IS contract, but a mutually beneficial IS or TC contract can be designed by imposing a participation contract. We refer to §2.6.3 as well as §1.2 for a summary of the remainder of the comparison results that have been obtained on the basis of a numerical study. Appendix A.4 establishes similar comparison results to settings where the contract terms are specified to generate a perfect coordination scheme, with aggregate profits split according to a Nash bargaining solution.

All of the above results have been obtained, first, in a base model without default risks. They have been extended to two generalized models with a general stochastic default and reorganization process. In the first such model, only the supplier is exposed to default risks, and in the second, both the supplier and the bank.

Future work should expand the comparisons among the three fundamental financing mechanisms to other alternatives. The Advance Purchase Discount (APD) contracts, for example, represent a hybrid combination of TC and IF mechanisms. The supplier offers the retailer two procurement options. In addition to traditional commitment orders at a given discounted price, the supplier makes available additional consignment inventory from which the retailer can draw, at a higher per unit price, to satisfy additional customer sales if and when they arise. Depending on who pays for the associated inventory carrying costs, the consignment inventory level may be selected by the supplier or by the retailer. More specifically, assume that at the beginning of each period, the retailer has access to *two* inventory pools. The primary pool is of size y_1 and consists of units the retailer has committed to and is charged a basic wholesale price w_1 . A second inventory pool of size y_2 is available in case the demand exceeds the first inventory pool. The second inventory pool is made available when needed, at a unit wholesale price w_2 along with an interest rate charge α for any unit in stock at the end of a period. A double base-stock policy is implemented to ensure that at the beginning of each period the inventory levels of the pools equal y_1 and y_2 respectively. It is easily verified that the profit functions

of the chain members are given by

$$\begin{aligned}\pi_r^{APD}(w_1, w_2, \alpha, y_1, y_2) &= p \min\{y_1 + y_2, D\} - w_1 y_1 - w_2 \min\{y_2, (D - y_1)^+\} \\ &\quad - \alpha w_2 (y_2 - (D - y_1)^+)^+ - (\alpha_r w_1 + h_0)(y_1 - D)^+ \\ \pi_s^{APD}(w_1, w_2, \alpha, y_1, y_2) &= w_1 y_1 + w_2 \min\{y_2, (D - y_1)^+\} \\ &\quad + (\alpha - \alpha_s) w_2 (y_2 - (D - y_1)^+)^+ - c(y_1 + y_2)\end{aligned}$$

Clearly, the wholesale prices w_1 and w_2 and the interest rate α are to be selected by the supplier while the retailer determines y_1 . As mentioned, the level y_2 may be selected by the supplier or the retailer. Cachon (2004) and Dong and Zhu (2007) consider the special case where $\alpha = 0$ and y_2 is selected by the supplier. Similar supply contracts with two procurement options, one based on early commitments at a lower price and the other based on last minute orders at a higher price, have been analyzed under the name “periodic commitment with flexibility contracts” by Anupindi and Bassok (1999) and Bassok and Anupindi (2008).

As explained in §2.2, bank loan rates are typically set as a spread over a common interest index such as the 3-month LIBOR rate. In our models, we have assumed that all parameters are stationary, the capital cost rates α_s and α_r included. This effectively assumes that the underlying LIBOR rate remains flat over the course of the planning horizon. A possible generalization would adopt one of the commonly used stochastic processes to represent fluctuations in the LIBOR rate, see, e.g., Chapter 30 in Hull (2008). The discrete time representations of these processes typically reduce to Markov chains. In this generalized model, the supplier may wish to adjust the wholesale price and the ECCR as a function of the state of the Markov chain. It can be easily shown that a state-dependent base-stock policy is optimal for the retailer. It would be interesting to analyze whether similarly robust comparison results among the three financing options (IF, IS and TC) can be obtained on the basis of the relative ranking of the *spreads* (over the LIBOR rate) that the supplier and the retailer obtain for their bank loans. We conjecture that these comparison results carry over to this generalized model.

Chapter 3

Inventory Models with Shelf Age and Delay Dependent Inventory Costs

We refer to Section 1.4 for an introduction of this chapter.

This chapter is organized as follows: we consider periodic review models with full backlogging, lost sales or partial backlogging in §3.1 while continuous review models with renewal demand processes or compound renewal demand processes are studied in §3.2. When not stated in the main text, proofs are deferred to Appendix B.1.

3.1 Periodic Review Models

In this section, we consider a general, single item periodic review inventory planning problem with a finite or infinite horizon. As in standard inventory models, we assume that demands in different periods are independent of each other. Orders arrive after a leadtime of L periods. To simplify the exposition we will initially assume that the leadtime L is deterministic. However, extensions to stochastic leadtime processes that are exogenous and sequential are straightforward, see below. As in standard inventory models, we assume the following sequence of events in each period:

(i) at the beginning of the period, the ending inventory of the previous period and all outstanding orders are observed, and a new order may be placed;

(ii) immediately thereafter, all units ordered L periods ago arrive and remain in stock until sold;

(iii) the period's demands occur thereafter.

(Assumption (ii) precludes inventory decreases resulting from factors other than the demand process, for example, random order yields or perishable goods.)

The planning problem is meaningful only if it consists of at least $L + 1$ periods. It is therefore convenient to denote the length of the planning horizon by $N + L$, with $N < \infty$ or $N = \infty$, depending upon whether a finite or infinite horizon problem is considered. All costs are discounted with a discount factor $\rho \leq 1$. We describe the general shelf age dependent holding cost structure as follows: an item ordered in period n , accrues an incremental carrying cost rate $\alpha_n(j)$ when reaching a shelf age of j periods, i.e., when still in stock at the end of period $n + L + j - 1$. We assume that, at any point in time, the incremental carrying costs of any unit in stock increases with its shelf age, i.e.,

(**ODY** - Old Dearer than Young): $\alpha_n(j) \geq \alpha_{n+1}(j - 1)$ for all n and j .

This condition ensures that it is optimal to deplete inventories on a FIFO basis. It is equivalent to the m-monotonicity property assumed in Levi et al. (2006, 2008) and Stauffer et al. (2011) in conjunction with other structural properties. Also, we assume $\alpha_n(0) = 0$ for all $n \geq 1$. Let

D_n = demand in period n ,

$D[t_1, t_2]$ = cumulative demand in periods $t_1, t_1 + 1, \dots, t_2 - 1$, with $t_2 > t_1$.

I_n^b = the beginning inventory level of period n (, after inclusion of any order placed one leadtime earlier).

I_n^e = the ending inventory level of period n (, after demands in this period occur).

q_n = order placed in period n .

q_n^j = the part of order q_n , placed in period n , which reaches a shelf age of at least j periods, i.e., which is still in inventory at the end of period $n + L + j - 1$, $j = 1, \dots, N - n + 1$.

By the definitions of I_n^e and I_n^b , $q_n = I_{n+L}^b - I_{n+L-1}^e$. It follows from the FIFO rule that

$$q_n^j = (q_n - (D[n+L, n+L+j] - I_{n+L-1}^e)^+)^+. \quad (3.1)$$

Since any order placed after period N does not arrive before the end of the planning horizon, we confine ourselves, without loss of optimality, to policies with $q_n = 0$ for all $n = N+1, \dots, N+L$. Similarly, since the first order q_1 does not arrive until period $L+1$, the costs in the first L periods cannot be controlled. We, therefore, define

C = the total expected discounted carrying cost over periods from $L+1$ to $L+N$.

All carrying costs in C can be attributed to orders placed in periods $n = 1, \dots, N$, as well as those associated with s_1 , the vector comprised of outstanding orders and the initial inventory at the beginning of the planning horizon. In view of the FIFO procedure, these initial units are used before any of the units ordered during the planning horizon. Their expected carrying costs are, therefore, independent of any ordering decisions, and hence denoted by $H_0(s_1)$.

Lemma 3.1 *Assume (ODY). C may be represented as a separable convex function of the sequence of beginning inventory levels $\{I_n^b, n = L+1, \dots, L+N\}$. More specifically,*

$$C = \sum_{n=1}^N \rho^n \tilde{G}_n(I_{n+L}^b) + A^0 \quad (3.2)$$

where

$$\tilde{G}_n(I_{n+L}^b) = \sum_{j=1}^{N-n+1} \rho^{L+j} [\alpha_n(j) - \alpha_{n+1}(j-1)] \mathbb{E}(I_{n+L}^b - D[n+L, n+L+j])^+. \quad (3.3)$$

and A^0 is a constant independent of any of the order decisions during the planning horizon.

Proof: Fix $n = 1, \dots, N$. For any $j = 1, \dots, N-n+1$, note that

$$\begin{aligned} q_n^j &= (I_{n+L}^b - I_{n+L-1}^e - (D[n+L, n+L+j] - I_{n+L-1}^e)^+)^+ \\ &= (I_{n+L}^b - D[n+L, n+L+j])^+ - (I_{n+L-1}^e - D[n+L, n+L+j])^+ \\ &= (I_{n+L}^b - D[n+L, n+L+j])^+ - (I_{n+L-1}^b - D[n+L-1, n+L+j])^+ \end{aligned} \quad (3.4)$$

where the first equality follows from (3.1), and the second equality can easily be verified by considering three cases: $I_{n+L-1}^e \geq D[n+L, n+L+j]$, $I_{n+L}^b \leq D[n+L, n+L+j]$ or $I_{n+L}^b > D[n+L, n+L+j] > I_{n+L-1}^e$. As to the third equality, note that regardless of whether all stockouts are backlogged, all result in lost sales, or *partial* backlogging takes place, at any period m :

$$I_m^b - D_m \leq I_m^e \leq (I_m^b - D_m)^+. \quad (3.5)$$

The upper bound in (3.5) corresponds with the case where *all* stockouts are lost, while the lower bound corresponds with the case where all are backlogged. Note that when $I_m^e > 0$, $I_m^e = I_m^b - D_m$, since in this case the lower and upper bounds coincide. Applying this for $m = n + L - 1$, this verifies the equality in (3.4) when $I_{n+L-1}^e > 0$. Moreover, when $I_{n+L-1}^e \leq 0$, $0 = (I_{n+L-1}^e - D[n+L, n+L+j])^+ \geq (I_{n+L-1}^b - D[n+L-1, n+L+j])^+ \geq 0$ where the first inequality is due to the lower bound in (3.5), thus verifying the equality in (3.4) for the case where $I_{n+L-1}^e \leq 0$, as well.

The total expected shelf age dependent holding costs associated with order q_n , discounted back to period 1, is given by:

$$\begin{aligned} & \rho^n \sum_{j=1}^{N-n+1} \rho^{L+j} \alpha_n(j) \mathbb{E} q_n^j \\ = & \rho^n \sum_{j=1}^{N-n+1} \rho^{L+j} \alpha_n(j) \mathbb{E} \{ (I_{n+L}^b - D[n+L, n+L+j])^+ \\ & - (I_{n+L-1}^b - D[n+L-1, n+L+j])^+ \}. \end{aligned}$$

Thus, summing the expected holding costs associated with *all* orders placed in periods $n = 1, \dots, N$, (as well as the costs associated with s_1), we obtain:

$$\begin{aligned} C &= H_0(s_1) + \sum_{n=1}^N \rho^n \sum_{j=1}^{N-n+1} \rho^{L+j} \alpha_n(j) \mathbb{E} (I_{n+L}^b - D[n+L, n+L+j])^+ \\ & - \sum_{n=1}^N \rho^n \sum_{j=1}^{N-n+1} \rho^{L+j} \alpha_n(j) \mathbb{E} (I_{n+L-1}^b - D[n+L-1, n+L+j])^+ \\ & = H_0(s_1) + \sum_{n=1}^N \rho^n \sum_{j=1}^{N-n+1} \rho^{L+j} \alpha_n(j) \mathbb{E} (I_{n+L}^b - D[n+L, n+L+j])^+ \end{aligned}$$

$$\begin{aligned}
& - \sum_{n=0}^{N-1} \rho^{n+1} \sum_{j=1}^{N-n} \rho^{L+j} \alpha_{n+1}(j) \mathbb{E}(I_{n+L}^b - D[n+L, n+L+j+1])^+ \text{ (shift } n \text{ by } -1) \\
= & H_0(s_1) + \sum_{n=1}^N \rho^n \sum_{j=1}^{N-n+1} \rho^{L+j} \alpha_n(j) \mathbb{E}(I_{n+L}^b - D[n+L, n+L+j])^+ \\
& - \sum_{n=0}^{N-1} \rho^n \sum_{j=2}^{N-n+1} \rho^{L+j} \alpha_{n+1}(j-1) \mathbb{E}(I_{n+L}^b - D[n+L, n+L+j])^+ \text{ (shift } j \text{ by } +1) \\
= & H_0(s_1) + \sum_{n=1}^N \rho^n \sum_{j=1}^{N-n+1} \rho^{L+j} \alpha_n(j) \mathbb{E}(I_{n+L}^b - D[n+L, n+L+j])^+ \\
& - \sum_{n=0}^N \rho^n \sum_{j=1}^{N-n+1} \rho^{L+j} \alpha_{n+1}(j-1) \mathbb{E}(I_{n+L}^b - D[n+L, n+L+j])^+ \\
= & H_0(s_1) + \sum_{n=1}^N \rho^n \sum_{j=1}^{N-n+1} \rho^{L+j} [\alpha_n(j) - \alpha_{n+1}(j-1)] \mathbb{E}(I_{n+L}^b - D[n+L, n+L+j])^+ \\
& - \sum_{j=1}^{N+1} \rho^{L+j} \alpha_1(j-1) \mathbb{E}(I_L^b - D[L, L+j])^+ = \sum_{n=1}^N \rho^n \tilde{G}_n(I_{n+L}^b) + A^0
\end{aligned}$$

where the fourth equality is obtained by adding zero terms with $j = 1$ and $n = 1, \dots, N$ to the expression after the minus since, by our assumption, $\alpha_n(0) = 0$, and where $A^0 \equiv H_0(s_1) - \sum_{j=1}^{N+1} \rho^{L+j} \alpha_1(j-1) \mathbb{E}(I_L^b - D[L, L+j])^+$. A^0 only depends on s_1 since the inventory level at the beginning of period L , I_L^b , does not depend on any order decisions. Since $\alpha_n(j-n) - \alpha_{n+1}(j-n-1) \geq 0$ by (ODY), and since each of the terms $\mathbb{E}(I_{n+L}^b - D[n+L, n+L+j])^+$ is convex in I_{n+L}^b , we conclude that $\tilde{G}_n(I_{n+L}^b)$ is convex. \square

3.1.1 Full Backlogging

As mentioned, under Assumption (ii) in a system with full backlogging, all inventory information can be aggregated into the single *inventory position* measure. Thus, for any $n = 1, 2, \dots$, let

y_n = the inventory position at the beginning of period n , *after* ordering.

The full backlogging assumption guarantees the simple dynamic recursion:

$$I_{n+L}^b = y_n - D[n, n+L]. \quad (3.6)$$

Using this identity, the following theorem follows from Lemma 3.1.

Theorem 3.1 (Equivalence of shelf age and inventory level dependent inventory carrying cost) *Assume (ODY).*

(a) *The model with shelf age dependent inventory carrying costs is equivalent to the one in which those costs are represented by the following convex cost functions of the inventory positions in periods $n = 1, \dots, N$:*

$$G_n(y_n) = \sum_{j=1}^{N-n+1} \rho^{L+j} [\alpha_n(j) - \alpha_{n+1}(j-1)] \mathbb{E}(y_n - D[n, L+n+j])^+. \quad (3.7)$$

(b) *Let $N = \infty$ and $\alpha_n(\cdot) = \alpha(\cdot)$ for all $n = 1, 2, \dots$. Assume $h \equiv \sum_{j=0}^{\infty} \rho[\alpha(j+1) - \alpha(j)] = \rho \lim_{t \rightarrow \infty} \alpha(t) < \infty$. Define a random leadtime Λ with the following distribution:*

$$P(\Lambda = j) = \frac{\rho[\alpha(j+1) - \alpha(j)]}{h}, \quad j = 0, 1, \dots. \quad (3.8)$$

The model with shelf age dependent inventory carrying costs is equivalent to one with linear holding costs at a constant rate h , but with an extended, stochastic leadtime $L + \Lambda$.

Proof: (a) follows from Lemma 3.1 and (3.6).

(b): Under the assumptions in this part, (3.7) can be written as

$$\begin{aligned} G(y_n) &= \sum_{j=1}^{\infty} \rho^{L+j} [\alpha(j) - \alpha(j-1)] \mathbb{E}(y_n - D[n, L+n+j])^+ \\ &= \sum_{j=0}^{\infty} \rho^{L+j+1} [\alpha(j+1) - \alpha(j)] \mathbb{E}(y_n - D[n, L+n+j])^+ \\ &= h \sum_{j=0}^{\infty} \rho^{L+j} P(\Lambda = j) \mathbb{E}(y_n - D[n, L+n+j])^+, \end{aligned}$$

which proves the claim. \square

Remark 1: When the nonlinear shelf age dependent cost structure arises because of trade credit arrangements (see Section 1.3), the function $\alpha(\cdot)$ is typically piecewise constant. A frequently used structure, referred to as a two-part credit scheme (see Cuñat (2007)), has an interest-free grace period (F), followed by a constant positive interest rate thereafter. In the infinite horizon model of Theorem 3.1(b), a piece-wise constant $\alpha(\cdot)$ function is equivalent to linear holding costs with

an additional leadtime component Λ which has support only on the period lengths in which the function $\alpha(\cdot)$ experiences an upward jump, see (3.8). In particular, in the two-part contract scheme, Λ has a two point distribution, i.e., the support is given by $\{0, F\}$.

Remark 2: More generally, order leadtimes may be characterized by a stochastic process $\{L(n) : n \geq 0\}$ with $L(n)$ the leadtime experienced by an order placed in period n . We assume the process is *exogenous*, i.e., it is independent of the demand process, as well as *sequential*, i.e., $n + L(n) \leq n' + L(n')$ for all $n < n'$, with probability one. Under sequential leadtime processes, orders do not cross. We refer to Zipkin (1986) for an extensive discussion of such processes and their applications.

Assume the exogenous sequential leadtime process has a steady-state distribution \mathcal{L} . It is easily verified from the proof of Theorem 3.1 that the equivalency result in part (a) continues to apply with the constant L replaced by the random variable \mathcal{L} :

$$G_n(y_n) = \mathbb{E} \left[\sum_{j=1}^{N-n+1} \rho^{\mathcal{L}+j} (\alpha_n(j) - \alpha_{n+1}(j-1)) \mathbb{E}(y_n - D[n, \mathcal{L} + n + j])^+ \right]. \quad (3.9)$$

Here, the expectation is taken both over the distribution of \mathcal{L} and that of the demand variables $\{D_n, D_{n+1}, \dots\}$. Part (b) can be generalized with the leadtime distribution in the equivalent model with linear holding costs, now given by $\mathcal{L} \oplus \Lambda$. (For any pair of random variables X and Y , $X \oplus Y$ denotes the convolution of X and Y .)

Assume, now, that in addition to the shelf age dependent inventory costs, the *remaining* inventory and backlogging costs may be represented by *convex* functions $\Gamma_n(y_n)$ of the inventory position after ordering, $n = 1, \dots, N + L$. The following theorem shows that the structural results pertaining to various standard models, can therefore be generalized to allow for general shelf age dependent costs. (We confine ourselves to a few basic models.)

Theorem 3.2 (Shelf age dependent costs: Structural results with backlogging) (a) *Assume $N < \infty$ and the ordering costs are proportional to the order sizes. A time dependent base-stock policy is optimal, i.e., in each period, $n = 1, \dots, N$, there*

exists a base-stock level S_n^* , such that it is optimal to raise the inventory position to S_n^* whenever it is below S_n^* .

(b) Assume $N < \infty$, and in the model of part (a), assume all parameters are stationary while any inventory at the end of the horizon can be returned at the original purchase price. Then $S_1^* \leq S_2^* \leq \dots \leq S_N^*$.

(c) Assume $N = \infty$ and the model of part (a), with all parameters stationary. Assume orders are subject to a capacity limit $C(\leq \infty)$. Under both the discounted and long-run average cost criteria, there exists a stationary modified base-stock policy with base-stock level S^* which is optimal: if the stationary inventory position in any given period is below S^* , an order is placed which brings the inventory position as close as possible to S^* .

(d) Assume an order in period n incurs a fixed cost K_n and a variable per unit cost c_n . If $N < \infty$, assume $K_n \geq \rho K_{n+1}$ for all $n = 1, \dots, N - 1$. There exists a time-dependent (s_n, S_n) -policy which is optimal. If $N = \infty$, an (s^*, S^*) policy is optimal both under the discounted total cost and the long-run average cost criteria.

(e) Let $N = \infty$ and consider the model with fixed-plus-linear order costs. Assume that all cost parameters, the cost functions $G_n(\cdot)$ and $\Gamma_n(\cdot)$, as well as the one-period demand distribution depend on a state of the world variable ω which evolves in accordance with a finite state irreducible Markov chain. Both under the total discounted and the long-run average cost criteria, a state-dependent (s, S) -policy is optimal.

(f) Let $N < \infty$ and $L = 0$. Assume that the demand distribution in each period depends on the sale price selected for that period, and that this price may be varied arbitrarily. Assume, further, that the demand variable for period n is of the form: $D_n = d_n(p_n) + \epsilon_n$, where $d_n(p_n)$ is a deterministic strictly decreasing function with the expected revenue $R_n(d) \equiv dd_n^{-1}(d)$ a concave function. The sequence $\{\epsilon_n\}$ is a sequence of independent random variables whose distribution is independent of the selected prices. Under fixed-plus-linear ordering costs, a so-called (s_n, S_n, p_n) policy is optimal, i.e., the inventory replenishment rules have an (s_n, S_n) -structure and price depends on the initial inventory level at the beginning of a period.

Under linear ordering costs, it is optimal to use a base-stock/list price strategy, i.e., the inventory replenishment strategy is of a base-stock type while the selected price p_n^ is a decreasing function of the inventory level of after ordering.*

Remark 3: When $L = 0$, the equivalency results in Theorem 3.1 and the structural results in Theorem 3.2 continue to apply in models with lost sales, as is easily verified from the proof of Theorem 3.1. We address the case of a general, positive, leadtime L in §3.1.2.

To our knowledge, the first and only structural result for models with general shelf age dependent carrying costs, was obtained in a recent paper of Gupta and Wang (2009). These authors proved part (a) of Theorem 3.2 for the case where the $\alpha(\cdot)$ function becomes flat after a certain period κ , i.e., $\alpha(j) = \alpha(\kappa)$, and thus permits a finite-dimensional state representation when disaggregating inventory levels according the items' shelf age. (The state of the system has dimension $\kappa + L$, independent of the planning horizon N .) The equivalency result in Theorem 3.1 establishes that the problem can be formulated as a dynamic program with a one-dimensional state space of a well-known structure, and this for arbitrary (increasing) $\alpha(\cdot)$ functions.

In their concluding section, Gupta and Wang (2009) question how the inclusion of fixed costs would impact the structure of the optimal policy. This question is resolved by part (d) of the above theorem, i.e., it is optimal to use an (s, S)-policy acting on the inventory position(, under the same assumption for the time dependence of the fixed costs that is required in the standard model with linear holding costs). Maddah et al. (2004) consider the long-run average cost criterion for the special case of a two-part credit scheme, see above, i.e., the $\alpha(\cdot)$ function adopts two distinct values. These authors restrict themselves to the class of (s, S) policies and develop heuristics to compute the best parameter pair. However, the equivalency result in Theorem 3.2 shows that this class of policies, in fact, contains the optimal policy and that the exact algorithms by Veinott and Wagner (1965) or Zheng and Federgruen (1992) can be used without modification, employing the transformed cost functions $G(\cdot)$, as well as the function $\Gamma(\cdot)$. Robb and Silver (2004)

address the same special case as Maddah et al. (2004), making an upfront restriction to (R, S) policies, and proposing heuristics for this class. (Under an (R, S) policy, the inventory position is increased to a level S every R time units.)

3.1.2 Systems with Lost Sales or Partial Backlogging

When all or some stockouts result in lost sales and orders get delivered after a positive leadtime L , it is no longer possible to represent the state of the system via the inventory position only. Instead, it is necessary to represent the system via a $(L + 1)$ -dimensional state vector, consisting of the inventory level and the size of the orders placed in the prior L periods. However, the result in Lemma 3.1 implies the following Theorem:

Theorem 3.3 (Systems with lost sales or partial backlogging) *Assume (ODY). Consider an inventory system with lost sales or partial backlogging of stockouts. The model with general shelf age dependent carrying costs is equivalent to one in which, in each period $n = 1, \dots, N$, an inventory level dependent convex function $\tilde{G}_n(I_{n+L}^b)$ is charged for the carrying costs.*

Consider, for example, the case of lost sales. The structure of the optimal policy is complex, even under the simplest setting with linear order costs. The model was first formulated and analyzed by Karlin and Scarf (1958) and Morton (1969). Recently, Zipkin(2008a, 2008b) identified bounds and monotonicity properties of the optimal order quantities. While the model in these references assumes that a *specific* convex cost function of each period's beginning inventory level is charged, it can easily be verified that the above results apply to arbitrary convex functions. This implies that all of bounds and monotonicity results for the standard model with linear holding costs continue to apply under general shelf age dependent carrying costs.

3.1.3 General Delay Dependent Backlogging Costs

As mentioned in Section 1.3, similar to the carrying costs, the backlogging costs associated with a demand unit may increase in a general non-linear way with the

amount of time that the unit has been delayed. Let $\beta(j)$ denote the *incremental* backlogging cost rate when a unit of demand waits for j periods. (See below for a discussion of more general structures which differentiate according to the period in which a backlogged demand unit arises.) Analogous to the (ODY)-assumption for shelf-age dependent carrying costs, we assume

$$\text{(ODY-B): } \beta(j) \geq \beta(j-1) \text{ for all } j \geq 1,$$

i.e., older outstanding demands incur higher costs. Under the (ODY-B) assumption, it is easily verified that demands are optimally filled on a FIFO basis. Also, as in §3.1.1, we assume that, in addition to the delay dependent backlogging costs, the remaining inventory and backlogging costs may be represented by convex functions $\Gamma_n(\cdot)$ of the inventory position after ordering. Together, these assumptions correspond with Assumption 1 in Huh et al. (2010).

As to the order leadtime L , we confine ourselves to the deterministic case, for reasons explained below. Huh et al. (2010) have shown that a model with general delay dependent backlogging costs is equivalent to one with traditional level-dependent backlogging cost, albeit under certain restrictions and assumptions.

Define $N_{n+L,j}$ as the number of backlogged demand units at the end of period $n+L$ that have been delayed for j periods, $j = 1, \dots, n+L$. Note that, for $j \leq L+1$,

$$\begin{aligned} N_{n+L,j} &= (D_{n+L-j+1} - (y_n - D[n, n+L-j])^+)^+ \\ &= (D[n, n+L-j+1] - y_n)^+ - (D[n, n+L-j] - y_n)^+, \end{aligned} \quad (3.10)$$

where the first equality follows from the FIFO rule and the second one may easily be verified by distinguishing between the case where y_n is greater than $D[n, n+L-j]$ and the the case where it is smaller. Thus, the part of the backlog at the end of period $n+L$ that has a delay duration $j \leq L+1$ may be expressed as a function of y_n , the inventory position after ordering at the beginning of period n , as well as the demand variables, not yet observed at that time. This allows us to express the expected (incremental) backlogging costs for all units with a delay of $j \leq L+1$ periods, as a function of y_n only:

$$\beta(j)\mathbb{E}\{(D[n, n+L-j+1] - y_n)^+ - (D[n, n+L-j] - y_n)^+\}, j = 1, \dots, L+1. \quad (3.11)$$

Unfortunately, a similar expression can not be obtained for backlogged units with a delay beyond $(L + 1)$ periods. (Their magnitudes depend on demands preceding period n , i.e., demands already observed by the beginning of period n , as well as orders placed prior to period n .) Only their *aggregate*, i.e., all backlogs with a delay beyond $L + 1$ periods, can be written as a function of y_n :

$$\sum_{j=L+2}^{n+L} N_{n+L,j} = (y_n)^-. \quad (3.12)$$

(If $y_n < 0$, $-y_n$ units are already backlogged at the beginning of period n and will not be filled by the end of period $n + L$; therefore, these $-y_n$ units will be part of the backlog at the end of period $n + L$ with a delay at least $L + 2$. At the same time, if $y_n > 0$, all units backlogged at the beginning of period n will be filled at the end of period $n + L$, so that any backlog at the end of that period has a delay at most equal to $L + 1$.) Thus, to enable the above mentioned equivalency with a traditional level dependent cost structure, and hence to enable a one-dimensional state space representation of the planning problem, requires one of the following two assumptions:

Assumption (NIP) (Non-negative Inventory Position after ordering): $y_n \geq 0$ for $n = 1, \dots, N$.

Assumption (CBL) (Marginal Cost Rate is Constant Beyond Leadtime): $\beta(n) = \beta(L + 2)$ for any $n \geq L + 2$.

Note, by (3.12), that Assumption (NIP) is equivalent to the assumption that no demand is backlogged for more than $L + 1$ periods.

Under Assumption (NIP), the total expected delay dependent backlogging costs for backlogs at the end of period of $n + L$ is obtained by summing (3.11) for $j = 1, \dots, L + 1$:

$$G_n^-(y_n) = \sum_{j=1}^{L+1} (\beta(j) - \beta(j-1)) \mathbb{E}(D[n, n+L-j+1] - y_n)^+ \quad (3.13)$$

with the convention $\beta(0) = 0$. Under Assumption (CBL), we have one additional term:

$$G_n^-(y_n) = \sum_{j=1}^{L+1} (\beta(j) - \beta(j-1)) \mathbb{E}(D[n, n+L-j+1] - y_n)^+ + \beta(L+2)(y_n)^-. \quad (3.14)$$

We conclude:

Lemma 3.2 (Huh et al. (2010)) *Assume (ODY-B) and either Assumption (NIP) or Assumption (CBL) holds. The model is equivalent to a traditional model with **convex, level** dependent cost functions $\{G_n^-(\cdot) + \Gamma_n(\cdot), n = 1, \dots, N\}$, with $G_n^-(\cdot)$ given by (3.13) or (3.14).*

It should be noted that Assumptions (NIP) and (CBL) impose significant restrictions, either in terms of the shape of the backlogging cost rate function (Assumption CBL) or by imposing a potentially restrictive constraint for the feasible order set (Assumption NIP). All the structural results in Theorem 3.2 continue to prevail in the presence of these delay dependent backlogging costs, either under Assumption (CBL) or by imposing the upfront restriction that the inventory position after ordering must be non-negative. In non-stationary settings or under discounting the latter assumption may be essentially restrictive.

The following Theorem shows that the equivalency results in Lemma 3.2 may be employed to obtain the optimality of base-stock and (s, S) -policies in models with independent and exogenously specified demands, and with linear and fixed-plus-linear order costs respectively.

Theorem 3.4 (Delay Dependent Costs: Structural Results in Periodic Review Systems) *Assume Assumption (ODY-B) applies. Consider a model in which the sequence of demand distributions $\{D_n\}$ is independent, with exogenously specified distributions.*

(a) *Let $N \leq \infty$. Assume order costs are linear, with a stationary variable cost rate c per unit ordered. Assume also that future costs are not discounted, i.e., $\rho = 1$. (When $N = \infty$, we employ the long-run average cost criterion.) A possibly time dependent base-stock policy is optimal.*

(b) *Assume order costs are linear, with arbitrary time-dependent order cost rates $\{c_n, n = 1, 2, \dots\}$ and an arbitrary discount factor $\rho \leq 1$. A time dependent base-stock policy is optimal, under either the (NIP) or (CBL) assumption.*

(c) *Let $N < \infty$. Assume order costs are fixed-plus-linear with arbitrary time-*

dependent fixed order costs $\{K_n\}$ that satisfy $K_n \geq \rho K_{n+1}$ and arbitrary variable per unit order cost rates $\{c_n\}$ and a general discount factor $\rho \leq 1$. A time-dependent (s, S) -policy is optimal, under the (NIP) or (CBL) assumption. The same applies when $N = \infty$ and the parameters are stationary, both under the discounted and long-run average cost criteria.

Part (a) of Theorem 3.4 was shown by Huh et al. (2010). (These authors present this result in Remark 4, albeit for a model without any order costs; the generalization to stationary linear order costs and $\rho = 1$ is straightforward.) For models with fixed-plus-linear order costs, Huh et al. (2010) establish the optimality of (s, S) policies but only under constraint (CBL) in combination with another assumption which restricts the shape of the demand distributions. (This is their Assumption 2; alternative (a) of this assumption, which restricts the class of demand distributions, permits more general incremental cost rates $\beta(\cdot)$ that may be unimodal as opposed to increasing; alternatives (b) and (c) allow for fixed costs in any period in which a backlog prevails.) Using Lemma 3.2's equivalency result, the structural results of parts (b), (c), (e) and (f) in Theorem 3.2 can be shown for a model with delay dependent backlogging costs, again under assumption (NIP) or (CBL).

Note also that, under Assumption (NIP), the above equivalency result, like that in §3.1.1, continues to apply under random leadtimes. The only required modification is that, in (3.13), the expectation needs to be taken over the leadtime, as well as the demand distributions. Under Assumption (CBL), it is possible to differentiate the incremental backlogging cost rate, not just as a function of the delay experienced, but also as a function of the period in which a backlogged unit is demanded. The required generalization of (CBL) calls for a uniform incremental backlogging cost rate for all units delayed beyond $L + 1$ periods irrespective of the time at which the unit is demanded.

3.2 Continuous Review Models

Stochastic *continuous review* inventory models with discrete demand epochs typically assume that demands are generated by a compound renewal process. As in §3.1, we assume a sequential and exogenous leadtime process. Let L denote a random variable whose distribution corresponds with the steady-state leadtime distribution. In this continuous review model, $\alpha(t)$ now denotes the *marginal* inventory cost rate incurred for an item that has a shelf age t and $\beta(t)$ the marginal backlogging cost rate when a unit of demand has been waiting for t time units. In accordance with the stationary version of the (ODY) and (ODY-B) properties, the functions $\alpha(\cdot)$ and $\beta(\cdot)$ are, again, assumed to be increasing with $\beta(0) > 0$. In view of the monotonicity of the $\alpha(\cdot)$ and $\beta(\cdot)$ functions and the fact that orders do not cross, it continues to be optimal to deplete inventories on a FIFO basis. To ensure that various expected carrying and backlogging costs are finite, we assume that both rate functions $\alpha(\cdot)$ and $\beta(\cdot)$ are polynomially bounded, i.e., there exists a power l such that $\alpha(t) = O(t^l)$ and $\beta(t) = O(t^l)$, as $t \uparrow \infty$.

We first, in §3.2.1, analyze *simple* renewal demand processes; in §3.2.2, we extend our results to general *compound* renewal processes, adopting a different and slightly more restrictive approach.

3.2.1 Renewal Demand Processes

Assume first that demands are generated by a renewal process, i.e., the interarrival times between demand epochs are iid random variables X_1, X_2, \dots , distributed like X with mean τ and $\mathbb{E}X^{l+1} < \infty$. A single unit is demanded at each demand epoch. We assume that stockouts are backlogged. Our objective is to minimize long-run average costs. We show that this model is, again, equivalent to a standard continuous review model with *level dependent* inventory costs, one of special structure, guaranteeing the optimality of procurement strategies of a specific, simple structure. We show, in particular, that under fixed-plus-linear ordering costs, an (r, q) policy is optimal in the presence of our general shelf-age/delay dependent holding/backlogging

costs.

Following the single-unit decomposition approach, first introduced by Axsäter(1990, 1993), we associate every ordered unit with a specific demand unit. (The term “single-unit decomposition approach” appears to have been coined by Muharremoglu and Tsitsiklis (2008).) More specifically, assuming the system starts empty, i.e., without inventory, backlogs or outstanding orders, and since ordered units are used on a FIFO basis, the j -th ordered unit (since time 0) is used to fill the j -th demand unit(, again, since time 0). This implies that, if at a given demand epoch, an order is placed that elevates the inventory position from x to $y > x$, the $(y - x)$ units in the order may be given an index $j = x + 1, x + 2, \dots, y$, such that the j -th item is used to satisfy the j -th ($(-j)$ -th) demand following (preceding) the order epoch if $j > (\leq)0$. Let A_j denote the difference between the arrival time of the demand unit matched to ordered item j and that of item j . If A_j is positive, it represents the shelf age of item j . If it is negative, the j -th item in the ordered batch arrives after the associated demand epoch, so that this demand unit experiences a backlog time equal to $-A_j$. By the definition of A_j , we have

$$A_j = \begin{cases} \sum_{i=1}^j X_i - L & \text{when } j > 0, \\ -\sum_{i=1}^{-j} X_i - L & \text{when } j \leq 0. \end{cases} \quad (3.15)$$

The functions $\alpha(\cdot)$ and $\beta(\cdot)$ are defined on \mathcal{R}^+ . Extend their definitions to the complete real line \mathcal{R} as follows: $\bar{\alpha}(s) = \alpha(s)$ for $s > 0$ and $\bar{\alpha}(s) = 0$ for $s \leq 0$; $\bar{\beta}(s) = \beta(s)$ for $s > 0$ and $\bar{\beta}(s) = 0$ for $s \leq 0$. Note that the extended functions $\bar{\alpha}(\cdot)$ and $\bar{\beta}(\cdot)$ continue to be increasing. Let $\hat{G}(j)$ denote the total expected inventory and backlogging costs associated with the j -th item. $\hat{G}(j) = \mathbb{E}H(A_j) + \mathbb{E}J(A_j)$, where $H(t) = \int_0^t \bar{\alpha}(s) ds$ and $J(t) = \int_0^{-t} \bar{\beta}(s) ds$. We first need the following Lemma:

Lemma 3.3 *The function $\hat{G}(y)$ is finite and convex in y .*

Optimality of an (r, q) -policy under fixed-plus-linear ordering costs

We now show that an (r, q) -policy is optimal under fixed-plus-linear ordering costs. We first consider a semi-Markov decision process (SMDP), embedded on

the demand epochs, with the same state space $S = \mathcal{Z}$ and the same action sets $A(x) = \{y \geq x : y \text{ is integer}\}$ as the original control problem, however with the following one-step expected cost functions:

$$\gamma(x, y) = K\delta(y - x) + c(y - x) + \sum_{j=x+1}^y \hat{G}(j), \quad (3.16)$$

where K and c denote the fixed and linear order costs, as in Section 3.1, and

$$\delta(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases}$$

The last term denotes the carrying and delay costs associated with an order that elevates the inventory position from x to $y \geq x$, according to the above described matching scheme. Finally, the state dynamics in this SMDP are identical to those in the original control problem, i.e., the state at the next decision epoch equals $(y - 1)$.

This SMDP is, strictly speaking, only a *relaxation* of the real control problem, in that for policies with an average order rate below $1/\mathbb{E}(X) = \tau^{-1}$, the carrying and backlogging costs associated with (infinitely) many demand units are unaccounted for, since unmatched with any orders. However, we will construct a solution to the long-run average optimality equation of the (approximating) SMDP and show that the stationary policy that satisfies this optimality equation is optimal in the original model. Moreover, this stationary policy will be shown to be of an (r, q) -type.

Theorem 3.5 (Renewal Demand Processes) *(a) Under fixed-plus-linear ordering costs, an (r, q) -policy is optimal. More specifically, the policy (r^*, q^*) which is optimal in the periodic review model with one-step expected carrying and shortage cost function $\hat{G}(\cdot)$ is optimal in the continuous model as well.*

(b) The long-run average cost under any (r, q) policy is given by

$$c(r, q) = c\tau^{-1} + \frac{K + \sum_{y=r+1}^{r+q} \hat{G}(y)}{q\tau}. \quad (3.17)$$

Proof: (a) The long-run average optimality equation in the approximate SMDP

is given by

$$\begin{aligned} v(x) &= \min_{y \geq x} \{\gamma(x, y) - g\tau + v(y-1)\} \\ &= \min_{y \geq x} \{K\delta(y-x) + c(y-x) + G(y) - G(x) - g\tau + v(y-1)\}. \end{aligned} \quad (3.18)$$

where we define

$$G(y) \equiv \begin{cases} \sum_{j=1}^y \hat{G}(j), & \text{if } y > 0, \\ -\sum_{j=y+1}^0 \hat{G}(j), & \text{if } y \leq 0, \end{cases}$$

with the convention that $\sum_{j=a}^b u(j) = 0$ when $a > b$ for any sequence $\{u(j)\}$. (One can easily verify that, with this definition, $G(y) - G(x) = \sum_{j=x+1}^y \hat{G}(j)$, regardless of the signs of x and y .) Adding $G(x)$ to both sides of equation (3.18) and defining $\hat{v}(x) \equiv v(x) + G(x)$, we obtain the following equivalent optimality equation in terms of $\{\hat{v}(\cdot), g\}$:

$$\begin{aligned} \hat{v}(x) &= \min_{y \geq x} \{K\delta(y-x) + c(y-x) + G(y) - G(y-1) - g\tau + \hat{v}(y-1)\} \\ &= \min_{y \geq x} \{K\delta(y-x) + c(y-x) + \hat{G}(y) - g\tau + \hat{v}(y-1)\}. \end{aligned} \quad (3.19)$$

Equation (3.19) may be interpreted as the optimality equation in the classical periodic review inventory model with periods of constant length τ , fixed-plus-linear ordering costs and immediate expected cost $\hat{G}(y)$ whenever the inventory position after ordering equals y . By Lemma 3.3, $\hat{G}(\cdot)$ is convex. It follows from Iglehart(1963a, 1963b), based on Scarf (1960), that this optimality equation has a -in fact bounded- solution $\{w^*(\cdot), g^*\}$, see also Veinott (1965). Moreover, an (s, S) -policy achieves the minimum in (3.19), for every state $x \in S$. (g^* denotes the long-run average cost of this (s, S) policy.) Finally, since at every demand epoch, a *unit* size demand occurs, this (s, S) policy is of an (r, q) type, say with reorder level r^* and fixed order size $q^* = S^* - s^*$. Clearly the function $v^*(x) = w^*(x) - G(x)$ is the solution to the original optimality equation (3.18), in conjunction with the scalar g^* , and the same (r^*, q^*) policy achieves the minimum in (3.18) for every state $x \in S$.

We now show that this (r^*, q^*) policy is optimal in the SMDP with optimality equation (3.18). Let $\pi = (f_1, f_2, \dots)$ denote an arbitrary Markov strategy, i.e., on

an arbitrary sequence of possibly randomized policies, such that policy f_t is used in period t . Given a specific starting state x_1 , let $\{x_t : t > 1\}$ denote the sequence of states adopted under the policy π . Note that in period t , the inventory position is raised from x_t to $y_t \equiv x_{t+1} + 1$. Since the (r^*, q^*) policy is optimal (among all Markov strategies) in the transformed periodic review model, we have

$$\begin{aligned}
g^* \tau &\leq \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}_\pi [K\delta(y_t - x_t) + c(y_t - x_t) + \hat{G}(y_t)] \\
&= \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}_\pi [K\delta(y_t - x_t) + c(y_t - x_t) + (G(y_t) - G(y_t - 1))] \\
&= \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}_\pi [K\delta(y_t - x_t) + c(y_t - x_t) + (G(y_t) - G(x_{t+1}))] \\
&= \liminf_{T \rightarrow \infty} \frac{1}{T} \left\{ \sum_{t=1}^T \mathbb{E}_\pi [K\delta(y_t - x_t) + c(y_t - x_t) + (G(y_t) - G(x_t))] - \mathbb{E}_\pi G(x_{T+1}) \right\} \\
&\leq \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}_\pi [K\delta(y_t - x_t) + c(y_t - x_t) + (G(y_t) - G(x_t))].
\end{aligned}$$

This establishes that the long-run average cost under any policy π is bounded from below by that under the (r^*, q^*) policy. (To verify the last equality, note that $\lim_{T \rightarrow \infty} \frac{G(x_1)}{T} = 0$. To verify the last inequality, note that $G(\cdot) \geq 0$).

It remains to be shown that the (r^*, q^*) policy is optimal in the original model, and not just in the SMDP. However, for any Markov policy π under which the long-run average ordering rate is at or above the average demand rate τ^{-1} , we have that the entire cost sequence and, in particular, the long-run average cost in the original model equal those in the SMDP. By the above argument, the latter is therefore bounded from below by the long-run average cost of the (r^*, q^*) policy. The remaining Markov policies π have a long run average ordering rate strictly lower than the demand rate τ^{-1} . Under such a policy, the state of the system $\{x_t\}$ is almost surely below any value $-M$ after a finite amount of time. This implies that backlogging costs are incurred at a rate at least equal to $M\beta(0) > 0$. This lower bound applies for all $M > 0$, thus establishing that any such policy π has infinitely large long-run average costs, because of experiencing ever larger backlogs. We conclude that the (r^*, q^*) policy is indeed optimal among *all* Markov polices in

the original model.

(b) The proof of part (a) shows that the long-run average cost of any given (r, q) policy is the same as that in the equivalent periodic review model. It is easily verified that the steady-state inventory position after ordering is uniformly distributed on the integers $\{r + 1, \dots, r + q\}$. It follows that the long-run average cost $c(r, q)$ is given by (3.17). \square

Given the cost representation in (3.17) and since the function \hat{G} is convex, see Lemma 1, the optimal policy can be computed efficiently with the algorithm in Federgruen and Zheng (1992). Finally it is of interest to characterize how various model parameters such as the shape of the marginal inventory cost rate function $\alpha(\cdot)$, that of the marginal backlogging cost rate function $\beta(\cdot)$ and the leadtime distribution impact the optimal policy parameters r^* and $R^* \equiv r^* + q^*$. To investigate the impact of any of these model primitives θ , we write the one-step expected cost function as $\hat{G}(y|\theta)$. In the first two examples, θ is a real-valued function in the space Θ of all increasing, non-negative functions, which we endow with the partial order implied by point-wise dominance, i.e., $\theta_1 \preceq \theta_2$ iff $\theta_1(t) \leq \theta_2(t)$ for all $t \geq 0$; in the last example, θ is an element of the space Θ of all distributions of non-negative random variables, endowed with the \leq_{st} partial order. We define the function $\hat{G}(\cdot|\theta)$ to be supermodular (submodular) if it has increasing (decreasing) differences, i.e., $G(y_2|\theta) - G(y_1|\theta)$ is increasing (decreasing) in $\theta \in \Theta$ for all $y_1 < y_2$.

Proposition 3.1 (a) *When the incremental inventory cost rate function $\alpha(\cdot)$ is replaced by a new function $\hat{\alpha}(\cdot)$ that is point-wise larger, i.e., $\alpha(s) \leq \hat{\alpha}(s)$ for all $s \geq 0$, the optimal values r^* and R^* decrease.*

(b) *When the incremental backlogging cost rate function $\beta(\cdot)$ is replaced by a new function $\hat{\beta}(\cdot)$ that is point-wise larger, i.e., $\beta(s) \leq \hat{\beta}(s)$ for all $s \geq 0$, the optimal values r^* and R^* increase.*

(c) *When the leadtime distribution L^1 is replaced by $L^2 \geq_{st} L^1$, the optimal values r^* and R^* increase.*

Proof: (a) By Theorem 4.2 from Chapter 4, it suffices to show that $G(y|\alpha(\cdot))$ is

supermodular in $(y, \alpha(\cdot))$, or that $\mathbb{E}H(A_y|\alpha(\cdot))$ is supermodular in $(y, \alpha(\cdot))$. Since

$$\begin{aligned}\mathbb{E}H(A_y|\alpha(\cdot)) &= \int_0^\infty \int_0^t \alpha(s) ds dF_{A_y}(t) = \int_0^\infty \int_s^\infty \alpha(s) dF_{A_y}(t) ds \\ &= \int_0^\infty \alpha(s)(1 - F_{A_y}(s)) ds,\end{aligned}\tag{3.20}$$

We get

$$\begin{aligned}\mathbb{E}H(A_{y+1}|\alpha(\cdot)) - \mathbb{E}H(A_y|\alpha(\cdot)) &= \int_0^\infty \alpha(s)(F_{A_y}(s) - F_{A_{y+1}}(s)) ds \\ &\leq \int_0^\infty \hat{\alpha}(s)(F_{A_y}(s) - F_{A_{y+1}}(s)) ds \\ &= \mathbb{E}H(A_{y+1}|\hat{\alpha}(\cdot)) - \mathbb{E}H(A_y|\hat{\alpha}(\cdot))\end{aligned}$$

where the inequality follows from the point-wise dominance $\alpha(\cdot) \leq \hat{\alpha}(\cdot)$ and the fact that the SIL property of A_y in y , see the proof of Lemma 3.3, implies that $A_y \leq_{st} A_{y+1}$.

(b) Analogous to (a), it suffices to show that $\mathbb{E}J(A_y|\beta(\cdot))$ is submodular in $(y, \beta(\cdot))$. Since

$$\begin{aligned}\mathbb{E}J(A_y|\beta(\cdot)) &= \mathbb{E} \int_{-L}^0 \int_0^{-t} \beta(s) ds dF_{A_y}(t) \\ &= \mathbb{E} \int_0^L \int_{-L}^{-s} \beta(s) dF_{A_y}(t) ds = \mathbb{E} \int_0^L \beta(s) F_{A_y}(-s) ds,\end{aligned}$$

similar to part (a), we have

$$\begin{aligned}\mathbb{E}J(A_{y+1}|\beta(\cdot)) - \mathbb{E}J(A_y|\beta(\cdot)) &= \mathbb{E} \int_0^L \beta(s)(F_{A_{y+1}}(-s) - F_{A_y}(-s)) ds \\ &\geq \int_0^\infty \hat{\beta}(s)(F_{A_{y+1}}(s) - F_{A_y}(s)) ds \\ &= \mathbb{E}J(A_{y+1}|\hat{\beta}(\cdot)) - \mathbb{E}J(A_y|\hat{\beta}(\cdot))\end{aligned}$$

(c) Once again, it suffices to show that $G(y|L) = \mathbb{E}H(A_y|L) + \mathbb{E}J(A_y|L)$ is submodular in (y, L) . We prove this for the first term; the proof for the second term is analogous. Note that

$$\mathbb{E}H(A_{y+1}|L) - \mathbb{E}H(A_y|L) = \mathbb{E} \left(\int_{\sum_{i=1}^y X_i - L}^{\sum_{i=1}^{y+1} X_i - L} \bar{\alpha}(s) ds \right) = \mathbb{E}_{\{X_i\}} \left\{ \mathbb{E}_L \int_{\sum_{i=1}^y X_i - L}^{\sum_{i=1}^{y+1} X_i - L} \bar{\alpha}(s) ds \right\}.$$

Also, for any given realization of the renewal process of $\{X_i\}$, the function $\int_{\sum_{i=1}^y X_i - L}^{\sum_{i=1}^{y+1} X_i - L} \bar{\alpha}(s) ds$ is decreasing in L , since the function $\bar{\alpha}(s)$ is increasing in y . This implies that

$$\mathbb{E}_{L^2} \int_{\sum_{i=1}^y X_i - L^2}^{\sum_{i=1}^{y+1} X_i - L^2} \bar{\alpha}(s) ds \leq \mathbb{E}_{L^1} \int_{\sum_{i=1}^y X_i - L^1}^{\sum_{i=1}^{y+1} X_i - L^1} \bar{\alpha}(s) ds$$

whenever $L^2 \geq_{st} L^1$, thus implying that $\mathbb{E}H(A_{y+1}|L^2) - \mathbb{E}H(A_y|L^2) \leq \mathbb{E}H(A_{y+1}|L^1) - \mathbb{E}H(A_y|L^1)$. \square

3.2.2 Compound Renewal Processes

Assume now that the demands are generated by a *compound* renewal process, i.e., at the renewal demand epochs, a random quantity is demanded as opposed to a single unit. Let $\{Z_n, n = 1, 2, \dots\}$, denote the sequence of demand quantities, assumed to be i.i.d., and distributed as the random variable Z , assumed to have a finite moment generating function in some interval around zero. Since the inventory path experiences downward jumps, of random magnitude, at the demand epochs, the system can no longer be governed by an (r, q) policy. Moreover, the above “single-unit decomposition approach”, matching an ordered unit to a specific demand unit, while still possible, fails to generate the structural insights that can be obtained otherwise.

For this general class of demand processes, our analysis therefore returns to the approach developed in §3.1. We will show that an (s, S) policy acting on the inventory position is optimal under general shelf-age dependent carrying costs, as long as the inter-demand distribution X has the New-Better-than-Used (NBU) property, a property weaker than the Increasing Failure Rate (IFR) or even the Increasing Failure Rate Average (IFRA) condition and therefore shared by most commonly used distributions for inter-demand times; see below. Under Assumption 1, the same applies when general delay dependent costs are added, indeed for completely general compound processes. In both cases, the model is equivalent to a periodic review model whose single period demand is distributed as Z , with a specific convex one-step expected cost function of the inventory position.

In addition to the shelf age dependent inventory costs, assume the cost structure

includes fixed-plus-linear order costs, specified by the parameters: K = the fixed cost incurred for any order, and c = the variable per unit procurement cost. For the sake of notational simplicity, we confine ourselves to the case where the leadtime is deterministic. Extensions to stochastic leadtimes arising from exogenous and sequential processes are straightforward.

3.2.2.1 General Shelf Age Dependent Carrying Costs

As in the periodic review model, note that every unit which resides in the system's inventory at some point in time, is part of a unique order. We therefore account for all expected inventory costs by assigning to each order all the expected carrying costs resulting from this order until all of its units are sold. These total expected carrying costs are charged at the time that the order is placed. Consider therefore the n -th demand epoch, for any $n \geq 1$, at which the inventory position is increased from a level x_n to a level $y_n \geq x_n$. The order arrives L time units later and is sequentially depleted at subsequent demand epochs, depending on the prevailing inventory at the time the order arrives and demand quantities at subsequent demand epochs. The first demand epoch following the order's arrival has a distribution

$R(L) \equiv$ the excess renewal distribution at time L for a renewal process starting at time 0.

All subsequent demand epochs follow this first epoch after an inter-renewal time distributed like X . In other words, the demand epochs following the order arrival represent a delayed renewal process and we define

$$\begin{aligned} \tilde{S}_j &\equiv \text{the time between the order arrival and the } j\text{-th subsequent demand epoch} \\ &= \begin{cases} R(L) \oplus \sum_{i=1}^{j-1} X_i, & j \geq 1 \\ R(L), & j = 0 \end{cases} \end{aligned}$$

Upon delivery, the part of the order which remains after filling any prevailing backlogs, is part of the system's inventory, for at least $R(L)$ time units; similarly, the part of the order that is left after filling demands at the j -th demand epoch following the order's arrival, sees its shelf age increase from \tilde{S}_j to \tilde{S}_{j+1} . Each such unit therefore incurs additional carrying costs between the j -th and $(j+1)$ -st demand

epochs, given by $\Delta H_j \equiv \mathbb{E}[H(\tilde{S}_{j+1}) - H(\tilde{S}_j)]$ where $H(t) = \int_0^t \alpha(s)ds$ is defined in §3.2.1, above Lemma 3.3.

To show that the total expected carrying cost associated with the order can be represented as a *convex* function of y_n (, only), we need to show that the sequence of incremental carrying cost contributions $\{\Delta H_j, j \geq 0\}$ is finite and increasing. This can be shown provided that the interrenewal time X is NBU, i.e., for all $t \geq 0$, $(X - t|X > t) \leq_{st} X$.

Lemma 3.4 *Assume the interrenewal distribution X has the NBU property. Then, $\Delta H_{j-1} \leq \Delta H_j < \infty$ for $j \geq 1$.*

Finally, as in §3.1, assume that the *remaining* inventory and backlogging costs may be represented by a *convex* functions $\Gamma(y_n)$ of the inventory position after ordering. The following theorem shows that this continuous review model is equivalent to a period review model with *convex, level dependent* inventory and backlogging costs only(, in addition to the order costs):

Periodic Review Model(PRM): This model has periods of constant length $\tau = \mathbb{E}(X)$, i.i.d. demands distributed as Z , the same fixed-plus-linear costs and one step expected inventory costs expressed as a function of the inventory position after ordering. More specifically, in addition to the cost function $\Gamma(y)$, the remaining expected shelf age dependent carrying costs are represented by

$$G(y) \equiv \sum_{j=1}^{\infty} (\Delta H_j - \Delta H_{j-1}) \mathbb{E}(y - \sum_{i=1}^{N(L)+j} Z_i)^+, \quad (3.21)$$

where $N(L)$ denotes the number of demand epochs in an interval of length L following a renewal epoch.

Theorem 3.6 (Compound renewal demand processes: shelf age dependent carrying costs) *Assume that the inter-demand distribution has the NBU property.*

(a) *The function $G(\cdot)$ is finite and convex.*

(b) *The above continuous review model is equivalent to the periodic review model PRM in a sense that, for any Markov policy π in the continuous review model, the long-run average cost $C(\pi)$ is the same as that of the same policy π in the PRM*

model. In particular, the (s^*, S^*) policy that is optimal in the equivalent PRM model is also optimal in the continuous model.

Proof: (a) Define $B_j \equiv \mathbb{E}(y - \sum_{i=1}^{N(L)+j} Z_i)^+$.

Note that $B_j = \mathbb{E}[(y - \sum_{i=1}^{N(L)+j} Z_i) \mathbf{1}\{\sum_{i=1}^{N(L)+j} Z_i < y\}] \leq y \mathbb{E} \mathbf{1}\{\sum_{i=1}^j Z_i < y\} = y \text{Prob}[\sum_{i=1}^j (-Z_i) \geq -y] = y \text{Prob}[e^{\theta(\sum_{i=1}^j -Z_i)} \geq e^{-\theta y}] \leq y e^{\theta y} \psi^j(\theta)$, where $\psi(\theta) = \mathbb{E}e^{-\theta Z} < 1$ and where the last inequality follows from Markov's inequality and the fact that the sequence $\{Z_i\}$ is i.i.d. Since the coefficients ΔH_j are polynomially bounded in j , this verifies that the infinite series in (3.21) converges. The convexity of $G(\cdot)$ follows from Lemma 3.4.

(b) Similar to its definition in §3.1, we use q_n^0 to denote the part of the demand epoch n 's order $q_n = y_n - x_n$ that remains in stock after the order's arrival and clearing any outstanding backlogs; similarly q_n^j , $j \geq 1$, denotes the part that remains in stock after the j -th subsequent demand epoch. It follows from the FIFO rule that, for $j \geq 0$,

$$\begin{aligned} q_n^j &= (q_n - (\sum_{i=1}^{N(L)+j} Z_{n+i} - x_n)^+)^+ \\ &= (y_n - \sum_{i=1}^{N(L)+j} Z_{n+i})^+ - (x_n - \sum_{i=1}^{N(L)+j} Z_{n+i})^+ \\ &= (y_n - \sum_{i=1}^{N(L)+j} Z_{n+i})^+ - (y_{n-1} - \sum_{i=0}^{N(L)+j} Z_{n+i})^+. \end{aligned} \quad (3.22)$$

The second equality may be verified using the same arguments as those used to derive (3.4) and the third one follows from the identity $x_n = y_{n-1} - Z_n$. As argued above, the total expected carrying costs associated with the order at the n -th demand epoch, which, by our accounting scheme, is charged at that time, are $\sum_{j=1}^{\infty} \Delta H_j \mathbb{E} q_n^{j-1}$. (Since, for all $j \geq 0$, the quantity q_n^j only depends on the process $\{Z_i, i \geq 1\}$ while each of its units incurs incremental carrying costs that depend on the process $\{X_i, i \geq 1\}$, these two quantities are independent so that the expectation of their product equals the product of their expectations.)

Thus for any Markov policy π , the long-run average costs $C(\pi)$ are

$$\lim_{T \rightarrow \infty} \frac{\sum_{n=1}^{N(T)} \{K\delta(y_n) + c(y_n - x_n) + \Gamma(y_n) + \sum_{j=1}^{\infty} \Delta H_j \mathbb{E}(q_n^{j-1} | y_n)\}}{T}$$

$$\begin{aligned}
&= \lim_{T \rightarrow \infty} \frac{\sum_{n=1}^{N(T)} \{K\delta(y) + c(y_n - x_n) + \Gamma(y_n) + \sum_{j=1}^{\infty} \Delta H_j \mathbb{E}(q_n^{j-1} | y_n)\} / N(T)}{T/N(T)} \\
&= \tau^{-1} \lim_{N \rightarrow \infty} \frac{\sum_{n=1}^N \{K\delta(y) + c(y_n - x_n) + \Gamma(y_n) + \sum_{j=1}^{\infty} \Delta H_j \mathbb{E}(q_n^{j-1} | y_n)\}}{N} \\
&= \tau^{-1} \lim_{N \rightarrow \infty} \frac{1}{N} \left\{ \sum_{n=1}^N \{K\delta(y) + c(y_n - x_n) + \Gamma(y_n) \right. \\
&\quad \left. + \sum_{j=0}^{\infty} \Delta H_j \mathbb{E}[(y_n - \sum_{i=1}^{N(L)+j-1} Z_{n+i})^+ - (y_{n-1} - \sum_{i=0}^{N(L)+j-1} Z_{n+i})^+] \right\} \\
&= \tau^{-1} \lim_{N \rightarrow \infty} \sum_{n=1}^N \{K\delta(y) + c(y_n - x_n) + \Gamma(y_n) + \sum_{j=1}^{\infty} (\Delta H_j - \Delta H_{j-1}) \mathbb{E}(y_n - \sum_{i=1}^{N(L)+j} Z_{n+i})^+ \\
&\quad + [\sum_{j=1}^{\infty} \Delta H_j \mathbb{E}(y_{N(T)} - \sum_{i=1}^{N(L)+j-1} Z_i)^+ - \sum_{j=1}^{\infty} \Delta H_j \mathbb{E}(x_1 - \sum_{i=1}^{N(L)+j} Z_i)^+] \} / N \\
&= \tau^{-1} \lim_{N \rightarrow \infty} \sum_{n=1}^N \{K\delta(y) + c(y_n - x_n) + \Gamma(y_n) + \sum_{j=0}^{\infty} (\Delta H_j - \Delta H_{j-1}) G(y_n)\}.
\end{aligned}$$

Thus establishing the equivalency result. (The third equality follows is based on $\lim_{T \rightarrow \infty} \frac{N(T)}{T} = \tau^{-1}$ a.s.-the basic renewal theorem- and $\lim_{T \rightarrow \infty} N(T) = \infty$ while the fourth equality follows from (3.22); the last follows from the last two terms in the expression between the square brackets to its left being bound in N .) Since both $\Gamma(\cdot)$ and $G(\cdot)$ are convex, the optimality of an (s, S) policy follows immediately, as in Theorem 3.2(d), see Iglehart(1963a, 1963b). \square

3.2.2.2 General Delay Dependent Backlogging Costs

We now show that the equivalency result in Theorem 3.6 can be extended in the presence of delay dependent backlogging cost represented by a general increasing function $\beta(\cdot)$.

Let $S_n \equiv \sum_{j=1}^n X_j$ denote the time of the n -th demand epoch. We account for the delay dependent backlogging costs by charging at time S_n the expected total backlogging costs incurred in the interval $[S_n + L, S_{n+1} + L]$ and showing that these expected costs can be expressed as a convex function of y_n , the inventory position after ordering at the n -th demand epoch, S_n . As in the periodic review models in §3.1 and unlike our treatment in §3.2.1 for simple renewal processes, this

representation requires an assumption guaranteeing either that no demand unit is delayed by more than L time units or, equivalently, the inventory position after ordering is non-negative (Assumption NIP), or that the marginal backlogging cost rate is constant for delays exceeding L (Assumption CBL). We confine ourselves to Assumption (NIP).

Similar to its definition in §3.1.3, for all $j \geq 1$, let $N_{n,j}$ denote the number of units demanded at time S_{n+j} that are filled at $S_{n+1} + L$ or later. (By Assumption (NIP), all demands occurring at earlier demand epochs are filled at time $S_n + L$.) All such units are not being filled at least until time $S_{n+1} + L$; they therefore incur delay dependent backlogging costs in the interval $[S_n + L, S_{n+1} + L)$, (unless the demands are after the time $S_{n+1} + L$). It follows from the FIFO rule that,

$$\sum_{j=1}^m N_{n,j} = \left(\sum_{j=1}^m Z_{n+j} - y_n \right)^+, m \geq 1. \quad (3.23)$$

During the interval $[S_{n-1} + L, S_n + L)$, every unit in $N_{n,j}$, since demanded at time S_{n+j} , sees its delay value increase from $[S_n + L - S_{n+j}]^+$ to $S_{n+1} + L - S_{n+j}$ and therefore incurs incremental backlogging costs in the amount:

$$DJ_j \equiv \int_{[S_n+L-S_{n+j}]^+}^{S_{n+1}+L-S_{n+j}} \bar{\beta}(u) du = J(S_{n+1} + L - S_{n+j}) - J([S_n + L - S_{n+j}]^+), \quad (3.24)$$

where the function $J(t) = \int_0^t \bar{\beta}(u) du$ is defined above Lemma 3.3, in §3.2.1. (For $j \leq N(L)$, i.e., when $S_{n+j} \leq S_n + L$, the units in $N_{n,j}$ have been backlogged from before the starting point of the interval $[S_{n-1} + L, S_n + L)$ but their backlogging costs prior to reaching the delay value of $(S_n + L - S_{n+j})$ are charged to some earlier demand epochs $r < n$. For demand epochs that occur after time $S_{n+1} + L$, no backlogging costs are incurred during the time window $[S_{n-1} + L, S_n + L)$, i.e., $DJ_j = 0 = \int_{[S_n+L-S_{n+j}]^+}^{S_{n+1}+L-S_{n+j}} \bar{\beta}(u) du$, since $\bar{\beta}(\cdot) = 0$ on the negative half line.) For $j \geq 1$, let

$$\Delta J_j = \mathbb{E}DJ_j = \mathbb{E}J(S_{n+1} + L - S_{n+j}) - \mathbb{E}J([S_n + L - S_{n+j}]^+). \quad (3.25)$$

The expected backlogging costs incurred during $[S_{n-1} + L, S_n + L)$ are given by

$$\begin{aligned}
& \mathbb{E} \left[\sum_{j=1}^{\infty} DJ_j N_{n,j} \right] \\
&= \sum_{j=1}^{\infty} \mathbb{E}(DJ_j N_{n,j}) = \sum_{j=1}^{\infty} \mathbb{E}(\Delta J_j) \mathbb{E} N_{n,j} \\
&= \sum_{j=1}^{\infty} (\Delta J_j) \mathbb{E} \left(\sum_{i=1}^j N_{n,i} - \sum_{k=1}^{j-1} N_{n,i} \right) = \mathbb{E} \left[\sum_{j=1}^{\infty} (\Delta J_j - \Delta J_{j+1}) \sum_{k=1}^j N_{n,k} \right] \\
&= \mathbb{E} \left[\sum_{j=1}^{\infty} (\Delta J_j - \Delta J_{j+1}) \left(\sum_{i=1}^j Z_{n+i} - y_n \right)^+ \right].
\end{aligned}$$

The first equality follows from the Monotone Convergence Theorem. The second equality follows from DJ_i being a function of the process $\{X_i\}$ while the quantity $N_{n,j}$ is a function of the process $\{Z_i\}$, and these two processes being independent, so that $\mathbb{E}(DJ_j N_{n,j}) = \mathbb{E}(\Delta J_j) \mathbb{E} N_{n,j}$. The last equality follows from (3.23).

Following the proof of Theorem 3.6, we show that our continuous review model with general delay-dependent backlogging costs is equivalent to the following periodic review model:

Periodic Review Model (PRM): This model has periods of constant length $\tau = \mathbb{E}(X)$, i.i.d. demands distributed as Z , the same fixed-plus-linear costs and one step expected inventory costs expressed as a function of the inventory position after ordering. More specifically, in addition to the cost function $\Gamma(y)$, the remaining expected delay dependent backlogging costs are represented by

$$G(y) \equiv \sum_{j=1}^{\infty} (\Delta J_j - \Delta J_{j+1}) \mathbb{E} \left(\sum_{i=1}^j Z_{n+i} - y_n \right)^+. \quad (3.26)$$

Theorem 3.7 (Compound renewal demand processes: delay dependent backlogging costs) (a) *The function $G(\cdot)$ is finite and convex.*

(b) *Assume Assumption (NIP) applies. The above continuous review model is equivalent to the periodic review model PRM in the sense that, for any Markov policy π in the continuous review model, the long-run average cost $C(\pi)$ is the same as that of the same policy π in the PRM model. In particular, the (s^*, S^*) policy that is optimal in the equivalent PRM model is also optimal in the continuous model.*

Chapter 4

Monotonicity Properties of Stochastic Inventory Systems

We refer to Section 1.4 for an introduction of this chapter. As mentioned in Section 1.4, the long-run average cost is of the following form, depending upon whether the sample paths of the leadtime demand process are continuous or step functions:

$$c(r, q|\theta) = \frac{\lambda K + \int_r^{r+q} G(y|\theta) dy}{q}, \quad (4.1)$$

or

$$c(r, q|\theta) = \frac{\lambda K + \sum_{y=r+1}^{r+q} G(y|\theta)}{q}. \quad (4.2)$$

In both (4.1) and (4.2), λ and K represent the long-run average demand rate and the fixed cost incurred for every order batch of size q respectively. All other model primitives $\theta \in \Theta$ impact the long-run average cost exclusively via the so-called instantaneous expected cost function $G(y|\theta)$. When the long-run average cost of an (r, q) or (r, nq) policy is given by (4.1)[(4.2)], we refer to the model as the continuous [discrete] model. Since the representations in (4.1) and (4.2) are common under (r, q) or (r, nq) policies, we henceforth confine ourselves to the former, without loss of generality. In Section 1.4, we introduced (r^*, q^*) as the optimal (r, q) - policy and $R^* \equiv r^* + q^*$ as the optimal order-up-to level.

This chapter is organized as follows. In §4.1, we review the relevant literature. §4.2 specifies a variety of inventory models governed by (r, q) or (r, nq) policies

under which the long-run average cost is of the form given by (4.1) or (4.2). This includes the standard inventory model as well as others, for example, those with shelf age and delay dependent inventory costs. This section also reviews several known properties of the optimal policy parameters, required in the subsequent analysis. The monotonicity results of r^* and R^* and their various important applications are derived in §4.3, and those pertaining to q^* , in §4.4. §4.5 is devoted to the monotonicity results of the optimal cost value, while §4.6 completes the paper with general conclusions.

4.1 Literature Review

The first studies of (r, q) policies in stochastic inventory systems go back to Whitin (1953) and the influential textbook by Hadley and Whitin (1963). The latter two focused on the case where the demand process is Poisson and the cost structure consists of linear holding/backlogging costs as well as fixed and variable ordering costs. This became the generally employed continuous review inventory model. Indeed, with Poisson demands and this cost structure, (r, q) policies are easily shown to be optimal, see Zipkin (2000). In practice, many inventory systems are, indeed, governed by an (optimal) (r, q) -policy, see, for example, Bagchi et al. (1986) reporting on an air force system with more than 500,000 SKUs.

Various authors consider alternative specifications of the demand process. These include Sahin (1979) modeling demands as a compound renewal process with a continuous batch size distribution, and Bather (1966) and Puterman (1975) addressing the case where the demand process is a Brownian motion. Song and Zipkin (1993) and Song and Zipkin (1996a) consider the case where the Poisson demand process is modulated by an underlying state of the world, which evolves according to a Markov process. (These authors confine themselves to the special case of base-stock policies, where $R = r + 1$, albeit that the reorder level r may be state-dependent.) A similar model with continuous demands and general (r, q) policies is analyzed in Browne and Zipkin (1991). Here, inventory declines at a continuous demand rate,

specified as a function of the “state of the world”, which evolves according to an ergodic Markov process.

As mentioned in the Introduction, relatively little is known about how the optimal policy parameters, r^* , R^* and q^* , vary with the model primitives, even in the above mentioned basic model (with Poisson demands and linear holding/backlogging costs). Most of the literature has focused on heuristic and exact algorithms to compute the optimal policy parameters, see, for example, Federgruen and Zheng (1992) and Gallego (1998) and the references therein. The first qualitative properties were derived by Zheng (1992). As mentioned, this author proved that the optimal policy parameters are monotone in the fixed cost K . In addition, he showed that the optimal order quantity q^* and the optimal cost value are always larger than their deterministic counterparts in the EOQ model, also demonstrating that the optimality gap incurred when employing the economic order quantity, is bounded by 12.5%. Axsäter (1996) and Gallego (1998) improved this optimality gap to $(\sqrt{5} - 2)/2$ and 6.07% respectively.

The study of monotonicity properties in elementary stochastic inventory models has, for the most part, confined itself to investigations of the effect of leadtime and/or demand uncertainties and this in newsvendor systems or systems governed by base-stock policies. (As mentioned, the latter arise as a special case of (r, q) policies with $q = 1$.) Examples include Gupta and Cooper (2005), Jemaï and Karaesmen (2005) and the references therein. To our knowledge, other than Zheng (1992) only two papers have addressed monotonicity properties in systems that are governed by general (r, q) policies. Song and Zipkin (1996b) address systems with i.i.d. leadtimes, assuming the optimal reorder level is selected in conjunction with any exogenously specified order quantity. Since under i.i.d. leadtimes the long-run average cost can only be approximated, these authors investigate the impact of increased leadtime variability on the average backlog/inventory size by conducting a numerical study. As in this paper, Song et al. (2010) consider systems governed by the globally optimal (r, q) policy. They show that both r^* and R^* increase when the steady-state leadtime or leadtime demand distribution becomes stochastically

larger. We obtain this as a special corollary of our general monotonicity result under submodular instantaneous expected cost functions $G(y|\theta)$. They also show that a stochastically smaller leadtime or leadtime demand is not guaranteed to result in a lower average cost, while the less *variable* leadtime or leadtime demand distribution does. The remainder of Song et al. (2010) establishes monotonicity properties of the policy parameters under increased (leadtime) demand variability, but only under certain conditions on the model parameters.

4.2 Model and Preliminaries

We consider a single-item inventory system which is reviewed continuously and where replenishment orders can be placed at any time. Leadtimes are characterized by a stochastic process $\{L(t) : t \geq 0\}$ with $L(t)$ the leadtime experienced by an order placed at time t . We assume the process is *exogenous*, i.e., it is independent of the demand process, as well as *sequential*, i.e., $t + L(t) \leq t' + L(t')$ for all $t < t'$, with probability one. Sequential order processes guarantee the well-known simple relationship between the net inventory process $\{IN(t) : t \geq 0\}$ (the net inventory equals stock on hand minus backorders) and the inventory position process $\{IP(t) : t \geq 0\}$ (the inventory position equals the inventory level plus all outstanding orders):

$$IN(t + L(t)) = IP(t) - D[t, t + L(t)]. \quad (4.3)$$

Here $D[t_1, t_2)$ represents the cumulative demand in the time interval $[t_1, t_2)$. We refer to Zipkin(2000,§7.4) for a survey of various exogenous and sequential leadtime processes.

The cost associated with any order batch has a fixed component K and a variable component which is proportional with the order size. Since the long-run average order size per unit of time equals the long-run average demand rate λ , this variable cost component is invariant with respect to the chosen replenishment policy, and can therefore be ignored. All other cost components are assumed to accrue continuously at a rate which depends on the prevailing inventory position y . This gives rise to the above mentioned instantaneous expected cost function $G(y|\theta)$, parameterized

by $\theta \in \Theta$, which, in some applications, refers to a model parameter, and in others to a distribution, stochastic process or cost rate function. We merely assume that Θ has a partial order \preceq .

As to the shape of $G(\cdot|\theta)$, we assume the following structural properties:

(Q): $G(\cdot|\theta)$ is strictly quasi-convex with $y_0(\theta)$ as its (unique) minimum and

$$\lim_{|y| \rightarrow \infty} G(y|\theta) = \infty.$$

(Strict quasi-convexity of $G(\cdot|\theta)$ means that $G(\cdot|\theta)$ is strictly decreasing for $y < y_0(\theta)$ and strictly increasing for $y > y_0(\theta)$.) See Veinott (1965) for a review of many settings, other than the above basic linear holding and backlogging cost structure, where quasi-convexity of the $G(\cdot|\theta)$ function prevails. Rosling (2002) considers settings where, in addition to the linear holding and backlogging costs, there is a one-time cost for every demand that is backlogged and a fixed cost for any time unit at which a backlog prevails. Such cost components arise, for example, under Lagrangian relaxation, when a fill rate or ready-rate constraint is added to the model, see the Introduction. Propositions 2.2 and 2.5 in Rosling (2002) identify conditions under which $G(\cdot|\theta)$ is quasi-convex.

In the main Sections 4.3 and 4.4, we assume, in addition, one of the following properties:

(SP): $G(y|\theta)$ is supermodular in $(y, \theta) \in \mathcal{R} \times \Theta$, i.e., for any given $\theta_1 \preceq \theta_2 \in \Theta$ and $y_2 > y_1$

$$G(y_2|\theta_1) - G(y_1|\theta_1) \leq G(y_2|\theta_2) - G(y_1|\theta_2). \quad (4.4)$$

If the inequality in (4.4) is reversed, $G(y|\theta)$ is submodular in (y, θ) , a property referred to as (SB). If the inequality is strict, we refer to the condition as (SP+) or (SB+), respectively.

Continuous Models

In the continuous model, we assume that the demand process has continuous sample paths and stationary increments. Examples include demands generated by a Brownian motion or those occurring at a continuous rate determined by an underlying “state of the world” $X(t)$ where $\{X(t) : t \geq 0\}$ evolves according to an ergodic process, either a continuous-time Markov chain or a diffusion process. See

Browne and Zipkin (1991) who state, for example, that many familiar forecasting models can be represented in this way. Under any of these demand processes, there exists a long-run average demand λ , i.e., $\lim_{t \rightarrow \infty} \frac{D[0,t]}{t} = \lambda$ with probability one.

When the inventory position y varies continuously, we assume $G(\cdot|\theta)$ is continuously differentiable. Also when $G(\cdot|\theta)$ is quasi-convex, it is, for some demand processes, well-known that an (r, q) policy is optimal. For example, Bather (1966) and Puterman (1975) show that an (r, q) policy is optimal when the demand process is a Brownian motion. In addition, even when an (r, q) policy fails to be optimal, many modelers and practitioners like to restrict themselves to this class of policies. Indeed, under any of the above demand processes, the inventory position process $\{IP(t) : t > 0\}$ has a unique steady-state distribution, which is the uniform distribution on the interval $(r, r + q]$, see, in particular, Browne and Zipkin (1991). The long-run average cost under an (r, q) policy is thus given by (4.1).

We first establish that the cost function $c(r, q)$ has a unique local (and global) minimum (r^*, q^*) . Zipkin (1986) proved this result in the special case of the standard inventory model with linear holding and backlogging costs, by showing that the cost function $c(r, q)$ is jointly convex. (See Zhang (1998) for a shorter proof than that in Zipkin (1986).)

As a first step, we show in Lemma 4.1 that, for any given quantity q , the function $c(r, q|\theta)$ is strictly quasi-convex in r , therefore possessing a unique minimum $r(q|\theta) = \arg \min_r c(r, q|\theta)$. Let $R(q|\theta) \equiv r(q|\theta) + q$ denote the corresponding optimal order-up-to level. Lemma 4.1, proved in Appendix C.1, also establishes several properties of the $r(q|\theta)$ and $R(q|\theta)$ functions, as required in the subsequent analysis.

Lemma 4.1 (Continuous models with a given order quantity) *Assume (Q).*

(P0) *For any given q , $c(r, q|\theta)$ is strictly quasi-convex in r achieving its minimum at $r(q|\theta)$.*

(P1) *$r(q|\theta)$ is the unique root of the equation $G(r|\theta) = G(r + q|\theta)$.*

(P2) *$r(q|\theta) < y_0(\theta) < R(q|\theta)$.*

(P3) *$r(q|\theta)$ is decreasing and differentiable in q with $\lim_{q \rightarrow \infty} r(q|\theta) = -\infty$.*

(P4) $R(q|\theta)$ is increasing and differentiable in q with $\lim_{q \rightarrow \infty} R(q|\theta) = \infty$.

Theorem 4.1, again proven in Appendix C.1, now show that just like there is a unique reorder level $r(q|\theta)$ for any given order quantity q , there exists a unique globally optimal order quantity $q^*(\theta)$ as well¹. This result is established by showing that the projected function $c(q|\theta) \equiv \min_r c(r, q|\theta)$ is strictly quasi-convex as well.

Theorem 4.1 (Continuous models: optimal model parameters) *Assume (Q). The function $c(q|\theta)$ is strictly quasi-convex in q . In particular, there exists a unique pair $(r^*(\theta), q^*(\theta))$, with $r^*(\theta) = r(q^*(\theta)|\theta)$, that minimizes the cost function $c(r, q|\theta)$. Moreover,*

(P5) $q \leq (=) q^*(\theta)$ if and only if $G(r(q|\theta)|\theta) \leq (=) c(q|\theta)$.

Define $g(y|\theta) \equiv \partial G(y|\theta)/\partial y$. The optimal policy parameters $r^*(\theta)$ and $q^*(\theta)$ are uniquely determined by:

$$G(r|\theta) = G(R|\theta), \quad (4.5)$$

$$\int_r^R yg(y|\theta)dy = K\lambda. \quad (4.6)$$

(4.5) follows from (P1) in Lemma 4.1 while (P5) in Theorem 4.1 implies that

$$\begin{aligned} G(r^*(\theta)|\theta) &= c(r^*(\theta), q^*(\theta)|\theta) = \frac{K\lambda + \int_{r^*(\theta)}^{R^*(\theta)} G(y|\theta)dy}{q^*(\theta)} \\ &= \frac{K\lambda + q^*(\theta)G(r^*(\theta)|\theta) - \int_{r^*(\theta)}^{R^*(\theta)} yg(y|\theta)dy}{q^*(\theta)}. \end{aligned}$$

(The last equality is obtained by integration by parts, using (4.5); (4.6) follows by subtracting $G(r^*(\theta)|\theta)$ from both sides of the equation.)

Even more common than the above demand processes with continuous sample paths, is the characterization of demands as a Poisson process. Under the above cost structure, (r, q) policies continue to be optimal and their long-run average cost value is given by (4.2), the discrete analogue of (4.1). More generally, if demand

¹Rosling (1999) proves a related but distinct property of the function $c(r, q)$, namely its pseudo-convexity.

is generated by a compound Poisson process, the reorder level r is likely to be overshoot, requiring an adaptation of the (r, q) policy structure. A frequently used class of policies are so-called (r, nq) policies: here, a replenishment order is placed whenever the inventory position drops to or *below* r . A sufficiently large multiple of q units is ordered to bring the inventory position back to the $(r, r + q]$ interval. If the compounding distribution is continuous, the long-run average cost under an (r, nq) policy is, again, given by (4.1). Alternatively, if the compounding distribution is discrete, it is given by (4.2). (Under compound Poisson demands, (r, nq) policies are suboptimal, but the best (r, nq) policy is typically close to optimal, see Zipkin (2000, p.228).)

Discrete Models

In the discrete case, we assume the same structural properties (Q) of $G(\cdot|\theta)$. Here, we redefine $r(q|\theta) \equiv \min\{r : c(r, q|\theta) = c(q|\theta)\}$ due to the possibility of multiple minimizers². Similarly, define $c^*(\theta) \equiv \min_q c(r(q|\theta), q|\theta)$ and $q^*(\theta) \equiv \min\{q : c(r(q|\theta), q|\theta) = c^*(\theta)\}$. Let $G_l(\theta)$ denote the l -th smallest value of $G(\cdot|\theta)$ function attained over all integers. (Clearly $G_1(\theta) = G(y_0|\theta)$.) The following lemma, proven in Appendix C.1 exhibits the discrete analogue of properties (P1)-(P5) of the continuous model:

Lemma 4.2 (Properties of discrete models) *Assume (Q). For any $q > 0$,*

$$\text{(P1')} \quad r(q|\theta) = \min\{r : G(r + 1 + q|\theta) \geq G(r + 1|\theta)\}.$$

$$\text{(P2')} \quad r(q|\theta) < y_0 \leq R(q|\theta).$$

$$\text{(P3')} \quad r(q|\theta) \text{ is decreasing in } q \text{ with } \lim_{q \rightarrow \infty} r(q|\theta) = -\infty.$$

$$\text{(P4')} \quad R(q|\theta) \text{ is increasing in } q \text{ with } \lim_{q \rightarrow \infty} R(q|\theta) = \infty.$$

$$\text{(P5')} \quad q < q^*(\theta) \iff G_{q+1}(\theta) < c(q|\theta). \text{ Moreover, } q^*(\theta) = \min\{q : G_{q+1}(\theta) \geq c(q|\theta)\}.$$

²Unlike in the continuous case, even under assumption (Q), there may be two optimal r values for a given value of q . For example, when y_0 is an integer and $G(y_0 - d) = G(y_0 + d)$ for all $d > 0$, both $r = y_0 - d$ and $r = y_0 - d - 1$ are optimal for $q = 2d$.

As mentioned in the Introduction, the model primitives can be partitioned into three categories, depending upon whether they impact (i) the first term in the numerator of the cost objective (4.1) or (4.2), (ii) the second term in the objective, or (iii) both terms:

(I) The fixed order cost K impacts *only* the first term. It is known that

$$r^* \text{ is decreasing in } K \text{ while } R^* \text{ and } q^* \text{ are increasing in } K. \quad (4.7)$$

See Zheng (1992) for proofs in the continuous model, and Federgruen and Zheng (1992) for an easy verification of the results in the discrete model.

(II) The average demand rate λ : this parameter changes *both* terms of the numerator of (4.1) and (4.2), and, as we shall see, the two changes sometimes impact the optimal policy parameters in opposite ways.

(III) All other model primitives θ impact the second term of the numerator of (4.1) and (4.2).

In the remainder of the paper, we focus on monotonicity properties of the policy parameters with respect to all these model primitives θ . After deriving our general results, we apply them to the primitives of the following special class of inventory models:

The Standard Inventory Model.

In the standard inventory model the following assumptions are made: the item is obtained at a given price per unit. Leadtimes are generated by a so-called exogenous and sequential process, ensuring that consecutive orders do not cross and that leadtimes are independent of the demand process. Inventory costs are accrued at a rate $h(\cdot)$ which is a convex increasing function of the inventory level with $h(0) = 0$; stockouts are backlogged where backlogging costs are, again, accrued at a rate $b(\cdot)$ which is a convex increasing function of the backlog size with $b(0) = 0$; (The basic model introduced by Hadley and Whitin (1963) with inventory and backlogging costs that are proportional with the inventory level and backlog size, respectively, is a special case of the standard inventory model.) (4.3) is used to account for these costs by assigning to epoch t the *expected* holding and backlogging costs to be incurred a leadtime $L(t)$ later, thus representing these cost components as a function

of the prevailing inventory *position* at time t . Here

$$G(y|\theta) = \mathbb{E}h((y - D_L)^+) + \mathbb{E}b((D_L - y)^+) \quad (4.8)$$

where D_L has the steady-state leadtime demand distribution with pdf $f(\cdot)$, cdf $F(\cdot)$ and ccdf $\bar{F}(\cdot)$.

There are, clearly, several other classes of inventory models that are optimally governed by an (r, q) or (r, nq) policy with the cost objective given by (4.1) or (4.2). One example is

Inventory Models with Shelf Age and Delay Dependent Inventory Costs.

In some systems, inventory cost rates depend on the so-called *shelf age* of a unit, i.e., the amount of time a unit has resided in a firm's inventory. This situation is prevalent under a variety of supplier financing schemes. One example is provided by trade credit arrangements under which a supplier allows the firm to defer payments of any given item until such time when it is sold to the consumer. The supplier may offer an initial interest-free period (e.g., 30 days) after which interest accumulates. Moreover, interest rates often increase as a function of the item's shelf age. These trade credit schemes have been considered in Gupta and Wang (2009) as well as Chapter 2. We refer to the latter for a discussion of how prevalent this practice is. Another setting with shelf age dependent inventory cost rates is when the supplier subsidizes part of the inventory cost. For example, in the automobile industry manufacturers pay the dealer so-called "holdbacks", i.e., a given amount for each month a car remains in the dealer's inventory, up to a given time limit (see, e.g., Nagarajan and Rajagopalan (2008)). The resulting inventory cost rate for any stocked item is, again, an increasing function of the item's shelf age.

Similar to shelf age dependent holding costs, backlogging cost rates may also depend on the amount of time by which delivery of a demand unit is delayed. This may reflect the structure of contractually agreed upon penalties for late delivery or, in case of *implicit* backlogging costs, the fact that customers become less or more impatient as they wait longer for an item. See Huh et al. (2010).

Let $\alpha(t)$ denote the *incremental* inventory cost rate incurred for an item that has a shelf age t and $\beta(t)$ the incremental backlogging cost rate when a unit of

demand has been waiting for t time units. Note that this class of inventory models and the standard inventory models above are distinct: the only models that are common to both classes have $\alpha(\cdot)$ and $\beta(\cdot)$ as constant functions, or equivalently $h(\cdot)$ and $b(\cdot)$ as linear functions. When the demand process is a renewal process, 3.5 has shown that the long-run average cost is of the form given by (4.2). Moreover, when the functions $\alpha(\cdot)$ and $\beta(\cdot)$ are general, merely assumed to be increasing, the instantaneous expected cost function $G(\cdot)$ has been shown to be convex *ibid.*

4.3 Monotonicity of r^* and R^*

In this section, we show that the supermodularity (submodularity) condition (SP) [(SB)] suffices to establish that the optimal reorder level r^* and order-up-to level R^* are decreasing [increasing] in $\theta \in \Theta$, both under an arbitrary order quantity and under the optimal order quantity.

Theorem 4.2 (Monotonicity of r^* and R^*) *Consider both the continuous and discrete models. Assume (Q).*

(a) *Under (SP) [(SB)], $r(q|\theta)$ and $R(q|\theta)$ are decreasing [increasing] in θ for any fixed $q > 0$.*

(b) *Under (SP) [(SB)], $r^*(\cdot)$ and $R^*(\cdot)$ are decreasing (increasing) in θ .*

Proof: The proof for the continuous model relies on the five properties (P1)-(P5) in Lemma 4.1 and Theorem 4.1. We omit the proof for the discrete model since it is analogous with each of the five properties replaced by its discrete counterpart (P1')-(P5') listed in Lemma 4.2. We also confine ourselves to the case where $G(y|\theta)$ is supermodular; the proof for the submodular case is similar.

(a) Fix $q > 0$. Apply (4.4) to $y_1 = r(q|\theta_1)$ and $y_2 = R(q|\theta_1)$ and use (P1) to conclude that

$$0 = G(y_2|\theta_1) - G(y_1|\theta_1) \leq G(y_2|\theta_2) - G(y_1|\theta_2). \quad (4.9)$$

This implies that

$$r(q|\theta_1) \geq r(q|\theta_2). \quad (4.10)$$

(Assume to the contrary that $r(q|\theta_1) < r(q|\theta_2)$, so that $R(q|\theta_2) = r(q|\theta_2) + q > y_2$; then $G(r(q|\theta_2)|\theta_2) < G(y_1|\theta_2) = G(y_2|\theta_2) < G(R(q|\theta_2)|\theta_2)$ where the first and third inequalities follow from (P2) and the strict quasi-convexity of $G(\cdot|\theta)$ while the second inequality follows from (4.9). However, this contradicts (P1) for $\theta = \theta_2$.) It follows immediately that $R(q|\theta) = r(q|\theta) + q$ is also decreasing in θ .

(b) We first show that $r^*(\theta)$ is decreasing. Assume to the contrary that $r^*(\theta_1) < r^*(\theta_2)$. Then we must have $q^*(\theta_1) > q^*(\theta_2)$ because, otherwise, $r^*(\theta_1) = r(q^*(\theta_1)|\theta_1) \geq r(q^*(\theta_2)|\theta_1) \geq r(q^*(\theta_2)|\theta_2) = r^*(\theta_2)$ where the first inequality follows from (P3) and the second from (4.10), which contradicts our assumption $r^*(\theta_1) < r^*(\theta_2)$. In view of (P3), let \hat{q} be the unique order quantity such that $r(\hat{q}|\theta_1) = r^*(\theta_2)$. Since, by (4.10), $r(q^*(\theta_2)|\theta_1) \geq r(q^*(\theta_2)|\theta_2) = r^*(\theta_2) = r(\hat{q}|\theta_1)$, (P3) implies:

$$q^*(\theta_2) \leq \hat{q}. \quad (4.11)$$

On the other hand, since by our assumption, $r(\hat{q}|\theta_1) = r^*(\theta_2) > r^*(\theta_1) = r(q^*(\theta_1)|\theta_1)$, it follows, again, from (P3) that

$$q^*(\theta_1) > \hat{q}. \quad (4.12)$$

By (P5), $G(r^*(\theta_2)) = c^*(q^*(\theta_2)|\theta_2) = \frac{K\lambda + \int_{r^*(\theta_2)}^{r^*(\theta_2)+q^*(\theta_2)} G(y|\theta_2)dy}{q^*(\theta_2)}$, or

$$\int_{r^*(\theta_2)}^{r^*(\theta_2)+q^*(\theta_2)} [G(r^*(\theta_2)|\theta_2) - G(y|\theta_2)]dy = \lambda K. \quad (4.13)$$

On the other hand, by (4.12) and (P5), $G(r^*(\theta_2)|\theta_1) = G(r(\hat{q}|\theta_1)|\theta_1) < c^*(\hat{q}|\theta_1)$, or

$$\int_{r^*(\theta_2)}^{r^*(\theta_2)+\hat{q}} [G(r^*(\theta_2)|\theta_1) - G(y|\theta_1)]dy < \lambda K. \quad (4.14)$$

Moreover,

$$0 \leq G(r^*(\theta_2)|\theta_2) - G(y|\theta_2) \leq G(r^*(\theta_2)|\theta_1) - G(y|\theta_1) \quad (4.15)$$

for all $r^*(\theta_2) \leq y \leq r^*(\theta_2) + q^*(\theta_2)$ where the first inequality follows from (P1) and the quasi-convexity of $G(\cdot)$ and the second inequality from (4.4). Therefore,

$$\begin{aligned} \lambda K &= \int_{r^*(\theta_2)}^{r^*(\theta_2)+q^*(\theta_2)} [G(r^*(\theta_2)|\theta_2) - G(y|\theta_2)]dy \\ &\leq \int_{r^*(\theta_2)}^{r^*(\theta_2)+q^*(\theta_2)} [G(r^*(\theta_2)|\theta_1) - G(y|\theta_1)]dy \\ &\leq \int_{r^*(\theta_2)}^{r^*(\theta_2)+\hat{q}} [G(r^*(\theta_2)|\theta_1) - G(y|\theta_1)]dy < \lambda K, \end{aligned}$$

where the first equality follows from (4.13), the first inequality from (4.15), the second from (4.15) as well as (4.11), and the last from (4.14). This contradicts our assumption $r^*(\theta_1) < r^*(\theta_2)$.

The proof that $R^*(\cdot)$ is decreasing also proceeds by contradiction. Assume, therefore, that $R^*(\theta_1) < R^*(\theta_2)$. By (P4), let \tilde{q} be the unique order level such that $R(\tilde{q}|\theta_2) = R^*(\theta_1) < R^*(\theta_2)$. It follows from (P4) that $\tilde{q} < q^*(\theta_2)$. This, together with (P5) and the definition of \tilde{q} , implies that $G(R^*(\theta_1)|\theta_2) = G(R(\tilde{q}|\theta_2)|\theta_2) < c^*(\tilde{q}|\theta_2)$, or

$$\int_{r(\tilde{q}|\theta_2)}^{R^*(\theta_1)} [G(R^*(\theta_1)|\theta_2) - G(y|\theta_2)] dy < \lambda K. \quad (4.16)$$

On the other hand, using the definition of \tilde{q} and part (a), $R^*(\theta_1) = R(\tilde{q}|\theta_2) \leq R(\tilde{q}|\theta_1)$. Hence, by (P4), $\tilde{q} \geq q^*(\theta_1)$. This, by using the definition of \tilde{q} once again, implies that

$$r(\tilde{q}|\theta_2) = R(\tilde{q}|\theta_2) - \tilde{q} = R^*(\theta_1) - \tilde{q} \leq R^*(\theta_1) - q^*(\theta_1) = r^*(\theta_1). \quad (4.17)$$

Therefore,

$$\begin{aligned} \lambda K &= \int_{r^*(\theta_1)}^{R^*(\theta_1)} [G(R^*(\theta_1)|\theta_1) - G(y|\theta_1)] dy \leq \int_{r^*(\theta_1)}^{R^*(\theta_1)} [G(R^*(\theta_1)|\theta_2) - G(y|\theta_2)] dy \\ &\leq \int_{r(\tilde{q}|\theta_2)}^{R^*(\theta_1)} [G(R^*(\theta_1)|\theta_2) - G(y|\theta_2)] dy < \lambda K \end{aligned} \quad (4.18)$$

Here the first equality follows from $G(R^*(\theta_1)) = c^*(q^*(\theta_1)|\theta_1) = \frac{K\lambda + \int_{r^*(\theta_1)}^{R^*(\theta_1)} G(y|\theta_1) dy}{q^*(\theta_1)}$ using (P5). The first inequality is due to the supermodularity of $G(y|\theta)$ (see (4.4)); the second inequality uses (4.17) and the fact that $G(R^*(\theta_1)|\theta_2) \geq G(y|\theta_2)$ for all $r(\tilde{q}|\theta_2) \leq y \leq R(\tilde{q}|\theta_2) = R^*(\theta_1)$, applying (P1) with $\theta = \theta_2$ and $q = \tilde{q}$ and quasi-convexity of $G(\cdot|\theta_2)$; finally, the last inequality follows from (4.16). The contradiction in (4.18) implies that our assumption $R^*(\theta_1) < R^*(\theta_2)$ is false. \square

Corollary 4.1 applies Theorem 4.2 to the standard inventory model.

Corollary 4.1 (Monotonicity properties of r^* and R^* in the standard inventory model) *Assume the instantaneous expected cost function $G(\cdot|\theta)$ is of the form (4.8), in both the continuous and the discrete model. Let Ξ denote the set of increasing*

convex and positive functions endowed with a partial order \preceq . For any $v_1(\cdot), v_2(\cdot) \in \Xi$, $v_1(\cdot) \preceq v_2(\cdot)$ means $v_1'(y) \leq v_2'(y)$ for all y .

- (a) r^* and R^* are decreasing in the holding cost rate function $h(\cdot) \in \Xi$.
- (b) r^* and R^* are increasing in the backlogging cost rate function $b(\cdot) \in \Xi$.
- (c) If, for any given y , the incremental holding cost rate $h'(y)$ is an increasing differentiable function of the purchase cost w , (e.g. $h(y) = (\alpha_0 w + h_0)y$ with $\alpha_0 > 0$ being the capital cost rate and h_0 the physical inventory cost rate,) then r^* and R^* are decreasing in w .
- (d) Consider two stochastically ordered steady-state leadtime demand distributions D_L^1 and D_L^2 with $D_L^1 \leq_{st} D_L^2$, however with the same average demand rate λ . Then r^* and R^* are larger under D_L^2 than those under D_L^1 .
- (e) Assume the demand process is compound Poisson. Consider two leadtime processes $\{L^1(t) : t \geq 0\}$ and $\{L^2(t) : t \geq 0\}$ such that $L^1(t) \leq_{st} L^2(t)$ for all $t > 0$. r^* and R^* are larger under the $\{L^2(t) : t \geq 0\}$ process than those under the $\{L^1(t) : t \geq 0\}$ process.
- (f) Consider a compound Poisson demand process. R^* is increasing in the Poisson rate ν .
- (g) Consider two demand processes, both of which are compound Poisson with Poisson rate ν , but with different compounding distributions Z^1 and Z^2 where $Z^1 \leq_{st} Z^2$. Then, R^* is larger under the second demand process as compared to the first.

Proof: Note that $g(y|\theta) \equiv \partial G(y|\theta)/\partial y = \mathbb{E}h'((y - D_L)^+) - \mathbb{E}b'((D_L - y)^+)$.

(a) Let $\theta = h(\cdot)$. The result follows immediately from part (b) of Theorem 4.2 as $g(y|h_1(\cdot)) \leq g(y|h_2(\cdot))$ for any $h_1(\cdot), h_2(\cdot) \in \Xi$ with $h_1(\cdot) \preceq h_2(\cdot)$.

(b) Let $\theta = b(\cdot)$. Similar to part (a), the result follows immediately from part (b) of Theorem 4.2 as $g(y|b_1(\cdot)) \geq g(y|b_2(\cdot))$ for any $b_1(\cdot), b_2(\cdot) \in \Xi$ with $b_1(\cdot) \preceq b_2(\cdot)$.

(c) Let $\theta = w$. The result is easily verified by noting that $\partial g(y|\theta)/\partial\theta = \mathbb{E} \left[\frac{\partial h'((y-D_L)^+)}{\partial w} \right] \geq 0$.

(d) Let $\theta = D_L$ and $\tilde{h}(x) = h(x^+) + b(x^-)$. Note that $\tilde{h}(\cdot)$ is convex and differentiable for all $x \neq 0$ with $\tilde{h}'(\cdot)$ an increasing function. $D_L^1 \leq_{st} D_L^2 \iff -D_L^1 \geq_{st} -D_L^2 \Rightarrow g(y|D_L^2) = \mathbb{E}\tilde{h}'(y - D_L^2) \leq \mathbb{E}\tilde{h}'(y - D_L^1) = g(y|D_L^1)$. Thus, the result follows from part (b) of Theorem 4.2.

(e) Song (1994) showed that the leadtime demand under $\{L^2(t) : t \geq 0\}$ is stochastically greater than that under $\{L^1(t) : t \geq 0\}$ when $L^1(t) \leq_{st} L^2(t)$. The result follows from part (d).

(f) Note that the average demand rate is

$$\lambda = \nu \mathbb{E}(Z) \tag{4.19}$$

where Z denotes the compounding distribution. While the parameters λ and ν are tied together by (4.19), consider for an arbitrary parameter pair $(\lambda, \nu) \in \mathcal{R}_+^2$, the expanded family of long-run average cost functions:

$$c(r, q|\lambda, \nu) = \frac{\lambda K + \int_r^{r+q} G(y|\nu) dy}{q}, \text{ and } c(r, q|\lambda, \nu) = \frac{\lambda K + \sum_{y=r+1}^{r+q} G(y|\nu)}{q}, \tag{4.20}$$

in the continuous and discrete model respectively. It follows from (4.7) that an independent increase in λ , with ν fixed, results in an increase in R^* . Similarly, by part (d), an independent increase of ν , with λ fixed, also results in an increase of R^* since D_L is stochastically increasing in ν by Theorem 1.A.5 of Shaked and Shanthikumar (2007), using the fact that a Poisson random variable is stochastically increasing in its mean. The same therefore applies as (λ, ν) are increased along the line described by (4.19), i.e., R^* is increasing in the Poisson rate ν .

(g) Similar to part (f), consider the expanded family of long-run average cost function $c(r, q|\lambda, Z)$. Since $\lambda_1 = \nu \mathbb{E}(Z^1) \leq \nu \mathbb{E}(Z^2) = \lambda_2$ and, again, by Theorem 1.A.5 of Shaked and Shanthikumar (2007) $D_L^1 \leq_{st} D_L^2$, the result follows from (4.7) and part (d). \square

To our knowledge, parts (a)-(c), (f) and (g) are new results. For the case of compound Poisson demand process, parts (d)-(e) are shown in Song et al. (2010),

see Lemma 4, Theorem 2 and discussions in §5.1 there. We include these parts to provide a comprehensive treatment of various primitives in the standard inventory model and to show that these monotonicity properties as well as those covered in (a)-(c), (f) and (g) all follow as direct corollaries of the general result in Theorem 4.2.

Summarizing parts (e)-(g), there are three ways in which the leadtime demand distribution under a compound Poisson demand process may increase (stochastically): (i) a stochastic increase in the leadtime distribution, (ii) an increase in the rate of the Poisson process and (iii) a stochastic increase in the compounding distribution. In the first case (i), both r^* and R^* increase. In the remaining cases, Corollary 4.1 shows the same monotonicity property for R^* . As to the impact on r^* , consider the cost representation given by (4.20) above. Following the proof of Corollary 4.1(f), an increase of ν results in increase of r^* ; an increase of λ results in an decrease of r^* , similar to the impact of an increase in K , see (4.7). Since λ and ν are tied together by equation (4.19), the stochastic increase in the demand process has two opposing effects. Our numerical explorations reveal that the *net* effect consists of an *increase* of r^* , (similar to the proven increase in R^*), but a formal proof is outstanding.

Finally, the following result has been shown in Proposition 3.1, addressing inventory models with general shelf age and delay dependent inventory cost rate functions, as described in Section 4.2.

Proposition 4.1 (Shelf age and delay dependent inventory costs) *Consider the shelf age and delay dependent inventory cost model of Section 4.2.*

(a) *When the incremental inventory cost rate function $\alpha(\cdot)$ is replaced by a new function $\hat{\alpha}(\cdot)$ that is point-wise larger, i.e., $\alpha(s) \leq \hat{\alpha}(s)$ for all $s \geq 0$, the optimal values r^* and R^* decrease.*

(b) *When the incremental backlogging cost rate function $\beta(\cdot)$ is replaced by a new function $\hat{\beta}(\cdot)$ that is point-wise larger, i.e., $\beta(s) \leq \hat{\beta}(s)$ for all $s \geq 0$, the optimal values r^* and R^* increase.*

4.4 Monotonicity of q^*

In this section, we identify monotonicity properties of the remaining policy parameter q^* , i.e., the optimal order quantity, with respect to various model parameters. These results require an additional property of the instantaneous expected cost function $G(y|\theta)$ beyond the simple supermodularity (SP) or submodularity property (SB) guaranteeing the monotonicity of r^* and R^* , see Theorem 4.2. To begin with, we confine ourselves to monotonicity properties with respect to a *scalar* model parameter θ , as opposed to more complex entities (demand distributions, leadtime distributions, cost rate functions etc.) addressed in §4.3. (In other words, in this section, we assume $\Theta \subseteq \mathcal{R}$.) We confine ourselves to the case where the function $G(y|\theta)$ is supermodular, i.e., (SP) prevails without giving explicit treatment to the parallel case (SB). Note that if the function $G(y|\theta)$ is submodular in (y, θ) , a supermodular structure is obtained by replacing the parameter θ by any decreasing transformation thereof, for instance, $\theta' = -\theta$ or $\theta' = \theta^{-1}$. As explained in the Introduction, the monotonicity properties of q^* differ between the continuous and discrete model. In subsection 4.4.1, we devote ourselves to the continuous model, while the discrete model is addressed in subsection 4.4.2.

4.4.1 The Continuous Model

In the continuous model, we henceforth assume that the function $G(y|\theta)$ is twice differentiable in (y, θ) . Thus let $g(y|\theta) \equiv \frac{\partial G(y|\theta)}{\partial y}$, $g_y(y|\theta) \equiv \frac{\partial g(y|\theta)}{\partial y}$, and $g_\theta(y|\theta) \equiv \frac{\partial g(y|\theta)}{\partial \theta}$.

We first derive a *necessary* and *sufficient* condition for q^* to be increasing or decreasing in θ , the proof of which is given in Appendix C.1.

Theorem 4.3 (Monotonicity properties of q^* in the continuous model) *Assume (Q) with $G(\cdot|\cdot)$ a twice differentiable function. $q^*(\theta)$ is decreasing (increasing) at $\theta = \theta_0$ if and only if*

$$\int_{r^*(\theta_0)}^{R^*(\theta_0)} yg_y(y|\theta_0)dy \int_{r^*(\theta_0)}^{R^*(\theta_0)} g_\theta(y|\theta_0)dy \leq (>) \int_{r^*(\theta_0)}^{R^*(\theta_0)} g_y(y|\theta_0)dy \int_{r^*(\theta_0)}^{R^*(\theta_0)} yg_\theta(y|\theta_0)dy. \quad (4.21)$$

In some cases, the necessary and sufficient condition (4.21) for q^* to be increasing or decreasing at any given θ may be verified directly. In other cases, it may be difficult or impossible to compute the integrals in a closed form. We therefore derive a sufficient condition, in Theorem 4.4 below, which is both more transparent and can be verified by computing the partial derivatives of the function $g(y|\theta)$ only. First we need the following lemma.

Lemma 4.3 *Let $u(\cdot)$ and $v(\cdot)$ be two continuous functions defined on the closed interval $[a, b] \subset \mathcal{R}$, with $a < b$ and $v(y) \neq 0$ for any $y \in [a, b]$. Assume $\rho(y) \equiv u(y)/v(y)$ is decreasing (increasing) in $y \in [a, b]$. Then $\int_a^b yv(y)dy \int_a^b u(y)dy \geq (\leq) \int_a^b v(y)dy \int_a^b yu(y)dy$. The inequality is strict unless the ratio function $\rho(\cdot)$ is constant.*

Proof: We give the proof for the case where $\rho(\cdot)$ is decreasing and $v(y) > 0$ for all $y \in [a, b]$; the proof for the three alternative cases is similar. Since $v(y) \in \mathcal{R}^+$,

$$\begin{aligned} & \int_a^b yv(y)dy \int_a^b u(y)dy \geq \int_a^b v(y)dy \int_a^b yu(y)dy \\ \iff & \frac{\int_a^b yv(y)dy}{\int_a^b v(y)dy} \int_a^b \rho(y)v(y)dy \geq \int_a^b y\rho(y)v(y)dy \\ \iff & \int_a^b \rho(y)\delta(y)v(y)dy \geq 0 \end{aligned} \tag{4.22}$$

where $\delta(y) = \frac{\int_a^b yv(y)dy}{\int_a^b v(y)dy} - y$. It is thus equivalent to show (4.22). Note that $\delta(\cdot)$ is a decreasing linear function of y . Moreover,

$$\int_a^b \delta(y)v(y)dy = 0, \tag{4.23}$$

precluding the case that $\delta(y)$ is uniformly positive or uniformly negative on $[a, b]$. In other words, $\delta(y) = 0$ has a unique root $c \in [a, b]$ with $\delta(y) \geq 0$ for $y \in [a, c]$ and $\delta(y) \leq 0$ for $y \in [c, b]$. As $v(y) \geq 0$ for all $y \in [a, b]$,

$$\delta(y)v(y) \geq 0 \text{ on } [a, c] \text{ and } \leq 0 \text{ on } [c, b]. \tag{4.24}$$

Hence,

$$\begin{aligned}
& \int_a^b \rho(y)\delta(y)v(y)dy = \int_a^c \rho(y)\delta(y)v(y)dy + \int_c^b \rho(y)\delta(y)v(y)dy \\
& \geq \rho(c) \int_a^c \delta(y)v(y)dy + \rho(c) \int_c^b \delta(y)v(y)dy \\
& = \rho(c) \int_a^b \delta(y)v(y)dy = 0,
\end{aligned} \tag{4.25}$$

verifying (4.22). (The inequality uses the fact that $\rho(y)$ is decreasing along with (4.24). The last equality follows from (4.23).) Note that the inequality in (4.25) is strict unless $\rho(\cdot)$ is constant. \square

Theorem 4.4 (Monotonicity of q^* in the continuous model: a simple sufficient condition) *Assume (Q) with $G(\cdot|\cdot)$ a twice differentiable function. Let (SP+) hold on a rectangle $[m, M] \times [\theta_1, \theta_2]$ with $\theta_1 < \theta_2$ and $m \leq r^*(\theta) < R^*(\theta) \leq M$ for all $\theta \in [\theta_1, \theta_2]$. $q^*(\cdot)$ is decreasing (increasing) on $[\theta_1, \theta_2]$ if $g_y(y|\theta)/g_\theta(y|\theta)$ is decreasing (increasing) in $y \in [m, M]$.*

Proof: Fix $\theta_0 \in [\theta_1, \theta_2]$. Apply Lemma 4.3 with $[a, b] = [r^*(\theta_0), R^*(\theta_0)]$, $v(y) = g_\theta(y|\theta_0)$ and $u(y) = g_y(y|\theta_0)$. $v(y) > 0$ since $G(y|\theta_0)$ is strictly supermodular. Lemma 4.3 shows that the necessary and sufficient condition in Theorem 4.3 for $q^*(\theta)$ to be increasing (decreasing) at θ_0 is satisfied when $g_y(y|\theta_0)/g_\theta(y|\theta_0)$ is increasing (decreasing) in y . \square

Remark 1. The following provides an alternative sufficient condition, the proof of which is analogous to that of Theorem 4.4: Assume $G(\cdot|\theta)$ is strictly convex in y on an interval $[m, M]$ with $m \leq r^*(\theta) < R^*(\theta) \leq M$. $q^*(\theta)$ is decreasing (increasing) if $g_\theta(y|\theta)/g_y(y|\theta)$ is increasing (decreasing) in $y \in [m, M]$.

We now apply the general monotonicity results in Theorems 4.3 and 4.4 to the standard inventory model with $G(y|\theta)$ specified in (4.8). We focus on the case of linear holding and backlogging costs. Corollary 4.1 showed that r^* and R^* are monotone in the holding cost rate function $h(\cdot)$ and backlogging cost rate function $b(\cdot)$, irrespective of the shape of the leadtime demand distribution. Corollary 4.2 below shows, inter alia, that, as far as the optimal order quantity q^* is concerned, its dependency on, say the backlogging cost rate b , *does* depend on the shape of the

leadtime demand distribution. q^* decreases in b when the complementary cdf of the leadtime distribution is log-concave, i.e., the logarithm of this function is concave, or equivalently when the distribution is IFR (Increasing Failure Rate); conversely, when the distribution is DFR (Decreasing Failure Rate), q^* is increasing in b .

Corollary 4.2 (Monotonicity properties of q^* in the continuous standard inventory model) *Assume the instantaneous expected cost function $G(\cdot|\theta)$ is of the form (4.8) with constant marginal inventory and backlogging cost rates h and b . The continuous leadtime demand D_L has support $[d_l, d_h]$ where $-\infty \leq d_l < d_h \leq \infty$.*

(a) q^* is decreasing in h if $F(y)$ is log-concave on $[d_l, d_h]$; q^* is increasing in h if $F(y)$ is log-convex on $[d_l, d_h]$ with $d_h = \infty$.

(b) q^* is decreasing in b if $\bar{F}(y)$ is log-concave on $[d_l, d_h]$; q^* is increasing in b if $\bar{F}(y)$ is log-convex with $r^*(b) \geq d_l$.

(c) Assume D_L is Normal with mean μ_L and standard deviation σ_L . q^* is invariant with respect to μ_L (assuming λ and σ_L remain constant).

(d) Assume D_L is Normal with mean μ_L and standard deviation σ_L . q^* is increasing in σ_L .

(e) Assume the demand process is a Brownian motion with drift μ and volatility σ while orders incur a constant leadtime L . q^* is increasing in μ , σ and the leadtime L .

Proof: We write $r^*(\theta)$, $R^*(\theta)$ and $q^*(\theta)$ as r^* , R^* and q^* , without ambiguity.

For parts (a) and (b), we have $g(y|\theta) = (h+b)F(y) - b$ and $g_y(y|\theta) = (h+b)f(y)$.

(a) Choose $\theta = h$. Thus $g_\theta(y|\theta) = F(y) > 0$ on $[d_l, d_h]$. Note first that (4.21), the necessary and sufficient condition for q^* to be decreasing (increasing) in θ may be replaced by

$$\begin{aligned} & \int_{\max\{d_l, r^*\}}^{R^*} y g_y(y|\theta) dy \int_{\max\{d_l, r^*\}}^{R^*} g_\theta(y|\theta) dy \\ \leq & (>) \int_{\max\{d_l, r^*\}}^{R^*} g_y(y|\theta) dy \int_{\max\{d_l, r^*\}}^{R^*} y g_\theta(y|\theta) dy, \end{aligned} \quad (4.26)$$

since all integrands are zero on $[r^*, \max\{d_l, r^*\})$. If $F(y)$ is log-concave on $[d_l, d_h]$, i.e., $f(y)/F(y)$ is decreasing on $[d_l, d_h]$, it is log-concave on $[d_l, \infty)$ since $f(y)/F(y) =$

0 for all $y > d_h$. Thus $\frac{g_y(y|\theta)}{g_\theta(y|\theta)} = (h+b)\frac{f(y)}{\bar{F}(y)}$ is decreasing on $[d_l, \infty)$. Apply the result in Remark 1 with $[m, M] = [\max\{d_l, r^*\}, R^*]$ to verify that (4.26) holds with the \leq -sign.

The proof of the second statement in part (a) is analogous since $d_h = \infty$.

(b) Choose $\theta = b$. Thus $g_\theta(y|\theta) = -\bar{F}(y) < 0$ on $[d_l, d_h]$. Note first that (4.21), the necessary and sufficient condition for q^* to be decreasing (increasing) in θ may be replaced by

$$\begin{aligned} & \int_{r^*}^{\min\{d_h, R^*\}} y g_y(y|\theta) dy \int_{r^*}^{\min\{d_h, R^*\}} g_\theta(y|\theta) dy \\ \leq & (>) \int_{r^*}^{\min\{d_h, R^*\}} g_y(y|\theta) dy \int_{r^*}^{\min\{d_h, R^*\}} y g_\theta(y|\theta) dy, \end{aligned} \quad (4.27)$$

since all integrands are zero on $[\min\{d_h, R^*\}, R^*]$. If $\bar{F}(y)$ is log-concave on $[d_l, d_h]$, i.e., $-f(y)/\bar{F}(y)$ is decreasing on $[d_l, d_h]$, it is log-concave on $(-\infty, d_h]$ since $-f(y)/\bar{F}(y) = 0$ for all $y < d_l$. Thus $\frac{g_y(y|\theta)}{g_\theta(y|\theta)} = -(h+b)\frac{f(y)}{\bar{F}(y)}$ is decreasing on $(-\infty, d_h]$. Apply the result in Remark 1 with $[m, M] = [r^*, \min\{d_h, R^*\}]$ to verify that (4.27) holds with the \leq -sign.

The proof of the second statement in part (b) is analogous since $d_l \leq r^*(b)$.

For parts (c)-(d) where the leadtime demand is Normal, let $\phi(\cdot)$ and $\Phi(\cdot)$ denote the pdf and cdf of the standard Normal distribution and $z_y = (y - \mu_L)/\sigma_L$. In this case, $G(y|\theta) = h(y - \mu_L) + \sigma_L(h+b) \int_{z_y}^{\infty} (1 - \Phi(x)) dx$. Thus $g(y|\theta) = h - (h+b)(1 - \Phi(z_y))$, and $g_y(y|\theta) = \frac{h+b}{\sigma_L} \phi(z_y) > 0$.

(c) Let $\theta = \mu_L$. $g_\theta(y|\theta) = -g_y(y|\mu)$. Thus, the sufficient and necessary condition (4.21) holds as an equality, which implies that q^* is invariant with respect to μ_L .

(d) Let $\theta = \sigma_L$. $g_\theta(y|\theta) = -(h+b)(y - \mu_L)\phi(z_y)/(\sigma_L)^2$. $\frac{g_\theta(y|\theta)}{g_y(y|\theta)} = -\frac{y-\mu_L}{\sigma_L}$ is decreasing in y . By Remark 1 of Theorem 4.4, q^* is increasing in σ_L .

(e) Since D_L is Normal with $\mu_L = \mu L$ and $\sigma_L^2 = \sigma^2 L$, the monotonicity with respect to σ and L follow from parts (c)-(d). When μ increases, it causes increases of μ_L as well as the demand rate $\lambda = \mu$. The first increase has no impact on q^* while the increase of λ results in an increase in q^* , see (4.7). This implies that the combined effect amounts to an increase of q^* . (See the proofs of parts (f)-(g) of Corollary 4.1 for a precise development.) \square

Thus, monotonicity of q^* with respect to h and b hinges on whether the cdf of the leadtime demand $F(\cdot)$ or the complementary cdf $\bar{F}(\cdot)$ are log-concave or log-convex. For some distributions such as the Uniform and Power distributions, it is straightforward to verify that they have log-concave distribution and complementary distribution functions. However, many popular distributions such as the Normals lack a closed-form cdf. The following proposition proven by Bagnoli and Bergstrom (2005) provides sufficient conditions to determine log-concavity (log-convexity) of $F(\cdot)$ and $\bar{F}(\cdot)$ by verifying specific properties of the density function $f(\cdot)$. Several recent papers on inventory models have identified log-convexity or log-concavity of the demand distribution as the key condition guaranteeing various structural properties, see Huh et al. (2010), Levi et al. (2011) and Rosling (2002).

Proposition 4.2 (*Bagnoli and Bergstrom (2005)*) *Suppose the probability density function $f(y)$ has support (d, e) .*

(a) *If $f(\cdot)$ is continuously differentiable and log-concave on (d, e) , then $F(\cdot)$ and $\bar{F}(\cdot)$ are also log-concave on (d, e) .*

(b) *If $f(\cdot)$ is monotone increasing, then $F(\cdot)$ and $\bar{F}(\cdot)$ are log-concave.*

(c) *If $f(\cdot)$ is continuously differentiable and log-convex on (d, e) and $f(d) = 0$, then $F(\cdot)$ is also log-convex on (d, e) .*

(d) *If $f(\cdot)$ is continuously differentiable and log-convex on (d, e) and $f(e) = 0$, then $\bar{F}(\cdot)$ is also log-convex on (d, e) .*

Table 4.1 summarizes the monotonicity results with respect to the parameters h and b under some of the most common demand distributions.

4.4.2 The Discrete Model

In this subsection, we turn our attention to monotonicity properties of the optimal order quantity q^* in the discrete model. As in the case of the continuous model, we confine ourselves to monotonicity properties with respect to a scalar parameter θ , i.e., $\Theta \subseteq \mathcal{R}$. In Section 4.3, we showed that under the broad supermodularity con-

Table 4.1: Monotonicity results of q^* under some common demand distributions

Name of Distribution	$F(y)$	$\bar{F}(y)$	$q^*(h)$	$q^*(b)$
Uniform, Normal, Exponential, Gamma($k > 1$), Beta($\alpha > 1, \beta > 1$)	log-concave	log-concave	decreasing	decreasing
Weibull($k < 1$), Gamma($k < 1$), Pareto	log-concave	log-convex	decreasing	increasing
Power, Lognormal	log-concave	neither	decreasing	unkown

dition (SP) for the function $G(y|\theta)$, both r^* and R^* are decreasing in θ^3 , i.e., $r^*(\theta)$ and $R^*(\theta)$ are decreasing step functions. In other words, there exists a sequence of break points $\{\theta_1, \theta_2, \dots\}$ such that all three optimal policy parameters r^* , R^* and $q^* = R^* - r^*$ are constant in between break points; at break points, r^* decreases or R^* decreases, where it is easily shown that, in view of the strict quasi-convexity of $G(y|\theta)$, the decrease is by one unit exactly. Thus, at break points, q^* decreases or increases by 1 unit, or it stays constant. Even in the simplest applications, q^* fails to be *uniformly* monotone. However, in this section, we show that, under the same structural conditions assumed in Theorem 4.4, q^* is either *roughly* decreasing or *roughly* increasing. Here we define an integer valued function to be roughly decreasing (increasing) if the step function does not exhibit any pair of consecutive increases (decreases). In other words, $q^*(\theta)$ is an increasing (decreasing) step function except possibly for a few isolated unit step increases (decreases). The fact that $q^*(\theta)$ is roughly decreasing [increasing] implies that $q^*(\theta') \leq q^*(\theta) + 1$ [$q^*(\theta') \geq q^*(\theta) - 1$] for all $\theta < \theta'$. The following simple example shows that “rough monotonicity” is the best one can hope for in general. While the monotonicity properties in the discrete model and the conditions guaranteeing these properties are very similar to those in the continuous model, the required analysis is fundamentally different.

Example 1: Consider the standard inventory model with Poisson demands and a constant leadtime. The instantaneous expected cost function is given by (4.8)

³Recall that when $G(y|\theta)$ is submodular, both r^* and R^* are increasing in θ .

with $h(y) = hy$ and $b(y) = by$. We have shown in Section 4.3 that r^* and R^* are increasing in b . The following graph exhibits q^* as a function of b , where the Poisson rate $\lambda = 20, K = 64, L = 1, h = 1$. Note that q^* is roughly decreasing in b under Poisson demands, a property we will prove in Corollary 4.3(b).

Figure 4.1: q^* is roughly decreasing with respect to backlogging cost rate b

In the discrete model, we define $g(y|\theta) \equiv G(y+1|\theta) - G(y|\theta)$. Similarly, $g_y(y|\theta) \equiv g(y+1|\theta) - g(y|\theta)$ and $g_\theta(y|\theta) \equiv \frac{\partial g(y|\theta)}{\partial \theta}$. The next theorem, proven in the Appendix, establishes sufficient conditions for $q(\theta)$ to be roughly increasing or roughly decreasing.

Theorem 4.5 (Monotonicity of q^* in the discrete model) *Assume (Q) and (SP+) holds on a rectangle $[m, M] \times [\theta', \theta'']$ with $\theta' < \theta''$ and $m \leq r^*(\theta) < R^*(\theta) \leq M$ for all $\theta \in [\theta', \theta'']$. $q^*(\cdot)$ is decreasing (increasing) on $[\theta', \theta'']$ if $g_y(y|\theta_1)/g_\theta(y|\theta_0)$ is decreasing (increasing) in $y \in [m, M]$ for any $\theta_0, \theta_1 \in [\theta', \theta'']$ with $\theta_0 \geq \theta_1$ and is not constant.*

We now apply Theorem 4.5 to derive the rough monotonicity property of q^* with respect to several parameters in the discrete standard inventory model. Similar to the continuous model, we obtain these properties for leadtime demand distributions that are log-concave or log-convex. See Corollary 4.3 below with its proof in the Appendix. Here, a non-negative discrete function $p(\cdot) : \mathcal{Z} \rightarrow \mathcal{R}^+$ is defined to be strictly log-concave (log-convex) on some interval $[a, b]$, with $0 < a < b \leq \infty$ if

$\log p(n)$ is strictly concave (convex), i.e., for all $n \in [a, b]$

$$\log p(n+1) - \log p(n) < (>) \log p(n) - \log p(n-1) \iff \frac{p(n+1)}{p(n)} < (>) \frac{p(n)}{p(n-1)}. \quad (4.28)$$

Corollary 4.3 (Monotonicity properties of q^* in the discrete standard inventory model) *Assume the instantaneous expected cost function $G(\cdot|\theta)$ is of the form (4.8) with constant marginal inventory and backlogging cost rates h and b . The continuous leadtime demand D_L has support $[d_l, d_h]$ where $-\infty \leq d_l < d_h \leq \infty$.*

(a) q^* is roughly decreasing in h if $F(y)$ is log-concave on $[d_l, d_h]$; q^* is roughly increasing in h if $F(y)$ is log-convex on $[d_l, d_h]$ with $d_h = \infty$.

(b) q^* is roughly decreasing in b if $\bar{F}(y)$ is log-concave on $[d_l, d_h]$; q^* is roughly increasing in b if $\bar{F}(y)$ is log-convex with $r^*(b) \geq d_l$.

Many common discrete probability density distributions are log-concave, including the discrete uniform distribution, the Poisson, the binomial, the hypergeometric, and the negative binomial. Lemma 4.4 below imply that their cdf and ccdf are log-concave as well. As a specific example, recall Example 1, displaying q^* as a roughly decreasing function of b when the leadtime demand distribution is Poisson. We first need the following properties for strictly log-concave (log-convex) functions, the proofs of which is in the Appendix.

Lemma 4.4 (Properties of discrete log-concave and log-convex functions) *Assume the function $p(\cdot)$ is log-concave (log-convex) on some interval $[a, b]$.*

(a) $P(n) = \sum_{i=a}^n p(i)$ is log-concave (log-convex) on $[a, b]$.

(b) If $b = \infty$ and $\lim_{n \rightarrow \infty} p(n) = 0$, then $P(n) = \sum_n^\infty p(i)$ is log-concave (log-convex) on $[a, \infty)$.

4.5 Monotonicity of the Optimal Cost Value

In this section, we briefly discuss monotonicity properties of the optimal cost value. Theorem 4.2 shows that a general sufficient condition for monotonicity of r^* and R^* with respect to the model primitives $\theta \in \Theta$ relates to the *second-order* properties

of the instantaneous expected cost function $G(\cdot)$ (, in particular supermodularity or submodularity which, for twice differentiable functions, is equivalent to the cross partial derivative with respect to y and θ being nonnegative or nonpositive). In a similar vein, a broad sufficient condition for monotonicity of the optimal cost value $c^*(\theta) \equiv \min_{(r,q)} c(r, q|\theta)$ is obtained from a *first-order* property of the instantaneous expected cost function $G(\cdot)$.

Proposition 4.3 (Monotonicity of $c^*(\theta)$) *$c^*(\theta)$ is increasing [decreasing] if the function $G(y|\theta)$ is increasing [decreasing] in $\theta \in \Theta$ for all y .*

Proof: We provide the proof for the continuous model. The proof for the discrete model is analogous. Let $\theta_1 \preceq \theta_2 \in \Theta$, $c^*(\theta_1) = \min_{(r,q)} \frac{\lambda K + \int_r^{r+q} G(y|\theta_1) dy}{q} \leq \min_{(r,q)} \frac{\lambda K + \int_r^{r+q} G(y|\theta_2) dy}{q} = c^*(\theta_2)$ where the inequality follows from the minimand to the left being dominated by the minimand to the right for every pair (r, q) . (The latter is immediate from $G(\cdot|\theta)$ being pair-wise increasing in θ .) \square

In the standard inventory model, Proposition 4.3 implies that the optimal cost value is increasing in the holding cost rate function $h(\cdot)$, backlogging cost rate function $b(\cdot)$ and the wholesale price w . Moreover, in the standard inventory model, $G(y|\theta) = \mathbb{E}h((y - D_L)^+) + \mathbb{E}b((D_L - y)^+)$. Since $h(\cdot)$ and $b(\cdot)$ are convex functions, it follows that, when comparing two leadtime demand distributions D_L^1 and D_L^2 that are convexly ordered, i.e., $D_L^1 \leq_{cx} D_L^2$, the corresponding instantaneous expected cost functions $G^1(\cdot)$ and $G^2(\cdot)$ are pointwise ordered as well, i.e., $G^1(y) \leq G^2(y)$ for all y . (For any pair of random variables X and Y , $X \leq_{cx} Y$ if $\mathbb{E}f(x) \leq \mathbb{E}f(Y)$ for any convex function $f(\cdot)$.) By Proposition 4.3, the optimal cost value under D_L^2 is larger than that under D_L^1 . The convex ordering $D_L^1 \leq_{cx} D_L^2$ implies that $\mathbb{E}D_L^1 = \mathbb{E}D_L^2$ and $var(D_L^1) \leq var(D_L^2)$. The monotonicity of $c^*(\theta)$ under convexly ordered leadtime demand distributions was proved by Song (1994), for the case of compound Poisson demands. Song (1994) also introduced a stronger, so-called variability ordering \leq_{var} , which is easier to verify, but still includes many commonly used classes of distributions, including gamma, uniform, Normal and truncated Normal distributions. See also Lemmas 5 and 6 and Theorem 5 in Song et. al (2010). Finally, Song (1994) showed that, under a Poisson demand process, for two lead-

time demand distributions to be convexly ordered it is sufficient that the leadtime distributions L^1 and L^2 are ordered according to this order, i.e., $L^1 \leq_{cx} L^2$.

One might conjecture that a (stochastically) longer leadtime also results in a larger optimal cost value. However, if $L_1 \leq_{st} L_2$ and demands are generated by a Poisson process, it is easily seen that the corresponding instantaneous expected cost rate functions $G^1(\cdot)$ and $G^2(\cdot)$ may fail to be pointwise ordered, i.e., the sufficient condition of Proposition 4.3 fails to apply. Indeed, Song (1994) and Song et al. (2010) exhibit examples where a stochastically larger leadtime results in a lower optimal cost value. In general, to our knowledge, all established monotonicity results pertaining to the optimal cost value $c^*(\cdot)$ reflect settings where there is pointwise dominance of the instantaneous expected cost functions.

4.6 Conclusions and Future Work

In this paper, we have addressed stochastic inventory systems which are governed by (r, q) or (r, nq) policies. For this class of models, we have provided general sufficient conditions under which each of the three optimal policy parameters (r^*, R^*, q^*) as well as the optimal cost value vary monotonically with various model primitives. The monotonicity properties for c^* , r^* and R^* are common to the continuous and discrete models and allow for common analyses. As far as the order quantity q^* is concerned, the continuous model, similarly, allows for broadly satisfied conditions under which q^* is purely monotone. However, in the discrete model, only a somewhat weaker form of monotonicity can be hoped for, where isolated unit step increases (decreases) may interrupt an otherwise purely decreasing (increasing) pattern. (We have referred to this as “rough monotonicity”.) After establishing sufficient conditions for the general model, we have derived various applications for the standard inventory model where the instantaneous expected cost function is of the form (4.8). The implications in Corollaries 4.1-4.3 are meant to be illustrative only.

Future work should attempt to investigate the monotonicity properties in other commonly used models, for example, those governed by an (s, S) policy (e.g., Por-

teus (2002) and Zipkin (2000)) or an (R, T) policy (e.g., Rao (2003)). Under an (s, S) policy, the inventory position is increased to a constant level S whenever it falls below s . Under an (R, T) policy, the inventory position is reviewed every T periods and increased to a constant level R . As far as the former class of policies is concerned, preliminary results indicate that the same full monotonicity properties for the optimal policy parameters may not (always) prevail, see the following example.

Example 2: Consider a periodic review inventory system. Demands in each period are independent and Poisson distributed with mean $\lambda = 25$. $G(y|D) = h\mathbb{E}(y - D)^+ + b\mathbb{E}(D - y)^+$ with $h = 1$ and $b = 9$ while $K = 64$. Consider two random leadtimes L^1 and L^2 , where $P(L^1 = 1) = 0.7$ and $P(L^1 = 2) = 0.3$ while $P(L^2 = 1) = 0.1$ and $P(L^2 = 2) = 0.9$. We have $D_L^1 \leq_{st} D_L^2$. However, the optimal (s, S) policy we found using the algorithm in Zheng and Federgruen (1992) has $(s_1, S_1) = (28, 86)$ and $(s_2, S_2) = (43, 82)$. In other words, contrary to Corollary (4.1)(d), the order-up-to level is *decreased* as we move to a stochastically larger leadtime distribution.

Appendix A

Appendices for Chapter 2

A.1 Proofs for Sections 1-7 of Chapter 2

Proof of Theorem 2.1.

Part (a): Rewrite (2.4)

$$w(y) = \frac{p - (p + h_0)F(y)}{1 - (1 - \delta)F(y)} = \frac{1}{1 - \delta} \left(p + h_0 - \frac{p\delta + h_0}{1 - (1 - \delta)F(y)} \right), \quad (\text{A.1})$$

so that

$$w'(y) = -\frac{(p\delta + h_0)f(y)}{(1 - (1 - \delta)F(y))^2} \leq 0. \quad (\text{A.2})$$

Differentiating (2.5) yields

$$\begin{aligned} \Pi'_s(y|\beta_g) &= w(y)\xi'(y) + \xi(y)w'(y) - c(1 - F(y)) \\ &= w(y)\xi'(y)g(y) \end{aligned} \quad (\text{A.3})$$

where

$$g(y) = 1 - c \frac{1 - F(y)}{w(y)\xi'(y)} - \frac{-\xi(y)w'(y)}{\xi'(y)w(y)} \quad (\text{A.4})$$

for $y < y_{max}$. We prove that $g(y)$ is decreasing on $[0, y_{max})$ by showing that each term on the right of (A.4) is decreasing, i.e.,

$$\frac{1 - F(y)}{w(y)\xi'(y)} \text{ is increasing,} \quad (\text{A.5})$$

and

$$\frac{-\xi(y)w'(y)}{w(y)\xi'(y)} \text{ is increasing.} \quad (\text{A.6})$$

$\frac{1-F(y)}{\xi'(y)} = \frac{1-F(y)}{1-(1+\beta_g)F(y)} = 1 + \frac{\beta_g F(y)}{1-(1+\beta_g)F(y)}$ is positive and increasing on $[0, y_{max})$ since $\xi'(y) > 0$ by (2.6). Thus (A.5) follows from the fact that $w(y)$ is positive and decreasing.

To show (A.6), note that $\xi(y)$ is a positive and increasing function on $[0, y_{max})$ since $\xi'(y) > 0$ by (2.6) and $\xi(0) = 0$. Again by (2.6), $\xi'(y)$ is a decreasing function. Hence $\frac{\xi(y)}{\xi'(y)}$ is positive and increasing. Thus, we only need to show that $\frac{-w'(y)}{w(y)}$ is positive and increasing. It is positive because $-w'(y) \geq 0$, see (A.2). Combining (2.4) and (A.2) we obtain that:

$$\begin{aligned} \frac{-w'(y)}{w(y)} &= \frac{(p\delta + h_0)f(y)}{(1 - (1 - \delta)F(y))(p - (p + h_0)F(y))}, \\ &= \frac{(\delta + h_0/p)f(y)}{(1 - (1 - \delta)F(y))(1 - (1 + h_0/p)F(y))}. \end{aligned}$$

Since $\delta = \bar{\alpha} - \beta_g \in [0, 1]$, $1 - (1 - \delta)F(y)$ is positive and decreasing. Therefore, it suffices to show that $\frac{f(y)}{1-(1+h_0/p)F(y)} = \frac{f(y)}{1-F(y)} \left(1 + \frac{F(y)h_0/p}{1-(1+h_0/p)F(y)}\right)$ is positive and increasing. It is positive since $1 - (1 + h_0/p)F(y) \geq 0$ on $[0, y_{max})$ as $F(y) < F(y_{max}) = \frac{1}{1+\max\{h_0/p, \beta_g\}}$. Finally, it is increasing as the product of two positive and increasing functions, using the fact that the demand distribution is IFR.

We next show that $g(0) > 0$ and $\lim_{y \uparrow y_{max}} g(y) < 0$. Since $g(\cdot)$ is decreasing on $[0, y_{max})$, this implies $g(\cdot)$ has a unique root $y_{\beta_g}^*$. By (A.3) and $\xi'(y) > 0$, it follows that $\Pi_s(\cdot|\beta_g)$ is increasing for $y \leq y_{\beta_g}^*$ and decreasing for $y > y_{\beta_g}^*$, thus establishing the theorem.

To verify the sign of the $g(\cdot)$ function at the boundaries of $[0, y_{max})$, note that $g(0) = 1 - c/p > 0$ since $F(0) = 0$, $w(0) = p$, $\xi(0) = 0$ and $\xi'(0) = 1$. If $y_{max} = y_s$, by (2.6) and $\xi'(y_{max}) = 0$, $\Pi'_s(y_{max}) = \xi(y_{max})w'(y_{max}) - c(1 - F(y_{max})) < 0$ since $w'(\cdot) \leq 0$ and $\xi(\cdot) \geq 0$. This implies that $\lim_{y \uparrow y_{max}} g(y) < 0$. Finally if $y_{max} = F^{-1}\left(\frac{1}{1+h_0/p}\right)$, then the profit value $\Pi_s(y_{max}|\beta_g)$ itself is negative by (2.5) as $w(y_{max}) = 0$. This, however, implies that $\Pi_s(y|\beta_g)$ is decreasing at some point $y^0 < y_{max}$, so that $g(y^0) < 0$. By the monotonicity of $g(\cdot)$, $\lim_{y \uparrow y_{max}} g(y) < 0$ in this case as well.

Part (b): $\lim_{w \uparrow p} \hat{\Pi}_s(w|\beta_g) = 0$ since $\lim_{w \uparrow p} y(w, \beta_g) = 0$. The remainder of part (b) is immediate from part (a). \square

Proof of Lemma 2.1.

Part (a): It is sufficient to show that $\frac{\int_0^y F(u)du}{F^2(y)}$ is an increasing function, i.e., its derivative $\frac{F^3(y)-2F(y)f(y)\int_0^y F(u)du}{F^4(y)} \geq 0 \iff F^2(y) - 2f(y)\int_0^y F(u)du \geq 0$ for all $y \geq L = 0$. However, the last inequality holds for $y = 0$ and the derivative of the function on the left side is $-2f'(y)\int_0^y F(u)du \geq 0$ for all $y \geq 0$.

Part (b): *Normal*: $\zeta(y) = b(x) \equiv \frac{x\Phi(x)+\phi(x)}{\Phi^2(x)} \frac{\phi(x)}{1-\Phi(x)}$ where $x = \frac{y-\mu}{\sigma}$. It follows from the proof in part (a) that the first factor is increasing if and only if $\Phi^2(x) - 2\phi(x)\int_{-\infty}^x \Phi(u)du = \Phi^2(x) - 2\phi(x)(x\Phi(x) + \phi(x)) \geq 0$. Since the left hand side of the inequality is increasing when $\phi(x)$ is decreasing, i.e., $x \geq 0$, and it already equals 0.026 at $x = 0.6$, it is increasing for $x \geq 0.6$, and so is the second factor since all Normal distributions have the IFR property. We conclude that $b(x)$ is increasing for all $x > 0.6$.

To show $b(x)$ is increasing on $[-1.8, 0.6]$, it suffices to prove that $\log b(x) = \log(x\Phi(x) + \phi(x)) - 2\log \Phi(x) + \log \phi(x) - \log(1 - \Phi(x))$ is increasing on this interval, i.e., $(\log b(x))' \geq 0$ where $(\log b(x))' = \frac{\Phi(x)}{x\Phi(x)+\phi(x)} - \frac{2\phi(x)}{\Phi(x)} - x + \frac{\phi(x)}{1-\Phi(x)}$. Let $m \equiv \min_{x \in [-1.8, 0.6]} \log b(x)$. Thus the proof of MIFR properties for all Normal distributions is reduced to verify that the single number m is positive. The single number can be evaluated by plotting the minimand function on $[-1.8, 0.6]$ or by applying a global minimization problem to the single variable function. Below we show the verification of $m \geq 0$ can be achieved by evaluating a related single variable function at a *finite* number of points. We distinguish between the intervals $[-1.8, 0]$ and $[0, 0.6]$, the first of which is divided into 900 subintervals and the latter into 13 subintervals. We prove the inequality $(\log b(x))'$ on each of the subintervals.

(I) $-1.8 \leq x \leq 0$: Divide the interval into subintervals of length 0.0002. For any subinterval $[x_1, x_2 = x_1 + 0.0002]$, we have for all $x_1 \leq x \leq x_2$, $(\log b(x))' \geq \frac{\Phi(x_1)}{x_2\Phi(x_2)+\phi(x_2)} - \frac{2\phi(x_2)}{\Phi(x_1)} - x_2 + \frac{\phi(x_1)}{1-\Phi(x_1)} > 0$, where the last inequality can be verified for all 900 starting points $x_1 = -1.8 + 0.0002k$, $k = 0, \dots, 899$.

(II) $0 < x \leq 0.6$: Divide the interval into subintervals of length 0.05. For any such subinterval $[x_1, x_2 = x_1 + 0.05]$, we have for all $x_1 \leq x \leq x_2$, $(\log b(x))' \geq \frac{\Phi(x_1)}{x_2\Phi(x_2)+\phi(x_2)} - \frac{2\phi(x_1)}{\Phi(x_1)} - x_2 + \frac{\phi(x_2)}{1-\Phi(x_1)} > 0$, where the last inequality can be verified

for all 13 starting points $x_1 = 0.05k$, $k = 0, \dots, 12$. \square

Proof of Theorem 2.2.

Parts (b) and (c) are immediate from part (a).

Part (a): Using (2.12), write

$$\Pi'_s(y|w) = (1 - F(y))n(y) - (w\bar{a} + h_0), \quad (\text{A.7})$$

where $n(y) = (p - c + w\bar{a} + h_0) - (p - w)\zeta(y)$. Note that $n(y)$ is decreasing in view of the (MIFR) property. Thus define $y_w \equiv \inf\{L \leq y : n(y) \leq 0\}$, unless $n(\cdot) > 0$ for all $y \geq L$, in which case we define $y_w \equiv \infty$. Note that, in either case, $n(\cdot) \geq 0$ for all $L \leq y < y_w$. Thus, for all $L \leq y < y_w$, $\Pi'_s(\cdot|w)$ is a decreasing function of y as a translation of the product of two positive and decreasing functions. In other words, $\Pi_s(\cdot|w)$ is concave on $[L, y_w]$. Moreover, $n(y) \leq 0$ and $\Pi'_s(\cdot|w) < 0$ for $y > y_w$, i.e., $\Pi_s(\cdot|w)$ is decreasing for $y > y_w$. This implies that the function $\Pi_s(\cdot|w)$ is quasi-concave in $[L, \infty)$, achieving its maximum for some $L \leq y^0 < y_w$. \square

Proof of Theorem 2.3.

Part (a): Note that $y_0(w) = \operatorname{argmax}_{g(w,y) \geq 0} \Pi_s(y|w)$ where $g(w, y) = y - L$. The function $\Pi_s(\cdot|w)$ is twice continuously differentiable, see (2.12), given that the demand distribution has a continuously differentiable pdf. Similarly, the constraint function $g(w, y)$ is a twice differentiable function as well. Since the constraint is linear in the decision variable y , the Karush-Kuhn-Tucker conditions necessarily hold at the optimal solution $y_0(w)$ for all $w \in [0, p]$. Continuity of $y_0(\cdot)$ follows from Theorem 4.1 of Fiacco and Kyparisis (1985), by verifying that the so-called SSOSC condition *ibid* is satisfied. When applied to a mathematical program with a single variable and linear constraints, the SSOSC condition is reduced to verifying that the second derivative of the objective function is strictly negative in the optimal solution unless one of the constraints is binding and the constraint gradient is nonzero. In our case, when $L < y_0(w)$, $\Pi''_s(y_0(w)|w) < 0$ follows from Theorem 2.2.

Part (b): Consider first the case where $L < y_0(w)$. It follows from the differentiability of $\Pi'_s(\cdot|w)$, see (2.13), and its quasi-concavity, see Theorem 2.2, that

$$\Pi'_s(y_0(w)|w) = 0 \text{ for all } w \in [0, p] \text{ with } y_0(w) > L. \quad (\text{A.8})$$

Since $\frac{\partial^2 \Pi_s(y_0|w)}{\partial y^2}$ and $\frac{\partial^2 \Pi_s(y_0|w)}{\partial y \partial w}$ both exist and are continuous functions, it follows from the Implicit Function Theorem that $y_0(\cdot)$ is a continuously differentiable function with $y'_0(w) = \frac{\partial^2 \Pi_s(y_0|w)}{\partial y \partial w} / \frac{\partial^2 \Pi_s(y_0|w)}{\partial y^2}$, where $\frac{\partial^2 \Pi_s(y_0|w)}{\partial y^2} < 0$ by Theorem 2.2 (a). Thus, by (2.13), $y'_0(w) \geq 0 \iff \frac{\partial^2 \Pi_s(y_0|w)}{\partial y \partial w} = -F(y_0(w))(\bar{\alpha} - \kappa(y_0(w))) \geq 0 \iff \kappa(y_0(w)) \geq \bar{\alpha}$. On the other hand, since $\Pi'_s(y_0(w)|w) = 0$, we obtain from (2.13) that for any wholesale price $w < p$ with $y_0(w) > L$, $\kappa(y_0(w)) \geq \bar{\alpha} \iff (p - c) - (p - c + p\bar{\alpha} + h_0)F(y_0(w)) \geq 0 \iff F(y_0(w)) \leq \frac{p-c}{p-c+p\bar{\alpha}+h_0} \iff y_0(w) \leq F^{-1}\left(\frac{p-c}{p-c+p\bar{\alpha}+h_0}\right) \leq \max\{L, F^{-1}\left(\frac{p-c}{p-c+p\bar{\alpha}+h_0}\right)\} = y^p$. Thus, for any $w < p$ and $y_0(w) > L$,

$$y_0(w) \text{ is increasing in } w \iff \frac{\partial^2 \Pi_s(y_0|w)}{\partial y \partial w} \geq 0 \iff \kappa(y_0(w)) \geq \bar{\alpha} \iff y_0(w) \leq y^p. \quad (\text{A.9})$$

The equivalent conditions for $y_0(\cdot)$ being *strictly* increasing are obtained by making all inequalities above *strict*.

In view of the continuity of $y_0(\cdot)$ on the complete interval $[0, p]$, only the following four situations can occur.

(i) $y_0(w) = L$ for all $w \in [0, p]$; (ii) $y_0(w) = y^p > L$ for all $w \in [0, p]$; (iii) there exists a value w^0 such that $L < y_0(w^0) < y^p$; (iv) there exists a value w^0 such that $y^p < y_0(w^0)$.

Case (i) can only occur when $y^p = L$ since $y_0(p) = y^p$, where it corresponds with a special case of (b-iv). In case (ii), $L < y^p = y_0(w)$ for all $w \in [0, p]$, which corresponds with part (b-ii). Fix a wholesale price w , with $0 \leq w < p$. Since $L < y_0(w)$, it follows from (A.8) that $\Pi'_s(y_0(w)|w) = \Pi'_s(y^p|w) = 0$, and by (2.13), $\kappa(y^p) = \bar{\alpha}$.

Case (iii): It follows from (A.9) that, when *decreasing* w downwards from w^0 , the $\{y_0(\cdot)\}$ -curve continues to decrease, possibly on the entire interval $[0, w^0]$ or until the level L is reached for some $0 \leq w_L \leq w^0$. In the latter case, $y_0(w) = L$ for all $0 \leq w \leq w_L$, for, otherwise, $0 < w^1 \equiv \sup\{0 < w < w_L : y_0(w) > L\}$. However, the latter case implies that as $w \uparrow w_1$, $L < y_0(w) < y^p$ and $y'_0(w) < 0$, which is contradicted by (A.9). Similarly, it follows from (A.9) that, when *increasing* w from w^0 , the $y_0(\cdot)$ -curve continues to increase until hitting the level y^p , either for $w = p$

or for some $w^0 < w_U < p$. It remains to be shown that the latter case cannot occur: after all, since $y_0(\cdot)$ is differentiable at w_U as $L < y_0(w_U) < \bar{y}(0)$, we have $\lim_{w \uparrow w_U} y'_0(w) = \lim_{w \uparrow w_U} y'_0(w) > 0$; In other words, the $\{y_0(\cdot)\}$ -curve crosses the level $y = y^p$ at $w = w_U$ from below. By the continuous differentiability of $y_0(\cdot)$ on the interior of $[L, \bar{y}(0)]$, there exists an interval $[w', w'']$ with $w' < w_U < w''$ such that (i) $L < y_0(w) < \bar{y}(0)$ for all $w \in [w', w'']$, (ii) $y'_0(w) > 0$ for all $w \in [w', w'']$ and (iii) $y_0(w) > y^p$ for $w_U < w < w''$. But (A.9) shows that $y_0(\cdot)$ is *decreasing* when $y_0(w) > y^p$, contradicting the assumption that the function $y_0(\cdot)$ crosses the level y^c from below. Therefore $y_0(\cdot)$ is increasing and approaches y^p at $w = p$.

We have shown that in case (iii) the function behaves as in part (b-i), since for all $w \uparrow p$, $L < y_0(w) < y^p$ while $y_0(\cdot)$ is strictly increasing; By (A.9), this implies that $\kappa(y^p) > \bar{\alpha}$. Therefore $\lim_{w \uparrow p} \kappa(y_0(w)) = \kappa(y_0(p)) \geq \bar{\alpha}$ by the continuity of the functions $\kappa(\cdot)$ and $y_0(\cdot)$, see part (a). Finally, the possibility of $\kappa(y^p) = \bar{\alpha}$ can be excluded since it implies $\Pi'_s(y^p|w) = 0$ for all $c \leq w \leq p$, i.e., $y_0(w) = y^p$ for all $0 \leq w \leq p$, contradicting the assumption of case (iii).

Case (iv): Since $L \leq y^p < y_0(w_0)$, it follows from (A.9) that as w is decreased downward from w^0 , the $y_0(\cdot)$ -curve increases on the entire interval $[0, w^0]$ since $y_0(w) > y^p$ for all $w \in [0, w^0]$. Similarly it follows from (A.9) that when w is *increased* from w^0 , the $y_0(\cdot)$ curve decreases either on the entire interval $[w^0, p]$ or until hitting the level $y = y^p$ at some point $w^0 < w^2 < p$. However, the latter case cannot occur for $y^p > L$ as can be shown with arguments analogous to those used in case (iii). Therefore, the function $y_0(\cdot)$ behaves as in part (b-iii) while the proof of $\kappa(y_0) < \bar{\alpha}$ is analogous to that of $\kappa(y_0) > \bar{\alpha}$ in case (iii). When $y^p = L$, it is possible for the $y_0(\cdot)$ -curve to hit the line $y = y^p = L$ at some wholesale price $w^2 < p$, but with arguments analogous to those employed in case (iii) when the $y_0(\cdot)$ -curve hits the level $y = L$, we can show that, in this case, $y_0 = y^p$ for all $w \in [w^2, p]$, i.e., the function $y_0(\cdot)$ behaves as in part (b-iv). \square

Proof of Proposition 2.1.

We first show that under the three families of distributions,

$$\kappa(\cdot) \text{ is strictly decreasing for } y \geq L, \kappa(M) > 1 - c/p \text{ and } \kappa(M) > \bar{\alpha} \quad (\text{A.10})$$

where M denotes the median of the distribution.

Uniform on an interval $[a, b]$: $F(y) = \frac{y-a}{b-a}$, $f(y) = \frac{1}{b-a}$, $h(y) = \frac{(y-a)^2}{2(b-a)}$. Thus $\kappa(y) = \frac{b-a}{2(y-a)}$ is strictly decreasing for all $y \in (a, b]$. Moreover, $M = (b+a)/2$ and $\kappa(M) = 1 > 1 - c/p$ as well as $\kappa(M) = 1 > \bar{\alpha}$.

Exponential with rate λ : For $y > 0$, $F(y) = 1 - e^{-\lambda y}$, $f(y) = \lambda e^{-\lambda y}$, $h(y) = y - \frac{1-e^{-\lambda y}}{\lambda}$. $\kappa(y) = \frac{(\lambda y - 1)e^{-\lambda y} + e^{-2\lambda y}}{(1-e^{-\lambda y})^3}$. We now show that $\kappa(y)$ is strictly decreasing in y . This is equivalent to showing that $c(x) \equiv \frac{(\ln x - 1)x^2 + x}{(x-1)^3}$ is strictly decreasing in $x \equiv e^{\lambda y} \geq 1$, i.e., $c'(x) = \frac{\bar{c}(x)}{(x-1)^4} < 0$ where $\bar{c}(x) = x^2(2 - \ln x) - 2x \ln x - x - 1$. To show that $\bar{c}(x) \leq 0$ for all $x \geq 1$, note that since $\bar{c}(1) = 0$, it is sufficient to show that $\bar{c}'(x) = 3(x-1) - 2(x+1) \ln x < 0$ for all $x > 1$. Similarly, to verify the latter, it suffices to show $\bar{c}''(x) = 1 - 2/x - 2 \ln x < 0$ for all $x > 1$ as $\bar{c}'(1) = 0$. Since $\bar{c}''(1) = -1$, this follows from $\bar{c}'''(x) = \frac{2(1-x)}{x^2} < 0$ for all $x > 1$.

Since $M = \ln 2/\lambda$, $\kappa(M) = 4(\ln 2 - 0.5) = 0.773 > 1 - c/p$ since we assume $p \leq 4c$. Finally, $\kappa(M) = 0.773 > \bar{\alpha}$ by assumption.

Normal with mean μ and standard deviation σ : We show that $\kappa(y)$ is strictly decreasing for $y \geq \mu - 1.8\sigma$ where $\kappa(y) = a(x) \equiv \frac{\phi(x)(x\Phi(x) + \phi(x))}{\Phi^3(x)}$ and $x = \frac{y-\mu}{\sigma}$, it suffices to show that $\log a(x) = \log \phi(x) + \log(x\Phi(x) + \phi(x)) - 3 \log \Phi(x)$ is decreasing on the interval $[-1.8, \infty)$. However,

$$\begin{aligned} (\log a(x))' &= -x + \frac{\Phi(x)}{x\Phi(x) + \phi(x)} - 3 \frac{\phi(x)}{\Phi(x)} \\ &= \frac{-x\Phi(x)(x\Phi(x) + \phi(x)) + \Phi^2(x) - 3\phi(x)(x\Phi(x) + \phi(x))}{\Phi(x)(x\Phi(x) + \phi(x))} \\ &= \frac{-(x^2 - 1)\Phi^2(x) - 4x\phi(x)\Phi(x) - 3\phi^2(x)}{\Phi(x)(x\Phi(x) + \phi(x))}. \end{aligned} \quad (\text{A.11})$$

It is easy to verify that $(\log a(x))' < 0$, where $x > 1$, since all three terms in the numerator of (A.11) are negative components when $x \geq 1$. Thus we only need to show that $(\log a(x))' < 0$ on $[-1.8, 1]$. Divide the interval into subintervals of length 0.1. In any subinterval $[x_1, x_2 = x_1 + 0.1]$ with $x_2 \leq 0$, we have for any $x_1 \leq x \leq x_2$, $(\log b(x))' \leq -x_1 + \frac{\Phi(x_2)}{x_1\Phi(x_1) + \phi(x_1)} - 3 \frac{\phi(x_1)}{\Phi(x_2)} < 0$, which can be verified for all 18 starting points $x_1 = -1.8 + 0.1k$, $k = 0, \dots, 17$. Similarly, $(\log b(x))' \leq -x_1 + \frac{\Phi(x_2)}{x_1\Phi(x_1) + \phi(x_1)} - 3 \frac{\phi(x_2)}{\Phi(x_2)} < 0$ for any $x_1 \leq x \leq x_2$ and $x_1 = 0.1k$, $k = 0, \dots, 10$.

The mean μ of a normal distribution is also its median and $\kappa(M) = \kappa(\mu) = a(0) = 1.27 > \max\{1 - c/p, \bar{\alpha}\}$.

So far we have completed the proof of (A.10) for the three families of distributions considered in the Proposition. While $\kappa(\cdot)$ is decreasing, recall from the MIFR property that $\zeta(y_0) = \kappa(y_0) \frac{F(y_0)}{1-F(y_0)}$ is increasing. These two functions therefore have a unique intersection at the point where $F(y) = 1/2$, i.e., at the median M .

Assume to the contrary that y_0 is decreasing in w . Then by (A.9) and also from the proof of Theorem 2.3:

$$\kappa(y_0) \leq \bar{\alpha}. \quad (\text{A.12})$$

From (2.12), rewrite $0 = \Pi'_s(y|w)$ as follows:

$$\Pi'_s(y_0|w) = (p - c)(1 - F(y_0)) - pF(y_0)\kappa(y_0) - h_0F(y_0) - wF(y_0)(\bar{\alpha} - \kappa(y)) \quad (\text{A.13})$$

Thus $(p - c)(1 - F(y_0)) - pF(y_0)\kappa(y_0) \geq 0$, i.e.,

$$\zeta(y_0) = \kappa(y_0) \frac{F(y_0)}{1 - F(y_0)} \leq 1 - c/p < \kappa(M) = \zeta(M). \quad (\text{A.14})$$

Since $\zeta(\cdot)$ is increasing, $y_0 < M$. Since $\kappa(\cdot)$ is strictly decreasing, $\kappa(y_0) > \kappa(M) > \bar{\alpha}$, contradicting (A.12) and hence the assumption that $y_0(w)$ is a decreasing. \square

Proof of Theorem 2.4

(a1) Note that $\Pi_s^*(w) = \max\{\Pi_s(y|w) : g_i(w, y) \geq 0\}$ where $g_1(w, y) = y - \underline{y}(w)$ and $g_2(w, y) = \bar{y}(w) - y$. The function $\Pi_s(y|w)$ is twice continuously differentiable, see (2.12), given that the demand distribution has a continuously differentiable pdf. Similarly, the constraint functions $g_1(w, y)$ and $g_2(w, y)$ are both twice differentiable functions as well. Since both constraints are linear in the decision variable y , the Karush-Kuhn-Tucker conditions necessarily hold at the optimal solution y_w^* for all $w \in [0, \bar{w}]$. The differentiability of the optimal value function $\Pi_s^*(\cdot)$ follows from Theorem 7.3 of Fiacco and Kyparisis (1985), by verifying that the so called (SSOSC) condition *ibid* is satisfied. (The only other condition for the theorem is that in the optimal solution of the mathematical program, the vectors of gradients of the binding constraints are linearly independent; in our case, at most one constraint is binding and its gradient is ± 1 .) When applied to a mathematical program with a single

variable and linear constraints, the SSOSC condition is reduced to verifying that the second derivative of the objective function is strictly negative in the optimal solution, unless one of the constraints is binding and the constraint gradient is nonzero. In our case, when $\underline{y}(w) < y_0(w) < \bar{y}(w)$, $y_w^* = y_0(w)$ and $\Pi_s''(y_0(w)|w) < 0$ follows from Theorem 2.2.

(a2) We distinguish between two cases:

Case 1: $\bar{w} = p$

It follows from Theorem 2.2 that the wholesale price range $[0, p]$ may be partitioned into a sequence of consecutive (, possibly single point) intervals that are of one of the following three types:

Type-1 interval: $(\beta^*(w), y_w^*) = (\underline{\alpha}, \underline{y}(w))$ on the whole interval.

Type-2 interval: $(\beta^*(w), y_w^*) = (\beta_g(w, y_0(w)), y_0(w))$ on the whole interval.

Type-3 interval: $(\beta^*(w), y_w^*) = (\bar{\alpha}, \bar{y}(w))$ on the whole interval.

When $w = 0$ the supplier's loss is increasing in y , see (2.2), which prompts him to set $\beta_g^*(0) = \underline{\alpha}$ so as to maximally suppress the base-stock level. Similarly, When $w = p$ the unconstrained optimal β_g value is $\beta_g(p, y_0(p)) = \bar{\alpha} + \frac{h_0}{p} > \bar{\alpha}$, see (2.8); hence, the quasi-concavity of the supplier's profit $\hat{\Pi}_s(\cdot|p)$ in β_g implies that $\beta_g^*(p) = \bar{\alpha}$, see Theorem 2.2(c). In other words, $[0, p]$ starts with a type-1 interval $[0, w_1]$, for some $0 < w_1 < p$ and ends with a type-3 interval $[w_2, p]$.

Consider first the interval $[0, w_1)$ or the *interior* of any other type-1 interval (w', w'') . On such intervals, $\Pi_s^*(\cdot) = \hat{\Pi}_s(\cdot|\underline{\alpha})$. By Theorem 2.1(b), $\hat{\Pi}_s(\cdot|\underline{\alpha})$ is quasi-concave on the full interval $[0, p]$ with an *interior* point $w_{\underline{\alpha}}^*$ as its unique local maximum. This implies that only $w_{\underline{\alpha}}^*$ may arise as a local maximum of $\Pi_s^*(\cdot)$, on $[0, w_1)$ and the interior of any other type-1 interval. A similar argument shows that only $w_{\bar{\alpha}}^*$ may arise as a local maximum of $\Pi_s^*(\cdot)$, on $(w_2, p]$ and the interior of any other type-3 interval.

It remains to be shown that, except for $w_{\underline{\alpha}}^*$ and $w_{\bar{\alpha}}^*$, no other local maximum may arise. As argued, any additional local maximum must be part of a type-2 interval. (With the exceptions of $w = 0$ and $w = p$, boundary points of type-1 or type-3 intervals are also of a type-2 interval. This follows from the fact that if a type-1 or

type-3 interval is adjacent to another interval, the adjacent interval must be type-2, an observation following from the continuity of

$$y_w^* = \min\{\max\{y_0(w), \underline{y}(w)\}, \bar{y}(w)\} \quad (\text{A.15})$$

and the fact that $\bar{y}(w) > \underline{y}(w)$. The continuity of $y_w^*(\cdot)$ simply follows from that of $y_0(\cdot)$, $\underline{y}(\cdot)$ and $\bar{y}(\cdot)$, see Theorem 2.3(a) and (2.10); $\bar{y}(w) > \underline{y}(w)$ is immediate from (2.10).

Consider first the special pattern of (b-ii) in Theorem 3 where $y_0(\cdot) = y^p$ is constant. We show that the interval $[w_1, w_2]$ is a *single* type-2 interval, i.e., it does not contain type-1 or type-3 subintervals. Since w_1 belongs to a type-1 and type-2 interval, $y_0(w_1) = \underline{y}(w_1)$, so that $y_0(w) > \underline{y}(w)$ for all $w > w_1$ as $y_0(\cdot)$ is constant and $\underline{y}(\cdot)$ is strictly decreasing, see (2.10). Thus no point in $(w_1, w_2]$ is of type-1. Similarly, since $y_0(w_2) = \bar{y}(w_2)$, we have $y_0(w) < \bar{y}(w)$ for all $w < w_2$ as $y_0(\cdot)$ is constant and $\bar{y}(\cdot)$ is strictly decreasing. Thus no point in $[w_1, w_2)$ is of type-3 either. We conclude that on $[w_1, w_2]$ $\Pi_s^*(w) = \Pi_s(y_w^*|w) = \Pi_s(y_0(w)|w) = \Pi_s(y^p|w)$; It follows from (2.9) that $\Pi_s^*(\cdot)$ is a linear function on $[w_1, w_2]$. If this linear function fails to be constant, none of the points in $[w_1, w_2]$ may arise as a local maximum since w_1 and w_2 are interior points of $[0, p]$. The remaining case within the pattern (b-ii) has $\Pi_s^*(\cdot)$ constant on the entire interval $[w_1, w_2]$. In this case, only w_1 and w_2 may arise as the local maxima, but $w_1 = w_{\underline{\alpha}}^*$ and $w_2 = w_{\bar{\alpha}}^*$. To verify that latter, note that since $\Pi_s^*(\cdot)$ is differentiable everywhere,

$$\begin{aligned} 0 &= \Pi_s^{*'}(w_1) = \hat{\Pi}'_s(w_1|\underline{\alpha}) \\ 0 &= \Pi_s^{*'}(w_2) = \hat{\Pi}'_s(w_2|\bar{\alpha}) \end{aligned}$$

where the second equality in each of the above equations follows from the fact that $\Pi_s^*(\cdot)$ coincides with $\hat{\Pi}'_s(w_1|\underline{\alpha})[\hat{\Pi}'_s(w_1|\bar{\alpha})]$ to the left [right] of $w_1[w_2]$. Thus if w_1 is a local maximum of the quasi-concave function $\hat{\Pi}'_s(\cdot|\underline{\alpha})$ and, by Theorem 1(b) $w_1 = w_{\underline{\alpha}}^*$; By the same argument, $w_2 = w_{\bar{\alpha}}^*$ if it is a local maximum.

To complete the proof, it suffices to show that under the remaining patterns (b-i), (b-iii) and (b-iv) no point in a type-2 interval $[w', w'']$ may arise as a local

maximum. On such an interval, $\Pi_s^*(\cdot) = \Pi_s^u(\cdot) \equiv \max_{y \geq L} \Pi_s(y|w)$. We will show that $\Pi_s^u(\cdot)$ is twice differentiable with $\Pi_s^{u''}(\cdot) > 0$, precluding any of the points in this interval from being a local maximum. (The conclusion is immediate for the interior of $[w', w'']$; Since $\Pi_s^u(\cdot)$ is differentiable, if w' were a local maximum, $\Pi_s^{*'}(w') = \Pi_s^{u'}(w') = 0$ and the strict convexity of $\Pi_s^u(\cdot)$ on $[w', w'']$ implies that $\Pi_s^*(\cdot)$ is strictly increasing to the right of w' , so that w' fails to be a local maximum; Similar arguments preclude w'' from being a local maximum.

We complete the proof by showing that $\Pi_s^u(\cdot)$ is twice differentiable with $\Pi_s^{u''}(\cdot) > 0$. It follows from the above mentioned envelope Theorem 7.3 in Fiacco and Kyparisis (1985) that $\Pi_s^u(w)$ is continuously differentiable with $\Pi_s^{u'}(w) = \frac{\partial \Pi_s(y_0(w)|w)}{\partial w}$. (The conditions for Theorem 7.3 in Fiacco and Kyparisis (1985) are the same as those in Theorem 4.1 there. The latter were verified in the proof of Theorem 2.3(a).) Employing (2.9) for any $w \leq w_2$, we get

$$\begin{aligned} \Pi_s^{u'}(w) &= \frac{\partial \Pi_s(y_0(w)|w)}{\partial w} \\ &= s(y_0(w)) - \left(\bar{\alpha} + \frac{1}{F(y_0(w))} - 1 \right) h(y_0(w)) \\ &= y_0(w) - \left(\bar{\alpha} + \frac{1}{F(y_0(w))} \right) h(y_0(w)). \end{aligned} \quad (\text{A.16})$$

Any point w in a type-2 interval has

$$y_0(w) \geq \underline{y}(w) > \underline{y}(\bar{w}) = L \quad (\text{A.17})$$

where the second inequality follows from the fact that $\underline{y}(\cdot)$ is strictly decreasing and $w \leq w_2 < \bar{w} = p$. The proof of Theorem 2.3(b) shows that since $y_0(w) \geq L$, $y_0(w)$ is continuously differentiable at w . Differentiating both sides of (A.16), we get

$$\Pi_s^{u''}(w) = F(y_0(w))[-\bar{\alpha} + \kappa(y_0(w))]y_0'(w) > 0$$

where the inequality follows from the fact that, under patterns (b-i), (b-iii) and (b-iv), $y_0'(w) \neq 0$ and always has the same sign as $-\bar{\alpha} + \kappa(y_0(w))$ whenever $y_0(w) > L$, see Theorem 2.3(b) and its proof.

Case 2: $\bar{w} < p$

As in the proof of case 1, the interval $[0, \bar{w}]$ is divided into a sequence of type-1, type-2 or type-3 intervals. This proof has also verified that neither $w = 0$ nor $w = p$ may arise as a local maximum; However, when $\bar{w} < p$, this end point of the feasible range $[0, \bar{w}]$ may be a local maximum. It remains to be shown which points in $(0, \bar{w})$ are candidates for a local maximum.

Following the proof of case 1, the only points in the *interior* of type-1 and type-3 interval that may arise as a local maximum are $w_{\underline{\alpha}}^*$ and $w_{\bar{\alpha}}^*$ respectively. All other points in $(0, \bar{w})$ either belong to a type-2 interval or are end points of a type-1 or type-3 interval. However, as shown in the proof of case 1, all such end points of type-1 or type-3 intervals all belong to a type-2 interval, and all points in a type-2 interval that belong to $(0, \bar{w})$ can be eliminated as a potential local maximum. Thus, we conclude that only $w_{\underline{\alpha}}^*$ and $w_{\bar{\alpha}}^*$ as well as \bar{w} may arise as a local maximum.

It remains to be shown that $\Pi_s^*(\cdot)$ is at most bimodal. Assume, to the contrary, that the function has three local maxima $w^1 < w^2 < \bar{w}$ with w^1 and w^2 being either $w_{\underline{\alpha}}^*$ or $w_{\bar{\alpha}}^*$. As shown above, w^2 is an interior point of a type-1 or type-3 interval $[w', w'']$, on which $\Pi_s^*(\cdot)$ coincides with the quasi-concave function $\hat{\Pi}_s(\cdot|\underline{\alpha})$ or $\hat{\Pi}_s(\cdot|\bar{\alpha})$, so that $\Pi_s^{*'}(w') > 0$ and $\Pi_s^{*'}(w'') < 0$, see Theorem 2.1. This implies that $w'' < \bar{w}$, so that $[w', w'']$ is *succeeded* by a type-2 interval. The presence of a smaller local maximum $w^1 < w^2$ reveals that $[w', w'']$ is also *preceded* by a type-2 interval. As shown in the proof of case 1, $\Pi_s^*(\cdot)$ coincides with the convex, differentiable function $\Pi_s^u(\cdot)$ on a type-2 interval, that is $\Pi_s^{u'}(w') = \Pi_s^{*'}(w') > 0 > \Pi_s^{*'}(w'') = \Pi_s^{u'}(w'')$, contradicting the convexity of $\Pi_s^u(\cdot)$ and hence the possibility of three local maxima.

(b1-1) In part (a2) we have shown that the interval $[w_1, w_2]$ is a single type-2 interval under pattern (b-ii) where $y_0(\cdot)$ is constant. The same arguments apply when $y_0(\cdot)$ is decreasing, except that $y_0(w_2) \leq \bar{y}(w_2)$ as opposed to $y_0(w_2) = \bar{y}(w_2)$ there (, where the strict inequality may occur when $w_2 = \bar{w}$).

(b1-2) It follows from the proof of part (a2) that $\beta_g^*(w) = \beta_g(w, \underline{y}(w)) = \underline{\alpha}$ when $w \leq w_1$ and $\beta_g^*(w) = \beta_g(\bar{y}(w)) = \bar{\alpha}$ when $w \geq w_2$. For $w_1 < w < w_2$, $\beta_g^*(w) < \bar{\alpha}$. Thus, by (2.8), $\bar{\alpha} - \beta_g^*(w) = \bar{\alpha} - \beta_g(w, y_0(w)) = \frac{(p-w)(1-F(y_0(w))) - h_0 F(y_0(w))}{w F(y_0(w))} > 0$. Moreover, since $y_0(\cdot)$ is an increasing function, $\bar{\alpha} - \beta_g^*(w)$ is decreasing as the ratio

of a positive decreasing function and a positive increasing function. Hence $\beta_g^*(w)$ is increasing for all $w_1 < w < w_2$.

(b2) It follows immediately from (A.15) and the fact that $y_0(\cdot)$, $\underline{y}(\cdot)$ and $\bar{y}(\cdot)$ are all decreasing functions of $w \in [0, \bar{w}]$.

(c) It follows immediately from part (a2). \square

Proof of Proposition 2.2.

It follows from Proposition 2.1 and Theorem 2.4(b1) that two threshold points $0 < w_1 \leq w_2 \leq \bar{w}$ exist such that (i) $[0, w_1]$ is of type-1, (ii) $(w_1, w_2]$ is of type-2 and (iii) $(w_2, \bar{w}]$ is of type-3. We first show that under the specified parameter conditions, $w_1 < w_2$, i.e., a complete type-2 interval exists. For the classes of uniform and exponential distributions, this follows from the proof of Theorem 2.4(a2) since $\bar{w} = p$. For the Normals, it suffices to show that $y_0(\bar{w}) > L = \underline{y}(\bar{w})$ where $L = \mu - 1.8\sigma$ and $\bar{w} = \frac{p(1-F(L))-F(L)h_0}{1-F(L)+F(L)(\bar{\alpha}-\underline{\alpha})}$, which implies that \bar{w} belongs to a type-2 or a type-3 interval. (Recall from the proof of Theorem 2.4(a2) that a type-1 interval can not be followed by a type-3 interval unless a type-2 interval is adjacent to it.) Since the function $\Pi_s(\cdot|\bar{w})$ is quasi-concave, see Theorem 2.2, $y_0(\bar{w}) > L = \underline{y}(\bar{w})$ follows by showing that

$$\begin{aligned}
\Pi'_s(L|\bar{w}) &= (p-c)(1-F(L)) - (\bar{w}\bar{\alpha} + h_0)F(L) - (p-\bar{w})\frac{h(L)f(L)}{F^2(L)} \\
&> (p-c)(1-F(L)) - pF(L) - p\frac{F(L)}{1-F(L)}\frac{h(L)f(L)}{F^2(L)} \\
&= (p-c)(1-0.036) - 0.036p(1+0.906) \\
&= 0.895p - 0.964c \\
&= 0.895(p - 1.077c) \geq 0.
\end{aligned}$$

The first equality follows from (2.12) and the first inequality follows from $\bar{w}\bar{\alpha} + h_0 \leq p$ (since $\bar{\alpha} \leq \max\{\alpha_r, \alpha_s\}$) and $p - \bar{w} = \frac{(p(\bar{\alpha}-\underline{\alpha})+h_0)F(L)}{1-F(L)+F(L)(\bar{\alpha}-\underline{\alpha})} < \frac{(p(\bar{\alpha}-\underline{\alpha})+h_0)F(L)}{1-F(L)} \leq \frac{pF(L)}{1-F(L)}$ (since $\bar{\alpha} - \underline{\alpha} \leq \max\{\alpha_r, \alpha_s\}$); The second equality follows from $F(L) = 0.036$ and $\frac{h(L)f(L)}{F^2(L)(1-F(L))} = \frac{(l\Phi(l)+\phi(l))\phi(l)}{\Phi^2(l)(1-\Phi(l))} = 0.906$ with $l = \frac{L-\mu}{\sigma} = -1.8$; Finally the last inequality follows from $\frac{p-c}{c} \geq 7.7\%$.

Since, in all parts (a)-(c), the initial type-1 interval is followed by a full type-2 interval $(w_1, w_2]$, the conclusions in these parts follow by showing that $\Pi_s^{*'}(\cdot) > 0$

on this type-2 interval. (Since $\Pi_s^*(\cdot)$ is quasi-concave on the type-1 interval $[0, w_1]$, among the points in $[0, w_2]$, only $w = w_2$ may arise as a local optimum and this only when $w_2 = \bar{w}$. If $w_2 < \bar{w}$, all local maxima are restricted to the last, type-3, interval $(w_2, \bar{w}]$, where only \bar{w} or w_α^* arises as the (unique) local maximum.) On type-2 intervals $\Pi_s^*(\cdot) = \Pi_s^u(\cdot)$, and, by (A.16),

$$\Pi_s^{u'}(w) = \omega(y_0(w)) \quad (\text{A.18})$$

where $\omega(y) = y - \left(\bar{\alpha} + \frac{1}{F(y)}\right) h(y)$ is strictly increasing in $y \in [L, y_0(w_1))$, since $\omega'(y) = 1 - \left(\bar{\alpha} + \frac{1}{F(y)}\right) F'(y) + \frac{h(y)F'(y)}{F^2(y)} = F'(y)(\kappa(y) - \bar{\alpha}) > F'(y)(\kappa(y_0(w_1)) - \bar{\alpha}) \geq 0$. (The first inequality holds since $\kappa(\cdot)$ is strictly decreasing in $y \geq L$ for all three classes of distributions considered here, see (A.10); the second inequality follows from (A.9) since $y_0(\cdot)$ is increasing, see Proposition 2.1(c).) Therefore, since $y_0(w_1) > L$, see (A.17), we have, for any $w \in [w_1, w_2]$, that $\omega(L) < \omega(y_0(w_1)) = \Pi_s^{u'}(w_1) \leq \Pi_s^{u'}(w)$ where the second inequality follows from the convexity of $\Pi_s^{u'}(\cdot)$. It remains to be shown that $\omega(L) \geq 0$. For the uniform and exponential distributions, it is immediate from the facts that these distributions have $L = 0$ and a support on the non-negative half line, as well as $\lim_{y \rightarrow 0} \omega(y) = \lim_{y \rightarrow 0} \frac{h(y)}{F(y)} = 0$ (To verify the second limit, note that $0 \leq \frac{h(y)}{F(y)} = \frac{\int_0^y F(t)dt}{F(y)} \leq y$.) For the Normal distributions, recall that $l = \frac{L-\mu}{\sigma} = -1.8$, $\phi(l) = 0.079$ and $\Phi(l) = 0.036$.

$$\begin{aligned} \omega(L) &= L - \left(\bar{\alpha} + \frac{1}{F(L)}\right) h(L) \\ &= \mu + l\sigma - \left(\bar{\alpha} + \frac{1}{\Phi(l)}\right) \sigma[l\Phi(l) + \phi(l)] \\ &= \sigma \left[\frac{\mu}{\sigma} - \frac{\phi(l)}{\Phi(l)} - \bar{\alpha}(l\Phi(l) + \phi(l)) \right] \\ &= \sigma \left(\frac{\mu}{\sigma} - 2.197 - 0.014\alpha_g \right) \\ &\geq \sigma \left(\frac{\mu}{\sigma} - 2.211 \right) > 0 \end{aligned}$$

when $\frac{\sigma}{\mu} \leq \frac{1}{2.211} \approx 0.45$ where the first inequality uses the fact that $\alpha_g \leq 1$. \square

Proof of Proposition 2.3.

Recall from Proposition 2.2 and its proof that $w^* = w_\alpha^*$ or \bar{w} , and that the wholesale price range $[0, \bar{w}]$ is partitioned into type-1, type-2 and type-3 intervals, at most one of each. We first prove the following claims:

Claim 1: If, for $\bar{\alpha} = \alpha_1$, $w_{\alpha_1}^*$ is in the type-3 interval, then $w_{\bar{\alpha}}^*$ is increasing at the point $\bar{\alpha} = \alpha_1$.

Claim 2: \bar{w} is continuous and decreasing in $\bar{\alpha}$.

Claim 3: If \bar{w} is the optimal wholesale price at $\bar{\alpha} = \alpha_1$, then it remains optimal for any $\bar{\alpha} > \alpha_1$.

Proof of claim 1: By (2.3), under the special case $\beta_g = \bar{\alpha}$, we have that $y_{\bar{\alpha}}^* = F^{-1}\left(1 - \frac{h_0}{p - w_{\bar{\alpha}}^* + h_0}\right)$. It is thus equivalent to show that $y_{\bar{\alpha}}^*$ is decreasing at $\bar{\alpha} = \alpha_1$. Note that

$$y_{\bar{\alpha}}^* = \operatorname{argmax}_{y \geq 0} \Pi_s(y|\bar{\alpha}) \quad (\text{A.19})$$

where $\Pi_s(y|\bar{\alpha}) = w(y)(s(y) - \bar{\alpha}h(y)) - cs(y)$ and $w(y) = \frac{p - (p+h_0)F(y)}{1 - F(y)} = p - h_0 \frac{F(y)}{1 - F(y)}$, see (2.4) and (2.5). Since it follows from Theorem 2.1 that $\Pi'_s(y_{\bar{\alpha}}^*|\bar{\alpha}) = 0$ and $\Pi''_s(y_{\bar{\alpha}}^*|\bar{\alpha}) < 0$, by the Implicit Function Theorem, it is sufficient to show that

$$\begin{aligned} \frac{\partial \Pi'_s(y|\beta_g = \bar{\alpha})}{\partial \bar{\alpha}} &= -w(y)F(y) - w'(y)h(y) \\ &= -\left(p - h_0 \frac{F(y)}{1 - F(y)}\right)F(y) + \frac{h_0 f(y)}{(1 - F(y))^2}h(y) \\ &= \frac{pF^2(y)}{1 - F(y)} \left(-\frac{1 - F(y)}{F(y)} + \frac{h_0}{p}(\zeta(y) + 1)\right) \leq 0 \end{aligned} \quad (\text{A.20})$$

at $y_{\alpha_1}^*$. (The first equality follows from the fact that $\frac{\partial \Pi'_s(y|\beta_g = \bar{\alpha})}{\partial \bar{\alpha}} = -(w(y)h(y))'$ since $\frac{\partial \Pi_s(y|\beta_g = \bar{\alpha})}{\partial \bar{\alpha}} = -w(y)h(y)$.) However, (A.20) holds at $y_{\alpha_1}^*$ since, in the second Stackelberg game with $w = w_{\alpha_1}^*$,

$$\begin{aligned} 0 &\leq \Pi'_s(y_{\alpha_1}^*|w = w_{\alpha_1}^*) \\ &= (p - c)(1 - F(y_{\alpha_1}^*)) - (w_{\alpha_1}^* \alpha_1 + h_0)F(y_{\alpha_1}^*) - (p - w_{\alpha_1}^*) \frac{h(y_{\alpha_1}^*)f(y_{\alpha_1}^*)}{F^2(y_{\alpha_1}^*)} \\ &= (p - c)(1 - F(y_{\alpha_1}^*)) - (w_{\alpha_1}^* \alpha_1 + h_0)F(y_{\alpha_1}^*) - \frac{h_0 F(y_{\alpha_1}^*)}{1 - F(y_{\alpha_1}^*)} \frac{h(y_{\alpha_1}^*)f(y_{\alpha_1}^*)}{F^2(y_{\alpha_1}^*)} \\ &= (p - c)(1 - F(y_{\alpha_1}^*)) - (w_{\alpha_1}^* \alpha_1 + h_0)F(y_{\alpha_1}^*) - h_0 F(y_{\alpha_1}^*) \zeta(y_{\alpha_1}^*) \\ &= pF(y_{\alpha_1}^*) \left(-\frac{c(1 - F(y_{\alpha_1}^*))}{pF(y_{\alpha_1}^*)} - \frac{w_{\alpha_1}^* \alpha_1}{p} + \frac{1 - F(y_{\alpha_1}^*)}{F(y_{\alpha_1}^*)} - \frac{h_0}{p}(1 + \zeta(y_{\alpha_1}^*))\right) \\ &\leq pF(y_{\alpha_1}^*) \left(\frac{1 - F(y_{\alpha_1}^*)}{F(y_{\alpha_1}^*)} - \frac{h_0}{p}(1 + \zeta(y_{\alpha_1}^*))\right) \\ \iff &-\frac{1 - F(y_{\alpha_1}^*)}{F(y_{\alpha_1}^*)} + \frac{h_0}{p}(1 + \zeta(y_{\alpha_1}^*)) \leq 0. \end{aligned}$$

(The first inequality follows from $\Pi_s(\cdot|w_{\alpha_1}^*)$ being quasi-concave and reaching its maximum at $y_0(w_{\alpha_1}^*)$, see Theorem 2.2, as well as the fact that $y_{\alpha_1}^* = y(w_{\alpha_1}^*, \alpha_1) = \bar{y}(w_{\alpha_1}^*) < y_0(w_{\alpha_1}^*)$ since $w_{\alpha_1}^*$ is part of the type-3 interval. The first equality follows from (2.12). The second equality is due to the fact that $(p - w_{\alpha_1}^*)(1 - F(y_{\alpha_1}^*)) = h_0 F(y_{\alpha_1}^*)$, by applying (2.3) with $\beta_g = \alpha_1$.)

Proof of claim 2:

The claim follows from (2.11) and the fact that $\underline{y}(\cdot)$ is point-wise decreasing in $\bar{\alpha}$, see (2.10), and continuous in $(\bar{\alpha}, w)$.

Proof of claim 3:

We attach a subscript α_1 to \bar{w} for its values when $\bar{\alpha} = \alpha_1$ and write $y_0(w|\bar{\alpha} = \alpha_1)$ to reflect the dependency of $y_0(w)$ on $\bar{\alpha}$, when necessary. Let $\alpha_0 = \inf\{\bar{\alpha} \geq \alpha_1 : \bar{w}_{\bar{\alpha}}$ is of type 2}, if $\bar{w}_{\bar{\alpha}}$ is of type-2 for some value of $\bar{\alpha}$; otherwise, $\alpha_0 = \infty$. We first prove the claim for $\bar{\alpha} \in [\alpha_1, \alpha_0)$ with $\alpha_0 > \alpha_1$. Since it follows from the proof of Proposition 2.2 that \bar{w} is either of type-2 or type-3, $\bar{w}_{\bar{\alpha}}$ is part of the type-3 interval for all $\bar{\alpha} \in [\alpha_1, \alpha_0)$. Since \bar{w}_{α_1} is in the type-3 interval and since it is optimal for $\bar{\alpha} = \alpha_1$, $\bar{w}_{\alpha_1} \leq w_{\alpha_1}^*$. (For $\bar{\alpha} = \alpha_1$, in the type-3 interval, $\Pi_s^*(\cdot) = \hat{\Pi}_s(\cdot|\beta_g = \bar{\alpha} = \alpha_1)$ has a unique local maximum of $w_{\alpha_1}^*$, see the proof of Proposition 2.2; therefore, if $\bar{w}_{\alpha_1} > w_{\alpha_1}^*$, $w_{\alpha_1}^*$ would be optimal for $\bar{\alpha} = \alpha_1$, contradicting the fact that \bar{w}_{α_1} is the optimal wholesale price.) As $\bar{\alpha}$ is increased continuously from α_1 to α_0 , $\bar{w}_{\bar{\alpha}}$ decreases while remains in the type-3 interval, by Claim 2 and the definition of α_0 . This, together with Claim 1, implies that $w_{\bar{\alpha}}^*$ increases on $[\alpha_1, \alpha_0)$ while remaining in the type-3 interval as well. In other words, for all $\bar{\alpha} \in [\alpha_1, \alpha_0)$, $\bar{w}_{\bar{\alpha}} \leq w_{\bar{\alpha}}^*$ so that $\bar{w}_{\bar{\alpha}}$ is the optimal wholesale price.

To complete the proof of Claim 3, it remains to be shown that $\bar{w}_{\bar{\alpha}}$ is optimal for all $\bar{\alpha} \geq \alpha_0$ and $\alpha_0 < \infty$. By the definition of α_0 , there exists a sequence of $\{\alpha_n\} \downarrow \alpha_0$, such that the values \bar{w}_{α_n} are of type-2, i.e.,

$$y_0(\bar{w}_{\alpha_n}|\bar{\alpha} = \alpha_n) < \bar{y}(\bar{w}_{\alpha_n}). \quad (\text{A.21})$$

y_0 is a continuous function of the wholesale price w and also of $\bar{\alpha}$ and $\bar{y}(\cdot)$ is continuous as well. It follows from Claim 2 that both the left and right hand sides of

(A.21) are continuous in $\bar{\alpha}$. Thus letting $n \rightarrow \infty$, we conclude that

$$y_0(\bar{w}_{\alpha_0} | \bar{\alpha} = \alpha_0) \leq \bar{y}(\bar{w}_{\alpha_0}). \quad (\text{A.22})$$

It follows that for any $\alpha_2 > \alpha_0$:

$$y_0(\bar{w}_{\alpha_2} | \bar{\alpha} = \alpha_2) \leq y_0(\bar{w}_{\alpha_2} | \bar{\alpha} = \alpha_0) \leq y_0(\bar{w}_{\alpha_0} | \bar{\alpha} = \alpha_0) \leq \bar{y}(\bar{w}_{\alpha_0}) \leq \bar{y}(\bar{w}_{\alpha_2}), \quad (\text{A.23})$$

so that \bar{w}_{α_2} is of type-2 and, by the proof of Proposition 2.2, optimal for $\bar{\alpha} = \alpha_2$. (To verify the first inequality of (A.23), note that, by (2.12), $\partial \Pi'_s(y|w)/\partial \bar{\alpha} = -wF(y) < 0$. Theorem 2.2 implies that $\Pi'_s(y_0(w)|w) = 0$ and $\Pi''_s(y_0(w)|w) \leq 0$. Hence, by the Implicit Function Theorem, for any given w ,

$$y_0(w) \text{ is decreasing in } \bar{\alpha}. \quad (\text{A.24})$$

The second inequality follows from the fact that

$$y_0(\cdot) \text{ is increasing in } w, \quad (\text{A.25})$$

see Proposition 2.1 while $\bar{w}_{\alpha_2} \leq \bar{w}_{\alpha_0}$, see Claim 2. The third inequality is given by (A.22) and the last one by the fact that $\bar{y}(\cdot)$ is decreasing and $\bar{w}_{\alpha_2} \leq \bar{w}_{\alpha_0}$.

We are now ready to prove the statements in parts (a) and (b).

Part (a): Recall that, under TC, $\bar{\alpha} = \alpha_s$. The monotonicity properties of w^* are immediate from Claims 1 and 3. Since $\underline{\alpha} = \alpha_s - \alpha_r$ so that $\bar{\alpha} - \underline{\alpha} = \alpha_r$ is independent of $\bar{\alpha} = \alpha_s$, and the function $y(\cdot)$ and hence \bar{w} are constant in $\bar{\alpha}$. $w^* = \bar{w}$ for some $\alpha_s = \alpha_s^c$, and w^* remains constant, see (2.11). Clearly, for $\alpha_s \leq \alpha_s^c$, $\beta_g^* = \bar{\alpha}$; Similarly, for $\alpha_s^c \leq \bar{\alpha} = \alpha_s \leq \alpha_0$, $\bar{w}_{\bar{\alpha}}$ is the optimal wholesale price and is of type-3 so that on this interval, as well, $\beta_g^* = \bar{\alpha}$. In other words, for all $\alpha_s \leq \alpha_0$, $\alpha^* = \beta_g^* = \bar{\alpha}$. Moreover, for $\bar{\alpha} \geq \alpha_0$, \bar{w} is of type-2 so that $\alpha^* = \bar{\alpha} - \beta_g^* = \bar{\alpha} - \beta_g(\bar{w}, y_0(\bar{w})) = \left(\frac{p}{\bar{w}} - 1\right) \left(\frac{1}{F(y_0(\bar{w}))} - 1\right) - \frac{h_0}{\bar{w}} = \frac{(p-\bar{w})(1-F(y_0(\bar{w}))) - h_0 F(y_0(\bar{w}))}{\bar{w} F(y_0(\bar{w}))}$ where the third equality follows from (2.8). It is easily verified that the numerator of this expression is decreasing in \bar{w} as well as $y_0(\bar{w})$, while the denominator is increasing in both quantities. Therefore α^* is decreasing in \bar{w} and $y_0(\bar{w})$. Since both of these quantities are decreasing in $\bar{\alpha}$ by Claim 2 and (A.24) as well as (A.25), it follows that α^* is increasing in $\bar{\alpha} = \alpha_s$.

Part (b): Under IS, $\bar{\alpha} = \alpha_r$. The proof is analogous to that of part (a). \square

Proof of Theorem 2.5.

Part (a): Recall that $\delta^{TC} = \bar{\alpha}^{TC} - \beta_g^{TC}$ and $\delta^{IS} = \bar{\alpha}^{IS} - \beta_g^{IS}$. Similarly let $\delta^{*TC}(w) = \bar{\alpha}^{TC} - \beta_g^{*TC}(w)$ and $\delta^{*IS}(w) = \bar{\alpha}^{IS} - \beta_g^{*IS}(w)$. We first prove the following claim:

$$\pi_s^{TC}(w, \beta_g^{TC}, y(w, \beta_g^{TC})) \geq \pi_s^{IS}(w, \beta_g^{IS}, y(w, \beta_g^{IS})) \text{ for } \beta_g^{IS} \text{ and } \beta_g^{TC} \text{ such that } \delta^{TC} = \delta^{IS}. \quad (\text{A.26})$$

(A.26) holds because the equality $\delta^{TC} = \delta^{IS}$, by (2.3), implies that $y(w, \beta_g^{TC}) = y(w, \beta_g^{IS})$. $\delta^{TC} = \delta^{IS}$ also implies that $\beta_g^{IS} \geq \beta_g^{TC}$ since $\bar{\alpha}^{IS} \geq \bar{\alpha}^{TC}$. The claim thus follows immediately from (2.2).

Choose β_g^{TC} such that $\delta^{TC} = \bar{\alpha}^{TC} - \beta_g^{TC} = \delta^{*IS}(w)$. (This is achievable since $\delta^{*IS}(w) \in [0, \alpha_r]$ and δ^{TC} continuously decreases from α_r to 0 as β_g^{TC} varies from $\underline{\alpha}^{TC}$ to $\bar{\alpha}^{TC}$.) Thus by (A.26),

$$\begin{aligned} \pi_s^{TC}(w, \beta_g^{*TC}(w), y(w, \beta_g^{*TC}(w))) &\geq \pi_s^{TC}(w, \beta_g^{TC}, y(w, \beta_g^{TC})) \\ &\geq \pi_s^{IS}(w, \beta_g^{*IS}(w), y(w, \beta_g^{*IS}(w))). \end{aligned}$$

Part (b): Since the above inequality implies that

$$\max_{w \in [0, p]} \pi_s^{TC}(w, \beta_g^{*TC}(w), y(w, \beta_g^{*TC}(w))) \geq \max_{w \in [0, p]} \pi_s^{IS}(w, \beta_g^{*IS}(w), y(w, \beta_g^{*IS}(w))),$$

part (b) follows immediately. \square

The following lemma is used in the proof of Theorem 2.6.

Lemma A.1 *Both w_1^{TC} and w_2^{TC} are smaller than or equal to (greater than) w_1^{IS} and w_2^{IS} , respectively, if $\alpha_s \leq (>) \alpha_r$.*

Proof of Lemma A.1.

Assume $\alpha_r > \alpha_s$. (The other case can be argued analogously.) By (2.12), $\partial \Pi'_s(y|w) / \partial \bar{\alpha} = -wF(y) < 0$. Theorem 2.2 implies that $\Pi'_s(y|w)$ is decreasing at y_0 . Hence, by the Implicit Function Theorem,

$$y_0 \text{ is decreasing in } \bar{\alpha} \text{ and } y_0^{IS}(w) \leq y_0^{TC}(w) \text{ for any } w \in [0, \bar{w}] \quad (\text{A.27})$$

since $y_0^{IS}(\cdot)$ and $y_0^{TC}(\cdot)$ differ only in the choice of $\bar{\alpha}$ with $\bar{\alpha}^{IS} = \alpha_r \geq \alpha_s = \bar{\alpha}^{TC}$.

We prove first that $w_2^{IS} \geq w_2^{TC}$. Recall that $\bar{y}(w) = F^{-1}(\frac{p-w}{p-w+h_0})$ under both TC and IS, which is invariant with respect to $\bar{\alpha}$ and decreasing in w . Only one of the following two cases can occur.

(i) $0 < w_2^{TC} < \bar{w}$: In this case w_2^{TC} is the point of intersection of the functions $y_0^{TC}(\cdot)$ and $\bar{y}(\cdot)$. It follows from (A.27) that $y_0^{IS}(w_2^{TC}) \leq y_0^{TC}(w_2^{TC}) = \bar{y}(w_2^{TC})$. Thus, as w is decreased from w_2^{TC} , the distance between the functions $y_0^{IS}(\cdot)$ and $\bar{y}(\cdot)$ gets larger since the former decreases and the latter increases. Thus $w_2^{IS} \geq w_2^{TC}$.

(ii) $w_2^{TC} = \bar{w}$: $w_2^{IS} = w_2^{TC} = \bar{w}$ since $y_0^{IS}(w) \leq y_0^{TC}(w) \leq \bar{y}(w)$ for any $w \in [0, \bar{w}]$ by (A.27).

Next we show that $w_1^{IS} \geq w_1^{TC}$. Assume to the contrary that $w_1^{IS} < w_1^{TC}$. Recall that $\underline{y}(w) = F^{-1}(\frac{p-w}{p-w+w\alpha_r+h_0})$ under both the IS and TC arrangements. The definition of w_1 has the following implications: If $w_1 < \bar{w}$, w_1 is the point of intersection of the functions $\underline{y}(\cdot)$ and $y_0(\cdot)$; If $w_1 = \bar{w}$, $y_0(w) \leq \underline{y}(w)$ throughout the entire $[0, \bar{w}]$ region with $y_0(w) < \underline{y}(w)$ for $w < w_1 = \bar{w}$. Therefore

$$y_0(w_1) \leq \underline{y}(w_1). \quad (\text{A.28})$$

Thus, as $0 < w_1^{IS} < w_1^{TC} \leq \bar{w}$, we get $y_0^{TC}(w_1^{TC}) \geq y_0^{IS}(w_1^{TC}) \geq y_0^{IS}(w_1^{IS}) \geq \underline{y}(w_1^{IS}) > \underline{y}(w_1^{TC})$ while the first inequality follows from (A.27); the second is due to the increasing property of $y_0(\cdot)$; the third follows from (A.28) and the last is due to the decreasing property of $\underline{y}(\cdot)$. Thus $y_0^{TC}(w_1^{TC}) > \underline{y}(w_1^{TC})$, which, however, contradicts (A.28). \square

Proof of Theorem 2.6. Recall that under condition (C) the following properties apply: (P1) The demand distribution has IFR and MIFR properties. (P2) $y_0(\cdot)$ is increasing. (P3) Either $(w_\alpha^*, \bar{\alpha})$ or $(\bar{w}, \beta_g^*(\bar{w}))$ is the (unique) equilibrium solution of the full Stackelberg game. Consider the case $\alpha_s \leq \alpha_r$. The proof for the remaining case $\alpha_s > \alpha_r$ is analogous.

Part (A-1): It follows from (2.1) that for any given wholesale price w and any given base-stock level y , the retailer's profit level $\pi_r(w, \beta_g, y)$ depends on the credit parameters $\bar{\alpha}$ and β_g only via their difference $\delta = \bar{\alpha} - \beta_g$, the effective capital cost rate incurred by the retailer under the supplier's selected ECCR. Moreover, π_r is

decreasing in δ for any given w and y . We first show that $\delta^{*TC}(w) = \bar{\alpha}^{TC} - \beta_g^{*TC}(w)$ is smaller than $\delta^{*IS}(w) = \bar{\alpha}^{IS} - \beta_g^{*IS}(w)$ for any $w \in [0, \bar{w}]$:

$$\delta^{*TC}(w) \leq \delta^{*IS}(w). \quad (\text{A.29})$$

Thus for any $w \in [0, \bar{w}]$ and $y \geq 0$, $\pi_r(w, y|\delta^{*TC}(w)) \geq \pi_r(w, y|\delta^{*IS}(w))$ and $\max_{y \geq 0} \pi_r(w, y|\delta^{*TC}(w)) \geq \max_{y \geq 0} \pi_r(w, y|\delta^{*IS}(w))$, i.e., the retailer is better off under TC, as opposed to IS.

We consider the following three cases, with the understanding that when $w_1^{IS} < w_2^{TC}$ the first two cases (i) and (ii) cover the whole range by themselves.

(i) $w \in [0, w_1^{IS}]$: $\delta^{*IS}(w) = \bar{\alpha} - \beta_g^{*IS}(w) = \bar{\alpha} - \underline{\alpha} = \alpha_r$. Since $\delta^{TC}(w) \in [\bar{\alpha}^{TC} - \bar{\alpha}^{TC}, \bar{\alpha}^{TC} - \underline{\alpha}^{TC}] = [0, \alpha_r]$, (A.29) holds.

(ii) $w \in (w_2^{TC}, \bar{w}]$: $\delta^{*TC}(w) = \bar{\alpha}^{TC} - \bar{\alpha}^{TC} = 0$. Since $\delta^{IS}(w) \in [\bar{\alpha}^{IS} - \bar{\alpha}^{IS}, \bar{\alpha}^{IS} - \underline{\alpha}^{IS}] = [0, \alpha_r]$, (A.29) holds.

(iii) $w \in (w_1^{IS}, w_2^{TC}]$: By Lemma A.1, $w_1^{TC} \leq w_1^{IS} \leq w_2^{TC} \leq w_2^{IS}$, so that $y_w^* = y_0(w)$ in this region under both mechanisms. Since $\bar{\alpha}^{TC} = \alpha_s < \bar{\alpha}^{IS} = \alpha_r$, $y_w^{*IS} < y_w^{*TC}$ follows from (A.27). On the other hand, by (2.8), $\delta^*(w) = \bar{\alpha} - \beta_g(w, y_w^*) = \left(\frac{p}{w} - 1\right) \left(\frac{1}{F(y_w^*)} - 1\right) - \frac{h_0}{w}$ is decreasing in y_w^* . Thus (A.29) holds.

Part (A-2): It follows from (2.3) that for any given wholesale price $w \in [0, \bar{w}]$, the retailer's optimal equilibrium base-stock level $y(w, \beta_g^*(w))$ is strictly decreasing in $\delta^*(w)$. The claim is thus a direct result of (A.29).

Part (B-1): Under either IS or TC, we distinguish between the following two cases: (I) $w_2 < \bar{w}$; (II) $w_2 = \bar{w}$. It follows from the proof of Proposition 2.2 that the optimal wholesale price/ECCR pair is $(\min\{w_{\bar{\alpha}}^*, \bar{w}\}, \bar{\alpha})$ in case (I) and $(\bar{w}, \beta_g^*(\bar{w}))$ in case (II). $(\Pi_s^*(\cdot))$ is increasing on $[0, w_2)$, immediately explaining the solution in case (II). In case (I), $\Pi_s^*(\cdot)$ achieves its maximum in the type-3 interval $[w_2, \bar{w}]$ where the function is quasi-concave and $\beta_g^* = \bar{\alpha}$.) Let $\delta^* \equiv \bar{\alpha} - \beta_g^* \geq 0$. Since, by Lemma A.1, $w_2^{TC} \leq w_2^{IS}$, we only need to consider the following three cases:

(a) $w_2^{TC} = w_2^{IS} = \bar{w}$: $w^{*TC} = w^{*IS} = \bar{w}$. Thus the retailer's preference for TC follows from part (a).

(b) $w_2^{TC} < w_2^{IS} = \bar{w}$: $w^{*TC} = \min\{w_{\bar{\alpha}}^{*TC}, \bar{w}\} \leq w^{*IS} = \bar{w}$ and $\delta^{*TC} = \bar{\alpha}^{TC} - \beta_g^{*TC} = 0 \leq \delta^{*IS}$. Since $\pi_r(w, y|\delta^*)$ is decreasing in both w and δ^* for any given y ,

the retailer's profit is higher under TC than IS for any given y , which remains true when the profit values are maximized over all $y \geq 0$.

(c) $w_2^{TC} \leq w_2^{IS} < \bar{w}$: Since $\beta_g^* = \bar{\alpha}$ for both TC and IS, by (2.1), for any base-stock level y

$$\begin{aligned} \pi_r(w^{*TC}, \beta_g^{*TC}, y) &= (p - w^{*TC})s(y) - h_0h(y) \\ &\geq (p - w^{*IS})s(y) - h_0h(y) = \pi_r(w^{*IS}, \beta_g^{*IS}, y), \end{aligned} \quad (\text{A.30})$$

provided $w^{*TC} \leq w^{*IS}$. Maximizing the far left and far right expression in (A.30) over all $y \geq 0$, we conclude that the retailer is better off under TC as opposed to IS, provided we can show $w^{*TC} = \min\{w_{\bar{\alpha}^{TC}}^{*TC}, \bar{w}\} \leq \min\{w_{\bar{\alpha}^{IS}}^{*IS}, \bar{w}\} = w^{*IS}$. However, this follows from $w_{\bar{\alpha}^{TC}}^{*TC} \leq w_{\bar{\alpha}^{IS}}^{*IS}$ which is implied by Claim 1 in the proof of Proposition 2.2 since $\bar{\alpha}^{TC} = \alpha_s < \alpha_r = \bar{\alpha}^{IS}$.

Parts (B-2) and (B-3): It follows from the proof of cases (a)-(c) in part (B-1) that $w^{*TC} \leq w^{*IS}$ and $\delta^{*TC} \leq \delta^{*IS}$. (Note that in case (1), the latter follows from part (a) since $w^{*TC} = w^{*IS}$.) Since $y^* = y(w^*, \beta_g^*)$ is decreasing in both w^* and δ^* , parts (B-2) and (B-3) follow. \square

A.2 Algorithm of the Full Stackelberg Game with Participation Constraint

In this appendix we show how the full Stackelberg game with a participation constraint can be solved efficiently. The participation constraint has the form:

$$\pi_r(w, \beta_g, y) \geq \Pi_r^0 \quad (\text{A.31})$$

for a given minimum retailer's profit value Π_r^0 . We first solve the unconstrained full Stackelberg game, see Theorem 2.4(c), and evaluate the associated retailer's profit value Π_r^* . If $\Pi_r^* \geq \Pi_r^0$, the unconstrained optimal contract also solves the problem with the participation constraint; otherwise, (A.31) holds as an equality.

As before, the retailer's best response function (2.3) may be written as $\beta_g = \beta_g(w, y)$, see (2.8). It is, again, advantageous to assume the supplier selects a wholesale price w and a targeted base-stock level y , which can be implemented by adopting the associated ECCR $\beta_g = \beta_g(w, y)$ assuming $\underline{\alpha} \leq \beta_g \leq \bar{\alpha}$. Substituting (2.8) into (A.31) - written as an equality - we obtain

$$(p - w)\hat{s}(y) = \Pi_r^0 \quad (\text{A.32})$$

where $\hat{s}(y) \equiv s(y) - (\frac{1}{F(y)} - 1)h(y) = y - \frac{h(y)}{F(y)}$ is an increasing differentiable function, since

$$\hat{s}'(y) = 1 - \frac{F^2(y) - h(y)f(y)}{F^2(y)} = \frac{h(y)f(y)}{F^2(y)} \geq 0, \quad (\text{A.33})$$

with $\hat{s}(0) = 0$ since $\lim_{y \downarrow 0} \frac{h(y)}{F(y)} \leq \lim_{y \downarrow 0} \frac{yF(y)}{F(y)} = 0$, and

$$\lim_{y \rightarrow \infty} \hat{s}(y) = \lim_{y \rightarrow \infty} \frac{yF(y) - h(y)}{F(y)} = \lim_{y \rightarrow \infty} \frac{yF(y) - \int_0^y F(t)dt}{F(y)} = \lim_{y \rightarrow \infty} \frac{\int_0^y t dF(t)}{F(y)} = \mu$$

. It follows that, for any given wholesale price w , any retailer profit value can be achieved as long as $\Pi_r^0 \leq (p - w)\mu$, i.e., as long as $w \leq p - \frac{\Pi_r^0}{\mu} \equiv w^0$, and this with the *unique* base-stock level

$$y_t(w) \equiv \hat{s}^{-1} \left(\frac{\Pi_r^0}{p - w} \right). \quad (\text{A.34})$$

To verify whether this base-stock level can be feasibly implemented with an ECCR $\beta_g \in [\underline{\alpha}, \bar{\alpha}]$, it suffices to verify that $\underline{y}(w) \leq y_t(w) \leq \bar{y}(w)$. Assume

$$y_t(0) \leq \bar{y}(0) = F^{-1}\left(\frac{p}{p+h_0}\right) = \underline{y}(0). \quad (\text{A.35})$$

Since $y_t(\cdot)$ is increasing (as the composition of two increasing functions, see (A.34)) with $\lim_{w \uparrow w^0} y_t(w) = \infty$, and $\bar{y}(\cdot)$ is decreasing, there exists a unique value $w^2 < w^0$ such that $y_t(w^2) = \bar{y}(w^2)$ and $y_t(w) < \bar{y}(w)$ for all $w < w^2$. (When (A.35) is violated, $y_t(w) > \bar{y}(w)$ for all $0 \leq w \leq w^0$, i.e., there is no feasible solution for the constrained Stackelberg game.) Also, define w^1 as the unique point of intersection of the increasing function $y_t(\cdot)$ and the decreasing function $\underline{y}(\cdot)$, which exists, under (A.35). (Note, $w^1 \leq w^2$ since $\underline{y}(\cdot) \leq \bar{y}(\cdot)$ pointwise.) Thus, $y_t(w) < \underline{y}(w)$ for all $w \leq w^1$, so that no feasible solution exists for any $w < w^1$. These observations substantiate the following algorithm:

Algorithm (Constrained Stackelberg Game):

Step 0: Compute the optimal solution (w^*, β^*) of the unconstrained Stackelberg game and evaluate Π_r^* . If $\Pi_r^* \geq \Pi_r^0$, (w^*, β^*) solves the constrained problem; exit.

Step 1: If $y_t(0) > \bar{y}(0)$, no feasible solution exists; exit.

Step 2: Calculate w^1 (w^2), the unique point of intersection of the function $y_t(\cdot)$ and $\underline{y}(\cdot)$ [$\bar{y}(\cdot)$].

Step 3: Optimize the single variable supply chain profit function $(p - c)s(y_t(w)) - (\bar{\alpha}w + h_0)h(y_t(w))$ on the interval of $[w^1, w^2]$.

A.3 Proof for Section 8 of Chapter 2

In this appendix we prove that Theorems 2.1-2.4 and Propositions 2.1-2.2 hold for the following profit functions in the second generalized model with default risks (see §2.7.2):

$$\pi_r(w, \beta_g, y) = (p - (1 - \gamma)w)s(y) - [((1 - \gamma)\bar{\alpha} - \beta_g)w + h_0]h(y)$$

$$\pi_s(w, \beta_g, y) = (\phi(\gamma)w - c)s(y) - \beta_g w h(y)$$

with $0 \leq \beta_g \leq (1 - \gamma)\bar{\alpha}$ and $0 \leq w \leq p/(1 - \gamma)$. To simplify the exposition, we apply the following transformation of variables: $w^n = w\phi(\gamma)$ and $\beta^n = \beta_g/\phi(\gamma)$. Thus the above profit functions become:

$$\pi_r(w^n, \beta_g^n, y) = (p - \tilde{\gamma}w^n)s(y) - [(\frac{1 - \gamma}{\phi(\gamma)}\bar{\alpha} - \beta_g^n)w^n + h_0]h(y) \quad (\text{A.36})$$

$$\pi_s(w^n, \beta_g^n, y) = (w^n - c)s(y) - \beta_g^n w^n h(y) \quad (\text{A.37})$$

with $0 \leq \beta_g^n \leq \bar{\alpha}_d \equiv \tilde{\gamma}\bar{\alpha}$ and $0 \leq w^n \leq p_d \equiv p/\tilde{\gamma}$ where $\tilde{\gamma} \equiv \frac{1 - \gamma}{\phi(\gamma)}$. Without loss of clarity, we henceforth suppress the superscript “n” in w^n and β^n , while appending a subscript “d” to differentiate a quantity in the model with default risks from its counterpart in the base model.

The Stackelberg Game under a Given ECCR

Proof of Theorem 2.1

(a) The retailer’s optimal base-stock level $y(w, \beta_g)$ in response to given trade terms (w, β_g) is given by a specific fractile of the demand distribution. More specifically, $y(w, \beta_g)$ is the fractile that satisfies

$$F(y) = 1 - \frac{(\tilde{\gamma}\bar{\alpha} - \beta_g)w + h_0}{p - \tilde{\gamma}w + (\tilde{\gamma}\bar{\alpha} - \beta_g)w + h_0}. \quad (\text{A.38})$$

Let $w_d(y)$ be the inverse demand function. $w_d(y) = \frac{p - (p + h_0)F(y)}{\tilde{\gamma}(1 - (1 - \delta_d)F(y))}$ with $\delta_d = \bar{\alpha} - \beta_g/\tilde{\gamma}$. The supplier’s equilibrium profit as a function of y can be written as

$$\begin{aligned} \Pi_s(y|\beta_g) &= w_d(y)(s(y) - \beta_g h(y)) - cs(y), \\ &= w_d(y)\xi(y) - cs(y). \end{aligned} \quad (\text{A.39})$$

where $\xi(y) \equiv y - (1 + \beta_g)h(y)$. Differentiating (A.39) yields

$$\begin{aligned}\Pi'_s(y|\beta_g) &= w_d(y)\xi'(y) + \xi(y)w'_d(y) - c(1 - F(y)) \\ &= w_d(y)\xi'(y)g_d(y)\end{aligned}\tag{A.40}$$

where

$$g_d(y) = 1 - c \frac{1 - F(y)}{w_d(y)\xi'(y)} - \frac{-\xi(y)w'_d(y)}{\xi'(y)w_d(y)}\tag{A.41}$$

for $y < y_{max}$ (, see (2.7) for the definition of y_{max}). We show that $g_d(y)$ is decreasing on $[0, y_{max})$ by applying the similar arguments as in the proof of Theorem 2.1 for the base model. $g_d(y)$ is the same as its counterpart $g(y)$ in the original proof except that $w(y)$ is replaced by $w_d(y)$. The only properties of $w(y)$ employed in the original proof are that $w(y)$ is positive and decreasing (see the proof of (A.5)) and that $\frac{-w'(y)}{w(y)}$ is positive and increasing (see the proof of (A.6)). We show these properties hold for $w_d(y)$ too. $w_d(y) \geq 0$, and is clearly decreasing since

$$w'_d(y) = -\frac{(p\delta_d + h_0)f(y)}{\tilde{\gamma}(1 - (1 - \delta_d)F(y))^2} \leq 0.$$

It follows from the decreasing property of $w_d(\cdot)$ that $\frac{-w'_d(y)}{w_d(y)} \geq 0$. Finally,

$$\begin{aligned}\frac{-w'_d(y)}{w_d(y)} &= \frac{(p\delta_d + h_0)f(y)}{(1 - (1 - \delta)F(y))(p - (p + h_0)F(y))}, \\ &= \frac{(\delta_d + h_0/p)f(y)}{(1 - (1 - \delta_d)F(y))(1 - (1 + h_0/p)F(y))}.\end{aligned}$$

Since $\delta_d = \bar{\alpha} - \beta_g/\tilde{\gamma} \in [0, 1]$, $1 - (1 - \delta_d)F(y)$ is positive and decreasing. Therefore, it suffices to show that $\frac{f(y)}{1 - (1 + h_0/p)F(y)} = \frac{f(y)}{1 - F(y)} \left(1 + \frac{F(y)h_0/p}{1 - (1 + h_0/p)F(y)}\right)$ is positive and increasing: both properties have been proved in the original proof of Theorem 2.1.

It is left to show that $g_d(0) > 0$ and $\lim_{y \uparrow y_{max}} g_d(y) < 0$. $g_d(0) = 1 - \frac{\tilde{\gamma}c}{p} > 0$ since $F(0) = 0$, $w(0) = p/\tilde{\gamma}$, $\xi(0) = 0$ and $\xi'(0) = 1$. If $y_{max} = y_s$, (see (2.6) for the definition of y_s), by (2.6) and $\xi'(y_{max}) = 0$, $\Pi'_s(y_{max}) = \xi(y_{max})w'_d(y_{max}) - c(1 - F(y_{max})) < 0$ since $w'_d(\cdot) \leq 0$ and $\xi(\cdot) \geq 0$. This implies that $\lim_{y \uparrow y_{max}} g_d(y) < 0$. Finally if $y_{max} = F^{-1}\left(\frac{1}{1 + h_0/p}\right)$, then the profit value $\Pi_s(y_{max}|\beta_g)$ itself is negative by (2.5) as $w_d(y_{max}) = 0$. This, however, implies that $\Pi_s(y|\beta_g)$ is decreasing at some point $y^0 < y_{max}$, so that $g_d(y^0) < 0$. By the monotonicity of $g_d(\cdot)$, $\lim_{y \uparrow y_{max}} g_d(y) < 0$ in this case as well.

Part (b): $\lim_{w \uparrow p/\tilde{\gamma}} \hat{\Pi}_s(w|\beta_g) = 0$ since $\lim_{w \uparrow p/\tilde{\gamma}} y(w, \beta_g) = 0$. The remainder of part (b) is immediate from part (a).

The Stackelberg Game under Given Wholesale Price

There exists a one-to-one mapping between the ECCR β_g and the resulting base-stock level y selected by the retailer. Indeed, it is apparent from (A.38) that β_g can be written as a closed-form function of y (the immediate generalization of (2.8)):

$$\beta_g(w, y) \equiv \tilde{\gamma}\bar{\alpha} - \left(\frac{p}{w} - \tilde{\gamma}\right) \left(\frac{1}{F(y)} - 1\right) + \frac{h_0}{w} \quad (\text{A.42})$$

Since the ECCR must be selected in the interval $[\underline{\alpha}, \tilde{\gamma}\bar{\alpha}]$, this implies that the targeted base-stock level y satisfies the bounds, $\underline{y} \leq y \leq \bar{y}$ where

$$\underline{y} \equiv F^{-1} \left(\frac{p - \tilde{\gamma}w}{p - \tilde{\gamma}w + w(\tilde{\gamma}\bar{\alpha} - \underline{\alpha}) + h_0} \right) \text{ and } \bar{y} \equiv F^{-1} \left(\frac{p - \tilde{\gamma}w}{p - \tilde{\gamma}w + h_0} \right). \quad (\text{A.43})$$

Substituting (A.42) into (A.37), we obtain the desired representation of the supplier's equilibrium profits as a function of the targeted base-stock level:

$$\Pi_s(y|w) = (w - c)s(y) - \left(w\tilde{\gamma}\bar{\alpha} - (p - \tilde{\gamma}w) \left(\frac{1}{F(y)} - 1 \right) + h_0 \right) h(y). \quad (\text{A.44})$$

The supplier's problem is $\max_{\underline{y} \leq y \leq \bar{y}} \Pi_s(y|w)$. The marginal equilibrium profit function for the supplier is given by

$$\begin{aligned} \Pi'_s(y|w) &= (w - c)(1 - F(y)) - [w\tilde{\gamma}\bar{\alpha} + h_0 - (p - \tilde{\gamma}w) \left(\frac{1}{F(y)} - 1 \right)] F(y) - (p - \tilde{\gamma}w) \frac{h(y)f(y)}{F^2(y)} \\ &= (p - c)(1 - F(y)) - (w\tilde{\gamma}\bar{\alpha} + h_0)F(y) - (p - \tilde{\gamma}w) \frac{h(y)f(y)}{F^2(y)} \\ &\quad + w(1 - \tilde{\gamma})(1 - F(y)) \end{aligned} \quad (\text{A.45})$$

Proof of Theorem 2.2 (with $\bar{\alpha}_d$ replacing $\bar{\alpha}$ in the statement of the theorem.)

Parts (b) and (c) are immediate from part (a).

Part (a): Using (A.45), write

$$\Pi'_s(y|w) = (1 - F(y))n(y) - (w\tilde{\gamma}\bar{\alpha} + h_0), \quad (\text{A.46})$$

where $n(y) = (p - c + w\tilde{\gamma}\bar{\alpha} + h_0) - (p - \tilde{\gamma}w)\zeta(y) + w(1 - \tilde{\gamma})$. The remainder of the proof is the same as in the base model: Note that $n(y)$ is decreasing in view of the (MIFR) property. Thus define $y_w \equiv \inf\{L \leq y : n(y) \leq 0\}$, unless $n(\cdot) > 0$ for

all $y \geq L$, in which case we define $y_w \equiv \infty$. Note that, in either case, $n(\cdot) \geq 0$ for all $L \leq y < y_w$. Thus, for all $L \leq y < y_w$, $\Pi'_s(\cdot|w)$ is a decreasing function of y as a translation of the product of two positive and decreasing functions. In other words, $\Pi_s(\cdot|w)$ is concave on $[L, y_w]$. Moreover, $n(y) \leq 0$ and $\Pi'_s(\cdot|w) < 0$ for $y > y_w$, i.e., $\Pi_s(\cdot|w)$ is decreasing for $y > y_w$. This implies that the function $\Pi_s(\cdot|w)$ is quasi-concave in $[L, \infty)$, achieving its maximum for some $L \leq y^0 < y_w$. \square

Comparative Statics and Full Stackelberg Game

We start with a characterization of the impact the wholesale price has on the *unconstrained* optimal base-stock level $y_0(w) = \operatorname{argmax}_{y \geq L} \Pi_s(y|w)$. Rewrite (A.45) as:

$$\begin{aligned} \Pi'_s(y|w) &= (p-c)(1-F(y)) + \bar{\alpha}(p-w\tilde{\gamma})F(y) - (p\bar{\alpha}+h_0)F(y) - (p-\tilde{\gamma}w)\frac{h(y)f(y)}{F^2(y)} \\ &\quad - \frac{1-\tilde{\gamma}}{\tilde{\gamma}}(p-\tilde{\gamma}w)(1-F(y)) + \frac{1-\tilde{\gamma}}{\tilde{\gamma}}p(1-F(y)) \\ &= \left(\frac{p}{\tilde{\gamma}}-c\right) - \left(\frac{p}{\tilde{\gamma}}-c+p\bar{\alpha}+h_0\right)F(y) \\ &\quad + (p-\tilde{\gamma}w)F(y)(\bar{\alpha}-\kappa_d(y)) \end{aligned} \tag{A.47}$$

where $\kappa_d(y) \equiv \frac{h(y)f(y)}{F^3(y)} + \frac{1-\tilde{\gamma}}{\tilde{\gamma}}\frac{1-F(y)}{F(y)}$. In general, $y_0(w)$ cannot be obtained in closed form except when the wholesale price $w = p_d$. For this maximal wholesale price value, it is easily verified from (A.47) that $y_0(p/\tilde{\gamma}) = y^p$ where

$$y_d^p \equiv \max\left\{L, F^{-1}\left(\frac{p/\tilde{\gamma}-c}{p/\tilde{\gamma}-c+p\bar{\alpha}+h_0}\right)\right\}. \tag{A.48}$$

Proof of Theorem 2.3 (with $\kappa_d(\cdot)$, y_d^p and p_d replacing their counterparts $\kappa(\cdot)$, y^p and p in the statement of the theorem)

It can be easily verified that the proof for the base model applies here as well.

Proof of Proposition 2.1 (with p_d replacing p in the statement of the proposition)

Since $\kappa_d(y) = \kappa(y) + \frac{1-\tilde{\gamma}}{\tilde{\gamma}}\frac{1-F(y)}{F(y)} \geq \kappa(y)$ and $\frac{1-F(y)}{F(y)}$ is increasing in y , the properties listed in (A.10) also hold for $\kappa_d(\cdot)$, i.e.,

$$\kappa_d(\cdot) \text{ is strictly decreasing for } y \geq L, \kappa_d(M) > 1 - c/p \text{ and } \kappa_d(M) > \bar{\alpha} \tag{A.49}$$

where M denotes the median of the distribution. Since

Assume to the contrary that y_0 is decreasing in w . Then by (A.9) and also from the proof of Theorem 2.3:

$$\kappa_d(y_0) \leq \bar{\alpha}. \quad (\text{A.50})$$

From (A.45), rewrite $0 = \Pi'_s(y|w)$ as follows:

$$\Pi'_s(y_0|w) = (p-c)(1-F(y_0)) - pF(y_0)\kappa(y_0) - h_0F(y_0) - \tilde{\gamma}wF(y_0)(\bar{\alpha} - \kappa_d(y_0)) \quad (\text{A.51})$$

Thus $(p-c)(1-F(y_0)) - pF(y_0)\kappa(y_0) \geq 0$, i.e.,

$$\zeta(y_0) = \kappa(y_0) \frac{F(y_0)}{1-F(y_0)} \leq 1 - c/p < \kappa(M) = \zeta(M). \quad (\text{A.52})$$

Since $\zeta(\cdot)$ is increasing, $y_0 < M$. Since $\kappa_d(\cdot)$ is strictly decreasing, $\kappa_d(y_0) > \kappa_d(M) > \bar{\alpha}$, contradicting (A.50) and hence the assumption that $y_0(w)$ is a decreasing. \square

Theorem 2.4 (with p_d and $\bar{\alpha}_d$ replacing p and $\bar{\alpha}$ in the statement of parts (b1) and (c))

(a1) The same proof applies to this model with default risk.

(a2) We distinguish between two cases:

Case 1: $\bar{w} = p/\tilde{\gamma} = p_d$

All the arguments in the original proof still apply up to (A.16), which is used to show $\Pi_s^{u''}(\cdot) > 0$). These arguments simply rely on the results on Theorems 2.2 and 2.3 as well as Proposition 2.1. The following completes the proof for this case:

Employing (A.44) for any $w \leq w_2$, we get

$$\begin{aligned} \Pi_s^{u'}(w) &= \frac{\partial \Pi_s(y_0(w)|w)}{\partial w} \\ &= s(y_0(w)) - \tilde{\gamma} \left(\bar{\alpha} + \frac{1}{F(y_0(w))} - 1 \right) h(y_0(w)) \end{aligned} \quad (\text{A.53})$$

$$= y_0(w) - \tilde{\gamma} \left(\bar{\alpha} + \frac{1}{F(y_0(w))} \right) h(y_0(w)) - \gamma h(y_0(w)). \quad (\text{A.54})$$

Any point w in a type-2 interval has

$$y_0(w) \geq \underline{y}(w) > \underline{y}(\bar{w}) = L \quad (\text{A.55})$$

where the second inequality follows from the fact that $\underline{y}(\cdot)$ is strictly decreasing and $w \leq w_2 < \bar{w} = p$. The proof of Theorem 2.3(b) shows that since $y_0(w) \geq L$, $y_0(w)$

is continuously differentiable at w . Differentiating both sides of (A.54), we get

$$\begin{aligned}\Pi_s^{u''}(w) &= \left\{ 1 - \tilde{\gamma} \left(\bar{\alpha}F(y_0(w)) + 1 - \frac{f(y_0(w))h(y_0(w))}{F^2(y_0(w))} \right) - \gamma F(y_0(w)) \right\} y_0'(w) \\ &= \left\{ \gamma(1 - F(y_0(w))) - \tilde{\gamma} \left(\bar{\alpha}F(y_0(w)) - \frac{f(y_0(w))h(y_0(w))}{F^2(y_0(w))} \right) \right\} y_0'(w) \\ &= \tilde{\gamma}F(y_0(w))[-\bar{\alpha} + \kappa_d(y_0(w))]y_0'(w) > 0\end{aligned}$$

where the inequality follows from the fact that, under patterns (b-i), (b-iii) and (b-iv), $y_0'(w) \neq 0$ always has the same sign as $-\bar{\alpha} + \kappa_d(y_0(w))$ whenever $y_0(w) > L$, see Theorem 2.3(b) and its proof.

Case 2: $\bar{w} < p/\tilde{\gamma} = p_d$

The proof for Case 2 is identical to that in the base model.

(b)-(c) Only the proof of (b1-2) needs to be adjusted, as follows:

It follows from the proof of part (a2) that $\beta_g^*(w) = \beta_g(w, \underline{y}(w)) = \underline{\alpha}$ when $w \leq w_1$ and $\beta_g^*(w) = \beta_g(\bar{y}(w)) = \bar{\alpha}$ when $w \geq w_2$. For $w_1 < w < w_2$, $\beta_g^*(w) < \bar{\alpha}$. Thus, by (A.42), $\tilde{\gamma}\bar{\alpha} - \beta_g^*(w) = \tilde{\gamma}\bar{\alpha} - \beta_g(w, y_0(w)) = \frac{(p-\tilde{\gamma}w)(1-F(y_0(w)))-h_0F(y_0(w))}{wF(y_0(w))} > 0$. Moreover, since $y_0(\cdot)$ is an increasing function, $\tilde{\gamma}\bar{\alpha} - \beta_g^*(w)$ is decreasing as the ratio of a positive decreasing function and a positive increasing function. Hence $\beta_g^*(w)$ is increasing for all $w_1 < w < w_2$.

Proposition 2.2

It follows from Proposition 2.1 and Theorem 2.4(b1) that two threshold points $0 < w_1 \leq w_2 \leq \bar{w}$ exist such that (i) $[0, w_1]$ is of type-1, (ii) $(w_1, w_2]$ is of type-2 and (iii) $(w_2, \bar{w}]$ is of type-3. We first show that under the specified parameter conditions, $w_1 < w_2$, i.e., a complete type-2 interval exists. For the classes of uniform and exponential distributions, this follows from the proof of Theorem 2.4(a2) since $\bar{w} = p_d$. For the Normals, it suffices to show that $\Pi'_s(L|\bar{w}) \geq 0$ where $\bar{w} = \frac{p(1-F(L))-F(L)h_0}{\tilde{\gamma}(1-F(L))+F(L)(\tilde{\gamma}\bar{\alpha}-\underline{\alpha})}$, (see the proof for the base model in Appendix I).

$$\begin{aligned}
\Pi'_s(L|\bar{w}) &= (p-c)(1-F(L)) - (\bar{w}\tilde{\gamma}\bar{\alpha} + h_0)F(L) - (p - \tilde{\gamma}\bar{w})\frac{h(L)f(L)}{F^2(L)} \\
&\quad + \bar{w}(1 - \tilde{\gamma})(1 - F(L)) \\
&> (p-c)(1-F(L)) - (\bar{w}\tilde{\gamma}\bar{\alpha} + h_0)F(L) - (p - \tilde{\gamma}\bar{w})\frac{h(L)f(L)}{F^2(L)} \\
&> (p-c)(1-F(L)) - pF(L) - p\frac{F(L)}{1-F(L)}\frac{h(L)f(L)}{F^2(L)} \\
&= (p-c)(1-0.036) - 0.036p(1+0.906) \\
&= 0.895p - 0.964c \\
&= 0.895(p - 1.077c) \geq 0.
\end{aligned}$$

The first equality follows from (A.45) and the second inequality follows from $\bar{w}\tilde{\gamma}\bar{\alpha} + h_0 \leq p\tilde{\gamma} \max\{\alpha_r, \alpha_s\} + h_0 \leq p \max\{\alpha_r, \alpha_s\} + h_0 \leq p$ and $p - \tilde{\gamma}\bar{w} = \frac{(p(\tilde{\gamma}\bar{\alpha} - \alpha) + \tilde{\gamma}h_0)F(L)}{\tilde{\gamma}(1-F(L)) + F(L)(\tilde{\gamma}\bar{\alpha} - \alpha)} < \frac{\tilde{\gamma}(p(\bar{\alpha} - \alpha) + h_0)F(L)}{\tilde{\gamma}(1-F(L))} \leq \frac{pF(L)}{1-F(L)}$ (since $\bar{\alpha} - \alpha \leq \max\{\alpha_r, \alpha_s\}$); The second equality follows from $F(L) = 0.036$ and $\frac{h(L)f(L)}{F^2(L)(1-F(L))} = \frac{(l\Phi(l) + \phi(l))\phi(l)}{\Phi^2(l)(1-\Phi(l))} = 0.906$ with $l = \frac{L-\mu}{\sigma} = -1.8$; Finally the last inequality follows from $\frac{p-c}{c} \geq 7.7\%$.

Since, in all parts (a)-(c), the initial type-1 interval is followed by a full type-2 interval $(w_1, w_2]$, the conclusions in these parts follow by showing that $\Pi_s^{*'}(\cdot) > 0$ on this type-2 interval. On type-2 intervals $\Pi_s^*(\cdot) = \Pi_s^u(\cdot)$, and, by (A.53),

$$\Pi_s^{u'}(w) = \omega_d(y_0(w)) \quad (\text{A.56})$$

where $\omega_d(y) = s(y) - \tilde{\gamma} \left(\bar{\alpha} + \frac{1}{F(y)} - 1 \right) h(y)$ is decreasing in $\tilde{\gamma}$ since $\bar{\alpha} + \frac{1}{F(y)} - 1 \geq 0$, and strictly increasing in $y \in [L, y_0(w_1))$ since $\omega'(y) = 1 - \tilde{\gamma} \left[\left(\bar{\alpha} + \frac{1}{F(y)} \right) F(y) + \frac{h(y)f(y)}{F^2(y)} \right] + (1 - \tilde{\gamma})F(y) = \tilde{\gamma}F(y)(\kappa_d(y) - \bar{\alpha}) > \tilde{\gamma}F(y)(\kappa_d(y_0(w_1)) - \bar{\alpha}) \geq 0$. (The first inequality holds since $\kappa_d(\cdot)$ is strictly decreasing in $y \geq L$ for all three classes of distributions considered here, see (A.49); the second inequality follows from (A.9) since $y_0(\cdot)$ is increasing, see Proposition 2.1(c).) Therefore, since $y_0(w_1) > L$, see (A.17), we have, for any $w \in [w_1, w_2]$, that $\omega(L) < \omega_d(y_0(w_1)) = \Pi_s^{u'}(w_1) \leq \Pi_s^{u'}(w)$ where the second inequality follows from the convexity of $\Pi_s^{u'}(\cdot)$. The proof is complete since $\omega_d(L) \geq \omega(L) \geq 0$, where the first inequality follows from the fact that $\omega_d(y)$ is decreasing $\tilde{\gamma}$ and it coincides with $\omega(y)$ (see the proof of the base model) when

$\tilde{\gamma} = \frac{1-\gamma}{\phi(\gamma)}|_{\gamma=0} = 1$ and the second inequality can be found in the proof of the base model.

A.4 Comparisons under a Perfect Coordination Scheme

In this appendix, we consider the case where the chain members agree to implement a perfect coordination scheme consisting of a combined wholesale price w and ECCR β_g that maximizes the aggregate profits, thereafter splitting these in accordance with an agreed upon allocation rule such as a Nash bargaining solution. In this context, we assume that all wholesale prices $c \leq w \leq p$ and all nonnegative ECCR values of β_g are feasible.

In the general model, specified by (2.1) and (2.2), the aggregate profit function $\Pi_{agg}(w, y) = (p - c)s(y) - (\bar{\alpha}w + h_0)h(y)$, which is clearly optimized over $c \leq w \leq p$ by selecting $w = c$. The corresponding optimal base-stock level y_{agg} satisfies

$$F(y) = 1 - \frac{\bar{\alpha}c + h_0}{p - c + \bar{\alpha}c + h_0}. \quad (\text{A1})$$

The corresponding first-best aggregate profits Π_{agg}^* can be achieved in a decentralized supply chain with a (w, β_g) - payment scheme. However the only such scheme achieving perfect coordination has $w = c$ and $\beta_g = 0$. (As argued above, any choice of $w > c$ results in double marginalization and suboptimal aggregate profits; it follows from (2.3) that, for fixed w , the optimal base-stock level is strictly increasing in β_g , thus only $\beta_g = 0$ induces the choice of $y = y_{agg}$.)

Assume now that the chain members agree to split the chain profits according to a Nash bargaining solution where the supplier (retailer) has a bargaining power index γ_s (γ_r) and minimum profit expectation Π_s^0 (Π_r^0). This means that the retailer's and the supplier's expected profit value Π_r^* and Π_s^* are the unique optimal solution of the optimization problem

$$\begin{aligned} \max \quad & (\Pi_s - \Pi_s^0)^{\gamma_s} (\Pi_r - \Pi_r^0)^{\gamma_r} \\ \text{s.t.} \quad & \Pi_r + \Pi_s = \Pi_{agg}^*. \end{aligned}$$

Clearly, both Π_s^* and Π_r^* are increasing functions of Π_{agg}^* , i.e., the larger the total profit, the larger the share of each of the chain members. Since $\Pi_{agg}(w, y)$ is decreasing in $\bar{\alpha}$, the same monotonicity property applies to Π_{agg}^* . The following conclusions

are immediate, employing equation (A1) and the fact that $\bar{\alpha} = \alpha_s$ under TC while $\bar{\alpha} = \alpha_r$ under IS.

Corollary A.1 *If $\alpha_s \leq (>)\alpha_r$, the expected profit of the supplier and that of the retailer, the base-stock level and the expected sales volume are all larger under TC as opposed to IS.*

Appendix B

Proofs for Chapter 3

B.1 Proofs

Proof of Theorem 3.2: (a) Follows from Scarf (1960).

(b) Note that the total expected cost over the finite planning horizon may be written as:

$$\begin{aligned} & \sum_{n=1}^N \rho^n \{c(y_n - x_n) + G_n(y_n) + \Gamma_n(y_n)\} - \rho^{N+1} c \mathbb{E}(y_N - D[N, N+L]) \\ &= \sum_{n=1}^N \rho^n \{c(1 - \rho)y_n + G_n(y_n) + \Gamma_n(y_n)\} - \rho c x_1 + c \sum_{n=1}^{N+1} (\rho^n + \rho^{N+1}) \mathbb{E}D_1 \end{aligned}$$

Let $L_n(y) \equiv c(1 - \rho)y + G_n(y) + \Gamma(y)$, a convex function by Theorem 1 (a) and our assumption about the $\Gamma(\cdot)$ function. Let $S_n^* = \operatorname{argmin} L_n(y)$ defined as the smallest minimizer of the function. We will show that $S_1^* \leq S_2^* \leq \dots \leq S_N^*$. This implies the optimality of the following myopic policy in each period n : adopt the base-stock policy with level S_n^* . The ordering decisions prescribed by this policy optimize each of the N terms separately, and therefore the aggregate expression, as well. We provide the proof assuming inventory levels vary continuously; the case where the demand distribution, and hence the inventory levels, are discrete, can be handled in a similar way.

With continuous inventory levels, it suffices to show that $L'_n(\cdot) \geq L'_{n+1}(\cdot)$ for all n , which is equivalent to showing that $G'_n(y) = \mathbb{E} \sum_{j=1}^{N-n+1} \rho^{L+j} [\alpha(j) - \alpha(j -$

1)] $\text{Prob}(D[1, L + j] \leq y) \geq \mathbb{E} \sum_{j=1}^{N-n} \rho^{L+j} [\alpha(j) - \alpha(j-1)] \text{Prob}(D[1, L + j] \leq y) = G'_{n+1}(y)$, where the inequality follows from the fact that the right hand side equals the left hand side plus one extra positive term.

(c) See Federgruen and Zipkin(1986a, 1986b);

(d) See Scarf (1960), Iglehart(1963a, 1963b) and Veinott (1965).

(e) See Sethi and Cheng (1997).

(f) See Chen and Simchi-Levi (2004) and Federgruen and Heching (1999).

□

Proof of Theorem 3.4: (a) In this case, it is easily verified that Assumption (NIP) holds, without loss of optimality: backlogs must be cleared ultimately and there is no incentive to delay the clearance of any previously backlogged demand at the beginning of any period; thus $y_n \geq 0$, without loss of optimality. The optimality of time-dependent base-stock policy follows immediately from Lemma 3.2.

(b) Under the (CBL) assumption, the proof of part (b) is identical to that of part (a), with $G_n^-(\cdot)$ specified by (3.14) as opposed to (3.13). Under the (NIP) assumption, the non-negative constraints for the action variable y_n , may be handled as follows: first consider the relaxed problem where the constraint $y_n \geq 0$ is relaxed. By the equivalency result of Lemma 3.1, a time-dependent base-stock policy is optimal in that “equivalent” model. If $S_n^* \geq 0$ for all $n = 1, \dots, N$, this policy satisfies the relaxed constraint and is therefore optimal in the original problem as well. Otherwise, transform the one-step expected cost functions $\{G_n^-(\cdot)\}$ to functions $\{G_n^-(\cdot|M)\}$ defined as follows: $G_n^-(y|M) = G_n^-(y)$ for $y \geq 0$ and $G_n^-(y|M) = e^{-My} + G(0) - 1$ for $y < 0$. Note that the functions $\{G_n^-(\cdot|M)\}$ continue to be convex for M sufficiently large, so that a time-dependent base-stock policy continues to be optimal. Moreover, for M sufficiently large, say $M \geq M_1$, $S_n^*(M) \geq 0$ for all $n \geq 1$. Finally, $S_n^*(M) = S_n^*(M_1)$ for all $M \geq M_1$ and all $n = 1, \dots, N$, since the value of the total expected costs is independent of M as long as $S_n^* \geq 0$.

(c) The proof of part (c) is analogous to that of part (b). In the transformed model, a level-dependent (s, S) - policy is optimal in view of the structural results by Scarf (1960) and Iglehart(1963a, 1963b). □

Proof of Lemma 3.3: The finiteness of $\hat{G}(\cdot)$ follows from the fact that $H(t)$ and $J(t)$ are both $O(t^{l+1})$ while $\mathbb{E}X^{l+1} < \infty$. The family of distributions $\{A_y : y \geq 0\}$ is SIL (Stochastically Increasing Linear) in y , see Example 8.A.16 in Shaked and Shanthikumar (2007). Similarly, we can easily verify that $\{A_y : y \in \mathcal{Z}^-\}$ is SIL in y using Example 8.B.7 and Theorem 8.B.9 there. This implies that $\mathbb{E}H(A_y)$ is a convex function since $H(\cdot)$ is increasing and convex in view of the monotonicity property of $\bar{\alpha}(\cdot)$, and see Definition 8.A.1(e) in Shaked and Shanthikumar (2007). Similarly, $\mathbb{E}J(A_y) = -\mathbb{E}(-J(A_y))$ is convex in y since $\mathbb{E}(-J(A_y))$ is a concave function of y , as $-J(\cdot)$ is increasing and concave in view of the monotonicity property of $\bar{\beta}(s)$. \square

Proof of Lemma 3.4: The fact that $\Delta H_j < \infty$ follows, again, from the fact that $\alpha(t) = O(t^l)$ and $\mathbb{E}(X^{l+1}) < \infty$ while $\mathbb{E}X^{p+1} < \infty$. For $j = 1$, $\Delta H_1 = \mathbb{E}[H(R(L) \oplus X) - H(R(L))] = \mathbb{E}_{R(L)}\{\mathbb{E}_X[H(r+X) - H(r) | R(L) = r]\} \geq \mathbb{E}_{R(L)}\mathbb{E}_X[H(X) - H(0)] \geq \mathbb{E}_{R(L)}\mathbb{E}_X[H(\tilde{S}_1) - H(0)] = \Delta H_0$, where the first inequality follows from the convexity of the $H(\cdot)$ function, since $\alpha(\cdot)$ is increasing, and where the second inequality follows from the fact that $H(\cdot)$ is increasing and $X \geq_{st} R(L)$; for the latter, see Lemma 3.10 in Barlow and Proschan (1975).

For $j \geq 2$,

$$\begin{aligned}
\Delta H_j &= \mathbb{E}[H(\tilde{S}_{j-1} \oplus X_{j-1} \oplus X_j) - H(\tilde{S}_{j-1} \oplus X_{j-1})] \\
&= \mathbb{E}_{\tilde{S}_{j-1}} \mathbb{E}_{X_{j-1}} \mathbb{E}[H(s + x + X_j) - H(s + x) | \tilde{S}_{j-1} = s, X_{j-1} = x] \\
&\geq \mathbb{E}_{\tilde{S}_{j-1}} \mathbb{E}_{X_{j-1}} \mathbb{E}_{X_j}[H(s + X_j) - H(s) | \tilde{S}_{j-1} = s] \\
&= \mathbb{E}_{\tilde{S}_{j-1}} \mathbb{E}_{X_j}[H(s + X_j) - H(s) | \tilde{S}_{j-1} = s] \\
&= \mathbb{E}_{\tilde{S}_{j-1}} \mathbb{E}_{X_{j-1}} \mathbb{E}[H(s + X_{j-1}) - H(s) | \tilde{S}_{j-1} = s] \\
&= \mathbb{E}[H(\tilde{S}_{j-1} \oplus X_{j-1}) - H(\tilde{S}_{j-1})] \\
&= \mathbb{E}[H(\tilde{S}_j) - H(\tilde{S}_{j-1})] = \Delta H_{j-1}.
\end{aligned}$$

(The first inequality follows from the convexity of $H(\cdot)$ and $X_{j-1} \geq 0$ a.s., and the third equality follows from X_j and X_{j-1} being identically distributed.) \square

Proof of Theorem 3.7: (a) The proof that $G(\cdot)$ is finite is analogous to that in Theorem 3.6(a). To show its convexity, it suffices to show that for all $j \geq 1$,

$\Delta J_j \geq \Delta J_{j+1}$. We show, in fact, that *almost surely*, $\Delta J_j \geq \Delta J_{j+1}$, i.e., $J(S_{n+1} + L - S_{n+j}) - J([S_n + L - S_{n+j}]^+) \geq J(S_{n+1} + L - S_{n+j+1}) - J([S_n + L - S_{n+j+1}]^+)$.

Since the function $J(\cdot)$ is convex, it suffices to show:

- (i) $[S_n + L - S_{n+j}]^+ \geq [S_n + L - S_{n+j+1}]^+$,
- (ii) $S_{n+1} + L - S_{n+j} - [S_n + L - S_{n+j}]^+ \geq S_{n+1} + L - S_{n+j+1} - [S_n + L - S_{n+j+1}]^+ \iff X_{n+j+1} + [S_n + L - S_{n+j+1}]^+ \geq [S_n + L - S_{n+j}]^+$.

The inequality in (i) is immediate, while the second inequality in (ii) follows from $X_{n+j+1} + [S_n + L - S_{n+j+1}]^+ \geq [X_{n+j+1} + S_n + L - S_{n+j+1}]^+ = [S_n + L - S_{n+j}]^+$, since $X_{n+j+1} \geq 0$, a.s.

(b) The proof of part (b) is analogous to that of Theorem 3.6, using part (a). Because of Assumption 1, the equivalent PRM has the non-standard feasibility constraints $y_n \geq 0$ for its action space. However, these constraints may be relaxed as in the proof of Theorem 3(b). This implies that an (s^*, S^*) policy continues to be optimal. To be feasible, in terms of the constraints $y_n \geq 0$, $s^* \geq 0$ and the optimal policy is the best among all (s, S) -policies, with $s \geq 0$. \square

Appendix C

Proofs for Chapter 4

C.1 Proofs

Algorithm for Optimal r and q (see Federgruen and Zheng (1992))

Step 0 (Initialize): Determine y_0 . $q := 1$; $r := y_0 - 1$. $C^* := \lambda K + G(y_0)$.

Step 1 (Determine next smallest $G(\cdot)$ -value): $c^* := \min\{G(r), G(r + q - 1)\}$.

Step 2 (Test for termination): If $c^* \geq C^*$, $q^* := q$, $r^* := r$. Stop

Step 3 (Update): $q := q + 1$, $C^* := C^* - (C^* - c^*)/q$. If $G(r)$ achieves the minimum in Step 1, $r := r - 1$.

Go to Step 1.

Proof of Lemma 4.1. Since θ is irrelevant in this lemma, we omit it for the sake of brevity.

Note that $\partial c(r, q)/\partial r = (G(r + q) - G(r))/q$. By the strict quasi-convexity and continuous differentiability of $G(\cdot)$, it follows that the function $\partial c(r, q)/\partial r$ is continuous and increasing in r on the interval $[y_0 - q, y_0]$. Moreover, $\partial c(r, q)/\partial r < 0$ for all $r \leq y_0 - q$ and $\partial c(r, q)/\partial r > 0$ for all $r \geq y_0$. This implies that the function $\partial c(r, q)/\partial r$ has a unique root r' on $[y_0 - q, y_0]$ with $\partial c(r, q)/\partial r \leq 0$ for $r \leq r'$ and $\partial c(r, q)/\partial r > 0$ for $r > r'$. This establishes that the function $c(\cdot, q)$ is strictly quasi-convex and achieves its minimum in a point $r' = r(q) < y_0$ with $R(q) = r' + q > y_0$. Also, $r(q)$ is the unique root of the equation $\partial c(r, q)/\partial r = 0$, or $G(r) = G(r + q)$.

Thus (P0)-(P2) follow.

See parts (3)-(4) of Lemma 3 in Zheng (1992) for proofs of (P3)-(P4), observing that the proof only relies on the strict quasi-convexity and differentiability of $G(\cdot)$.

□

Proof of Theorem 4.1. Since θ is irrelevant in this theorem as well, we omit it from the notation.

As in Zheng (1992), define $H(q) \equiv G(r(q))$ and $H(0) \equiv \lim_{q \rightarrow 0^+} G(r(q)) = G(y_0)$. $H'(q) = G'(r(q))r'(q) > 0$ since $G'(r(q)) < 0$ by (P2) and $r'(q) < 0$ by (P3). $c(q) = \frac{\lambda K + \int_0^q H(y)dy}{q}$ as shown in Zheng (1992). Applying L' Hôpital's rule, we observe that $\lim_{q \rightarrow \infty} c(q) = \lim_{q \rightarrow \infty} H(q) = \lim_{q \rightarrow \infty} G(r(q)) = \infty$, by properties in (Q) and since $\lim_{q \rightarrow \infty} r(q) = -\infty$ by (P3). (Note that the numerator of $c(q)$ is strictly increasing and convex, hence tends to infinity when q goes to infinity so that the conditions for L' Hôpital's rule are satisfied.)

Moreover, $c'(q) = \frac{A(q) - \lambda K}{q^2}$ where $A(q) \equiv H(q)q - \int_0^q H(y)dy$ with $A(0) = 0$ and $A'(q) = H'(q)q > 0$. Thus, define q^0 as the unique root of the equation $A(q) = \lambda K$. (q^0 must exist, for, otherwise, $c'(q) < 0$ for all q , so that $q^* = \infty$. This, however, contradicts $\lim_{q \rightarrow \infty} c(q) = \infty$ as shown above.) Then $c'(q) \leq 0$ for $q \leq q^0$, which is equivalent to $H(q)q - \int_0^q H(y)dy - \lambda K < 0 \iff H(q) = G(r(q)) \leq c(q)$. Similarly, $c'(q) > 0$ for $q > q^0$, which is equivalent to $H(q) = G(r(q)) \geq c(q)$. This establishes both the strict quasi-convexity of $c(q)$ and (P5). □

Proof of Lemma 4.2. Since θ is irrelevant in this lemma, again, we omit it from the notation.

Note that $\Delta_r c(r, q) \equiv c(r, q|\theta) - c(r-1, q|\theta) = (G(r+q) - G(r))/q$. It follows from the quasi-convexity of $G(\cdot)$ that $\Delta_r c(r, q) < 0$ for all $r \leq y_0 - q$ and $\Delta_r c(r, q) > 0$ for all $r \geq y_0$. Let $r' \equiv \min\{r : \Delta_r c(r, q) \geq 0\} = \min\{r : G(r+q) \geq G(r)\}$. Thus

$$y_0 - q < r' \leq y_0. \quad (\text{C.1})$$

For any $r' < r < y_0$, $\Delta_r c(r, q) > 0$ since $G(r+q) > G(r'+q) \geq G(r') > G(r)$ by the strict quasi-convexity of $G(\cdot)$. Therefore, $\Delta_r c(r, q|\theta) \leq 0$ for $r \leq r'$ and $\Delta_r c(r, q|\theta) \geq 0$ for $r \geq r'$, which implies that $c(r, q)$ is quasi-convex in r achieving its minimum at $r' - 1$, establishing (P1'). (P2') follows immediately from (C.1).

Next we show that $r(q)$ is decreasing in q . Let $r_0 \equiv r(q_0)$. For any $q_1 > q_0$, $G(r_0 + 1 + q_1) > G(r_0 + 1 + q_0) \geq G(r_0 + 1)$ where the first inequality follows from (P2') and the strict quasi-convexity of $G(\cdot|\theta)$ while the second inequality follows from (P1') with $q = q_0$. Thus, by (P1'), $r(q_1) \leq r_0$. To show that $R(q)$ is increasing, assume to the contrary that $R(q_0) = r(q_0) + q_0 > R(q_1) = r(q_1) + q_1$ for some $q_0 < q_1$. Hence

$$r(q_1) < r^0 \equiv r(q_1) + q_1 - q_0 < r(q_0). \quad (\text{C.2})$$

Thus, $G(r^0 + 1) < G(r(q_1) + 1) \leq G(r(q_1) + q_1 + 1) = G(r^0 + q_0 + 1)$. (The first inequality follows from the strict quasi-convexity of $G(\cdot)$ and $r(q_1) + 1 < r^0 + 1 < r(q_0) + 1 \leq y_0$. The second inequality is by applying (P1') with $q = q_1$.) This implies that $r(q_0) \leq r^0$, which contradicts (C.2). The proof of the limits in (P3') and (P4') is analogous to that of Lemma 2, part (4) in Zheng (1992), using (P1'). (P5') follows from Lemma 2 and its proof in Federgruen and Zheng (1992). \square

Proof of Theorem 4.3. We write $r^*(\theta)$, $R^*(\theta)$ and $q^*(\theta)$ as r^* , R^* and q^* , without ambiguity. It follows from the implicit function theorem applied to (4.5)

and (4.6) that $\mathbf{H} \begin{bmatrix} dr \\ dq \end{bmatrix} + \mathbf{b}d\theta = 0$ where

$$\mathbf{H} = \begin{bmatrix} g(r|\theta) - g(R|\theta) & -g(R|\theta) \\ Rg(R|\theta) - rg(r|\theta) & Rg(R|\theta) \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} \frac{\partial G(r|\theta)}{\partial \theta} - \frac{\partial G(R|\theta)}{\partial \theta} \\ \int_r^R yg_\theta(y|\theta) dy \end{bmatrix} = \begin{bmatrix} -\int_r^R g_\theta(y|\theta) dy \\ \int_r^R yg_\theta(y|\theta) dy \end{bmatrix}.$$

Thus $\begin{bmatrix} dr^*/d\theta \\ dq^*/d\theta \end{bmatrix} = -\mathbf{H}^{-1}\mathbf{b}$. We have

$$\begin{aligned} \frac{dq^*}{d\theta} &= \frac{(g(R^*|\theta) - g(r^*|\theta)) \int_{r^*}^{R^*} yg_\theta(y|\theta) dy - (R^*g(R^*|\theta) - r^*g(r^*|\theta)) \int_{r^*}^{R^*} g_\theta(y|\theta) dy}{|\mathbf{H}|} \\ &= \frac{\int_{r^*}^{R^*} g_y(y|\theta) dy \int_{r^*}^{R^*} yg_\theta(y|\theta) dy - \int_{r^*}^{R^*} yg_y(y|\theta) dy \int_{r^*}^{R^*} g_\theta(y|\theta) dy}{|\mathbf{H}|} \end{aligned} \quad (\text{C.3})$$

since $g(R^*|\theta) - g(r^*|\theta) = \int_{r^*}^{R^*} g_y(y|\theta) dy$ and $R^*g(R^*|\theta) - r^*g(r^*|\theta) = \int_{r^*}^{R^*} yg_y(y|\theta) dy + \int_{r^*}^{R^*} g(y|\theta) dy = \int_{r^*}^{R^*} yg_y(y|\theta) dy + G(R^*|\theta) - G(r^*|\theta) = \int_{r^*}^{R^*} yg_y(y|\theta) dy$ where the last equation follows from (4.5). $|\mathbf{H}| = q^*g(r^*|\theta)g(R^*|\theta) < 0$ as $g(R^*|\theta) > 0$ and $g(r^*|\theta) < 0$, by (P2) in Lemma 4.1 and the strict quasi-convexity of $G(\cdot|\theta)$. The theorem thus follows from (C.3). \square

Proof of Theorem 4.5. We provide proof for the case where $g_y(y|\theta_1)/g_\theta(y|\theta_0)$ is decreasing. The other case can be proved analogously.

Recall from Federgruen and Zheng (1992), see also Zipkin (2000), that

$$r^*(\theta) = r, q^*(\theta) = q \iff G(j) < (\geq) c(r, q|\theta) \text{ for any } j \in (\neq)\{r+1, \dots, r+q\}. \quad (\text{C.4})$$

Suppose $q^*(\theta)$ increases at $\theta = \theta_1$. As explained above, the increase is by one unit, say from q^0 to $q^0 + 1$. Since $g_\theta(y|\theta)$ is uniformly positive, i.e., $G(y|\theta)$ is supermodular, both $r^*(\theta)$ and $R^*(\theta)$ are decreasing by Theorem 4.2. This implies that $r^*(\theta)$ decreases by 1 unit at θ_1 , say from r^0 to $r^0 - 1$. Let $\Delta(\theta) \equiv c(r^0, q^0|\theta) - c(r^0 - 1, q^0 + 1|\theta)$ and note that $\lim_{\theta \uparrow \theta_1} \Delta(\theta) \leq 0$ and $\lim_{\theta \downarrow \theta_1} \Delta(\theta) \geq 0$. Since $\Delta(\theta)$ is a continuous function of θ , $\Delta(\theta_1) = 0$, i.e.,

$$c(r^0 - 1, q^0 + 1|\theta_1) = c(r^0, q^0|\theta_1). \quad (\text{C.5})$$

(The continuity of $\Delta(\cdot)$ follows from (4.2) and the continuity in θ of $G(y|\theta)$.) Hence

$$G(r^0 + q^0|\theta_1) < c(r^0, q^0|\theta_1) = c(r^0 - 1, q^0 + 1|\theta_1) = G(r^0|\theta_1) < G(r^0 - 1|\theta_1). \quad (\text{C.6})$$

The first inequality of (C.6) follows from the fact that (r^0, q^0) is the optimal policy for $\theta = \theta_1$ and (C.4), while the second inequality follows from the strict quasi-convexity of $G(\cdot|\theta)$. The first equality follows from (C.5), while the second equality follows from (C.5) and

$$c(r^0 - 1, q^0 + 1|\theta_1) = \frac{K\lambda + \sum_{r^0+1}^{r^0+q^0} G(j|\theta_1) + G(r^0|\theta_1)}{q^0 + 1} = \frac{q^0 c(r^0, q^0|\theta_1) + G(r^0|\theta_1)}{q^0 + 1}.$$

Let $\theta_2 \equiv \inf\{\theta \geq \theta_1 : G(r^0 + q^0|\theta) \geq c(r^0 - 1, q^0 + 1|\theta)\}$. We will show that

$$c(r^0 - 1, q^0 + 1|\theta) < G(r^0 - 1|\theta) \text{ for all } \theta_1 \leq \theta \leq \theta_2. \quad (\text{C.7})$$

This will allow us to conclude that the first change of $q^*(\theta)$ for $\theta > \theta_1$, if any, involves a decrease, thus proving rough monotonicity of $q^*(\cdot)$. Therefore, let θ_3 denote the first value of $\theta > \theta_1$ at which $q^*(\theta)$ changes from its current value $q^0 + 1$, if any. If $\theta_3 < \theta_2$, assume by contradiction that q^* increases at $\theta = \theta_3$. Then, analogous

to (C.6), we get $c(r^0 - 1, q^0 + 1|\theta_3) = G(r^0 - 1|\theta_3)$, contradicting (C.7). Thus, if $\theta_3 < \theta_2$, q^* decreases at θ_3 .

The remaining case has $\infty > \theta_3 \geq \theta_2$. So, r^* , R^* and q^* remain constant between θ_1 and θ_2 , while by the definition of θ_2 and the continuity in θ of the functions $G(r^0 + q^0|\theta)$ and $c(r^0 - 1, q^0 + 1|\theta)$,

$$G(r^0 + q^0|\theta_2) = c(r^0 - 1, q^0 + 1|\theta_2). \quad (\text{C.8})$$

Note first that the parameter pair $(r^0 - 1, q^0 + 1)$ continues to be optimal at θ_2 : for any pair (r, q) , $c(r^0 - 1, q^0 + 1|\theta) \leq c(r, q|\theta)$ for all $\theta_1 \leq \theta < \theta_2$. This inequality continues to hold when letting $\theta \uparrow \theta_2$ on both sides and using the continuity of the functions on both sides of the inequality. Thus $c(r^0 - 1, q^0 + 1|\theta_2) = \min_{(r,q)} c(r, q|\theta)$. By (C.8) and

$$\begin{aligned} c(r^0 - 1, q^0|\theta_2) &= \frac{K\lambda + \sum_{r^0}^{r^0+q^0} G(j|\theta_2) - G(r^0 + q^0|\theta_2)}{q^0} \\ &= \frac{(q^0 + 1)c(r^0 - 1, q^0 + 1|\theta_2) - G(r^0 + q^0|\theta_2)}{q^0}, \end{aligned}$$

we have $c(r^0 - 1, q^0|\theta_2) = c(r^0 - 1, q^0 + 1|\theta_2) = \min_{(r,q)} c(r, q|\theta)$, which proves that $q^*(\theta_2) \leq q^0$. We conclude that, irrespective of whether $\theta_3 < \theta_2$ or $\infty \geq \theta_3 \geq \theta_2$, if a change in $q^*(\theta)$ occurs for some $\theta > \theta_1$, the first change is a decrease, thus proving the roughly decreasing property of $q^*(\cdot)$.

The remainder is to prove (C.7). Fix $\theta = \theta'$ with $\theta_1 \leq \theta' \leq \theta_2$. (C.7) is equivalent to

$$\begin{aligned} 0 &< G(r^0 - 1|\theta') - c(r^0 - 1, q^0 + 1|\theta') \\ &= [G(r^0 - 1|\theta_1) - c(r^0 - 1, q^0 + 1|\theta_1)] \\ &\quad + \int_{\theta_1}^{\theta'} (G'_\theta(r - 1|\theta) - c'_\theta(r - 1, q + 1|\theta))d\theta \end{aligned} \quad (\text{C.9})$$

Note that

$$\begin{aligned}
a_1 &\equiv c(r^0 - 1, q^0 + 1|\theta_1) - G(r^0 + q^0|\theta_1) \\
&= \int_{\theta_1}^{\theta'} (G'_\theta(r^0 + q^0|\theta) - c'_\theta(r^0 - 1, q^0 + 1|\theta))d\theta + c(r^0, q^0|\theta') - G(r^0 + q^0|\theta') \\
&\geq \int_{\theta_1}^{\theta'} (G'_\theta(r^0 + q^0|\theta) - c'_\theta(r^0 - 1, q^0 + 1|\theta))d\theta \equiv a_2 \geq 0
\end{aligned}$$

where the first inequality holds because $c(r^0, q^0|\theta') - G(r^0 + q^0|\theta') \geq 0$ for all $\theta_1 \leq \theta' \leq \theta_2$ by the definition of θ_2 , and the second inequality holds because $G'_\theta(r^0 + q^0|\theta) - c'_\theta(r^0 - 1, q^0 + 1|\theta) = \sum_r^{r+q} (G'_\theta(r^0 + q^0|\theta) - G'_\theta(y|\theta))/(q + 1) \geq 0$ for any θ by the supermodularity of $G(y|\theta)$. Hence,

$$\begin{aligned}
\text{(C.9)} \quad &\iff b_1 \equiv G(r^0 - 1|\theta_1) - c(r^0 - 1, q^0 + 1|\theta_1) \\
&> \int_{\theta_1}^{\theta'} (c'_\theta(r^0 - 1, q^0 + 1|\theta) - G'_\theta(r^0 - 1|\theta))d\theta \equiv b_2 \\
&\iff a_1 b_2 < a_2 b_1 \iff (a_1 + b_1)b_2 = a_1 b_2 + b_1 b_2 < a_2 b_1 + b_1 b_2 = b_1(a_2 + b_2) \\
&\iff [G(r^0 - 1|\theta_1) - G(r^0 + q^0|\theta_1)] \int_{\theta_1}^{\theta'} (c'_\theta(r^0 - 1, q^0 + 1|\theta) - G'_\theta(r^0 - 1|\theta))d\theta \\
&< [G(r^0 - 1|\theta_1) - c(r^0 - 1, q^0 + 1|\theta_1)] \int_{\theta_1}^{\theta'} (G'_\theta(r^0 + q^0|\theta) - G'_\theta(r^0 - 1|\theta))d\theta \\
&\iff [G(r^0 - 1|\theta_1) - G(r^0 + q^0|\theta_1)][c'_\theta(r^0 - 1, q^0 + 1|\theta_0) - G'_\theta(r^0 - 1|\theta_0)] \\
&< [G(r^0 - 1|\theta_1) - c(r^0 - 1, q^0 + 1|\theta_1)][G'_\theta(r^0 + q^0|\theta_0) - G'_\theta(r^0 - 1|\theta_0)] \\
&\quad \text{for all } \theta_1 \leq \theta_0 \leq \theta' \tag{C.10}
\end{aligned}$$

Thus, to prove (C.7), it suffices to show (C.10). The first factor of the expression to the left side of (C.10) can be written as:

$$G(r^0 - 1|\theta_1) - G(r^0 + q^0|\theta_1) = - \sum_{t=r^0-1}^{r^0+q^0-1} g(t|\theta_1). \tag{C.11}$$

The second factor to the left side of (C.10) can be written as

$$\begin{aligned}
& c'_\theta(r^0 - 1, q^0 + 1|\theta) - G'_\theta(r^0 - 1|\theta) = \sum_{y=r^0}^{r^0+q^0} [G'_\theta(y|\theta) - G'_\theta(r^0 - 1|\theta)]/(q^0 + 1) \\
& = \sum_{y=r^0}^{r^0+q^0} \sum_{s=r^0-1}^{y-1} g_\theta(s|\theta)/(q^0 + 1) = \sum_{s=r^0-1}^{r^0+q^0-1} \sum_{y=s+1}^{r^0+q^0} g_\theta(s|\theta)/(q^0 + 1) \\
& = \sum_{s=r^0-1}^{r^0+q^0-1} (r^0 + q^0 - s)g_\theta(s|\theta)/(q^0 + 1). \tag{C.12}
\end{aligned}$$

Similarly, the first and second factors to the right side of (C.10) can be written, respectively, as:

$$G(r^0 - 1|\theta_1) - c(r^0 - 1, q^0 + 1|\theta_1) = G(r^0 - 1|\theta_1) - G(r^0|\theta_1) = -g(r^0 - 1|\theta_1), \tag{C.13}$$

where the first equality follows from $c(r^0 - 1, q^0 + 1|\theta_1) = G(r^0|\theta_1)$, see (C.6);

$$G'_\theta(r^0 + q^0|\theta) - G'_\theta(r^0 - 1|\theta) = \sum_{s=r^0-1}^{r^0+q^0-1} g_\theta(s|\theta). \tag{C.14}$$

Substituting (C.11)-(C.14) into (C.10), multiplying both sides with $q^0 + 1$ and defining $R^0 \equiv r^0 + q^0$, we get

$$\begin{aligned}
& \text{(C.10)} \\
& \iff - \sum_{t=r^0-1}^{R^0-1} g(t|\theta_1) \sum_{s=r^0-1}^{R^0-1} (R^0 - s)g_\theta(s|\theta_0) < -(q^0 + 1)g(r^0 - 1|\theta_1) \sum_{s=r^0-1}^{R^0-1} g_\theta(s|\theta_0) \\
& \iff - \sum_{t=r^0-1}^{R^0-1} g(t|\theta_1) \sum_{s=r^0-1}^{R^0-1} sg_\theta(s|\theta_0) \\
& > \left[(q^0 + 1)g(r^0 - 1|\theta_1) - R^0 \sum_{t=r^0-1}^{R^0-1} g(t|\theta_1) \right] \sum_{s=r^0-1}^{R^0-1} g_\theta(s|\theta_0)
\end{aligned}$$

Adding $\sum_{t=r^0}^{R^0} g(t|\theta_1) \sum_{s=r^0-1}^{R^0-1} sg_\theta(s|\theta_0)$ to both sides of the inequality above, we

obtain that (C.10) is equivalent to

$$\begin{aligned}
& [g(R^0|\theta_1) - g(r^0 - 1|\theta_1)] \sum_{s=r^0-1}^{R^0-1} sg_\theta(s|\theta_0) = \sum_{t=r^0-1}^{R^0-1} g_y(t|\theta_1) \sum_{s=r^0-1}^{R^0-1} sg_\theta(s|\theta_0) \equiv A \\
& > \sum_{t=r^0}^{R^0} g(t|\theta_1) \sum_{s=r^0-1}^{R^0-1} sg_\theta(s|\theta_0) \\
& + \left[(q^0 + 1)g(r^0 - 1|\theta_1) - R^0 \sum_{t=r^0-1}^{R^0-1} g(t|\theta_1) \right] \sum_{s=r^0-1}^{R^0-1} g_\theta(s|\theta_0) \equiv B. \tag{C.15}
\end{aligned}$$

To show (C.15), it is sufficient to prove the following:

$$A = \sum_{t=r^0-1}^{R^0-1} g_y(t|\theta_1) \sum_{s=r^0-1}^{R^0-1} sg_\theta(s|\theta_0) > \sum_{t=r^0-1}^{R^0-1} tg_y(t|\theta_1) \sum_{s=r^0-1}^{R^0-1} g_\theta(s|\theta_0) \equiv B', \tag{C.16}$$

since $B' \geq B$ by Lemma C.1 below. However, (C.16) holds by applying Lemma 4.3 (the discrete version) as $\frac{g_y(t|\theta_1)}{g_\theta(t|\theta_0)}$ is decreasing in t and is not constant.

Lemma C.1 $B' \geq B$.

Proof: Note first that, for any $a \leq b$,

$$\begin{aligned}
\sum_{t=a}^b tg_y(t|\theta_1) &= \sum_{t=a}^b t(g(t+1|\theta_1) - g(t|\theta_1)) = \sum_{t=a}^b tg(t+1|\theta_1) - \sum_{t=a}^b tg(t|\theta_1) \\
&= \sum_{t=a+1}^{b+1} (t-1)g(t|\theta_1) - \sum_{t=a}^b tg(t|\theta_1) \\
&= \sum_{t=a+1}^{b+1} tg(t|\theta_1) - \sum_{t=a+1}^{b+1} g(t|\theta_1) - \sum_{t=a}^b tg(t|\theta_1) \\
&= (b+1)g(b+1|\theta_1) - ag(a|\theta_1) - \sum_{t=a+1}^{b+1} g(t|\theta_1) \tag{C.17}
\end{aligned}$$

Hence we have $\sum_{t=r^0-1}^{R^0-1} tg_y(t|\theta_1) = R^0g(R^0|\theta_1) - (r^0-1)g(r^0-1|\theta_1) - \sum_{t=r^0}^{R^0} g(t|\theta_1)$.

Therefore,

$$\begin{aligned}
& B' - B \\
&= \left[R^0 g(R^0 | \theta_1) - (r^0 - 1)g(r^0 - 1 | \theta_1) - \sum_{t=r^0}^{R^0} g(t | \theta_1) \right] \sum_{s=r^0-1}^{R^0-1} g_\theta(s | \theta_0) \\
&\quad - \sum_{t=r^0}^{R^0} g(t | \theta_1) \sum_{s=r^0-1}^{R^0-1} s g_\theta(s | \theta_0) \\
&\quad - \left[(q^0 + 1)g(r^0 - 1 | \theta_1) - R^0 \sum_{t=r^0-1}^{R^0-1} g(t | \theta_1) \right] \sum_{s=r^0-1}^{R^0-1} g_\theta(s | \theta_0) \\
&= \left[R^0 g(R^0 | \theta_1) - R^0 g(r^0 - 1 | \theta_1) - \sum_{t=r^0}^{R^0} g(t | \theta_1) + R^0 \sum_{t=r^0-1}^{R^0-1} g(t | \theta_1) \right] \\
&\quad \times \sum_{s=r^0-1}^{R^0-1} g_\theta(s | \theta_0) - \sum_{t=r^0}^{R^0} g(t | \theta_1) \sum_{s=r^0-1}^{R^0-1} s g_\theta(s | \theta_0) \\
&= (R^0 - 1) \sum_{t=r^0}^{R^0} g(t | \theta_1) \sum_{s=r^0-1}^{R^0-1} g_\theta(s | \theta_0) - \sum_{t=r^0}^{R^0} g(t | \theta_1) \sum_{s=r^0-1}^{R^0-1} s g_\theta(s | \theta_0) \\
&= \sum_{t=r^0}^{R^0} g(t | \theta_1) \sum_{s=r^0-1}^{R^0-1} (R^0 - 1 - s) g_\theta(s | \theta_0) \\
&= [G(R^0 + 1 | \theta_1) - G(r^0 | \theta_1)] \sum_{s=r^0-1}^{R^0-1} (R^0 - 1 - s) g_\theta(s | \theta_0) \\
&\geq 0
\end{aligned}$$

where the inequality holds since $g_\theta(s | \theta_0) > 0$ by the strict supermodularity of $G(y | \theta)$ and $G(R^0 + 1 | \theta_1) - G(r^0 | \theta_1) = G(R^0 + 1 | \theta_1) - c(r^0, q^0 | \theta_1) \geq 0$ where the equality follows from (C.6) and the inequality follows from (C.4) and $(r^*(\theta_1), q^*(\theta_1)) = (r^0, q^0)$.

□

Proof of Corollary 4.3. Note that $g(y | \theta) = (h + b)F(y) - b$ and $g_y(y | \theta) = (h + b)f(y)$ where $f(y) \equiv F(y + 1) - F(y)$.

(a) Choose $\theta = h$. Thus $g_\theta(y | \theta) = F(y) > 0$ on $[d_l, \infty]$. It follows from (C.16) in the proof of Theorem 4.5 that the following is the sufficient condition for q^* to be roughly decreasing: for any $\theta_0 \equiv h_0 > h_1 \equiv \theta_1$ with $(r^0 - 1, R^0)$ being the optimal

policy at θ_1 ,

$$\sum_{t=r^0-1}^{R^0-1} (h_1 + b)f(t) \sum_{s=r^0-1}^{R^0-1} sF(s) > \sum_{t=r^0-1}^{R^0-1} t(h_1 + b)f(t) \sum_{s=r^0-1}^{R^0-1} F(s). \quad (\text{C.18})$$

It may be replaced by

$$\sum_{t=\max\{d_l, r^0-1\}}^{R^0-1} f(t) \sum_{s=\max\{d_l, r^0-1\}}^{R^0-1} sF(s) > \sum_{t=\max\{d_l, r^0-1\}}^{R^0-1} tf(t) \sum_{s=\max\{d_l, r^0-1\}}^{R^0-1} F(s). \quad (\text{C.19})$$

since all integrands are zero on $[r^0 - 1, \max\{d_l, r^0 - 1\})$. If $F(y)$ is log-concave on $[d_l, d_h]$, i.e., $f(y)/F(y)$ is decreasing on $[d_l, d_h]$, it is log-concave on $[d_l, \infty)$ since $f(y)/F(y) = 0$ for all $y > d_h$. Thus $\frac{g_y(y|\theta_1)}{g_\theta(y|\theta_0)} = (h_1 + b)\frac{f(y)}{F(y)}$ is decreasing on $[d_l, \infty)$ and obviously not constant. The result follows from Theorem 4.5 with $[m, M] = [\max\{d_l, r^0 - 1\}, R^0 - 1]$.

The proof of the second statement in part (a) is analogous since $d_h = \infty$.

(b) Choose $\theta = -b$. for any $\theta_0 \equiv -b_0 > -b_1 \equiv \theta_1$ with $(r^0 - 1, R^0)$ being the optimal policy at θ_1 ,

$$\sum_{t=r^0-1}^{R^0-1} (h + b_1)f(t) \sum_{s=r^0-1}^{R^0-1} sF(s) > \sum_{t=r^0-1}^{R^0-1} t(h + b_1)f(t) \sum_{s=r^0-1}^{R^0-1} F(s). \quad (\text{C.20})$$

It may be replaced by

$$\sum_{t=r^0-1}^{\min\{d_h, R^0-1\}} f(t) \sum_{s=r^0-1}^{\min\{d_h, R^0-1\}} sF(s) > \sum_{t=r^0-1}^{\min\{d_h, R^0-1\}} tf(t) \sum_{s=r^0-1}^{\min\{d_h, R^0-1\}} F(s). \quad (\text{C.21})$$

since all integrands are zero on $[\min\{d_h, R^0 - 1\}, R^0 - 1)$. If $F(y)$ is log-concave on $[d_l, d_h]$, i.e., $f(y)/F(y)$ is decreasing on $[d_l, d_h]$, it is log-concave on $[d_l, \infty)$ since $f(y)/F(y) = 0$ for all $y > d_h$. Thus $\frac{g_y(y|\theta_1)}{g_\theta(y|\theta_0)} = (h + b_1)\frac{f(y)}{F(y)}$ is decreasing on $[d_l, \infty)$ and obviously not constant. The result follows from Theorem 4.5 with $[m, M] = [r^0 - 1, \min\{d_h, R^0 - 1\}]$.

The proof of the second statement in part (b) is analogous since $d_l \leq r^*(b)$. \square

Proof of Lemma 4.4. We prove the lemma for $p(n)$ being log-concave. The proof for the other case where $p(n)$ is log-convex is analogous. We have the following property of $p(n)$: for any $t > 0$ and $n \leq n'$ with $a \leq n + t < n' + t \leq b$,

$$p(n + t)/p(n) \geq (\leq)p(n' + t)/p(n'), \quad (\text{C.22})$$

since $\frac{p(n+t)}{p(n)} = \prod_{i=1}^t \frac{p(n+i)}{p(n+i-1)} \geq \prod_{i=1}^t \frac{p(n'+i)}{p(n'+i-1)} = \frac{p(n'+t)}{p(n')}$ where the inequality follows from $\frac{p(n+i)}{p(n+i-1)} \geq \frac{p(n'+i)}{p(n'+i-1)}$ for all $i \geq 1$ and $a \leq n+i < n'+i \leq b$ by (4.28).

(a) By the non-negativity of $P(\cdot)$ and (4.28), it is equivalent to show that

$$\begin{aligned} P(n+1)P(n-1) &\leq P^2(n) \\ \iff (P(n) + p(n+1))P(n-1) &\leq (P(n-1) + p(n))P(n) \\ \iff P(n-1)p(n+1) &\leq P(n)p(n). \end{aligned} \tag{C.23}$$

However, (C.23) is satisfied since

$$P(n-1)\frac{p(n+1)}{p(n)} = \sum_{s=0}^{n-1} \frac{p(s)p(n+1)}{p(n)} \leq \sum_{s=0}^n p(s+1) < \sum_{s=0}^n p(s) = P(n)$$

where the first inequality follows from $\frac{p(n+1)}{p(n)} \leq \frac{p(s+1)}{p(s)}$ for all $s \leq n-1$ by (C.22).

(b) There exists an integer N such that $p(N+1)/p(N) < 1$, for, otherwise, $p(n+1) \geq p(n) \geq \dots \geq p(0) > 0$ for all n . However, this implies $\lim_{n \rightarrow \infty} p(n) \geq p(0) > 0$, which contradicts our assumption. It follows from (4.28) that $\frac{p(n+1)}{p(n)} \leq \frac{p(N+1)}{p(N)} < 1$ for all $n \geq N$. Thus $\lim_{n \rightarrow \infty} \frac{p(n+1)}{p(n)} < 1$. Hence $P(n) \leq P(0) = \sum_0^\infty p(i) < \infty$. To show that $P(n)$ is log-concave, it is equivalent to show that

$$\begin{aligned} P(n+1)P(n-1) &\leq P^2(n) \\ \iff (P(n) - p(n))P(n-1) &\leq (P(n-1) - p(n-1))P(n) \\ \iff P(n)p(n-1) &\leq P(n-1)p(n). \end{aligned} \tag{C.24}$$

(C.24) holds since

$$P(n)\frac{p(n-1)}{p(n)} = \sum_{s=n}^{\infty} \frac{p(s)p(n-1)}{p(n)} \leq \sum_{s=n}^{\infty} p(s-1) = P(n-1)$$

where the first inequality follows from $\frac{p(n)}{p(n-1)} \geq \frac{p(s)}{p(s-1)}$ for all $s \geq n$ by (C.22). \square

Bibliography

- Anupindi, R., Y. Bassok. 1999. Centralization of stocks: retailers vs. manufacturer. *Management Sci.* **45**(2) 178–191.
- Anupindi, R., Y. Bassok. 1999. Supply Contracts with Quantity Commitments and Stochastic Demand, S. Tayur, R. Ganeshan, and M. Magazine Eds. *Quantitative Models in Supply Chain Management*, Chapter 7. Kluwer Academic Publishers, Norwell, Massachusetts.
- Axsäter, S. 1990. Simple solution procedures for a class of two-echelon inventory problems. *Oper. Res.* **38** 64–69.
- Axsäter, S. 1993. Exact and approximate evaluation of batch-ordering policies for two-level inventory systems. *Oper. Res.* **41** 777–785.
- Axsäter, S. 1996. Using the deterministic EOQ formula in stochastic inventory control. *Management Science.* **42**(6) 830–834.
- Bagchi, U., J. Hayya, C. Chu. 1986. The effect of lead time variability: The case of independent demand. *J. Oper. Management.* **6**(2) 159–177.
- Bagnoli M. and T. Bergstrom. 2005. Log-concave Probability and Its Applications. *Economic Theory.* **26**(2) 445–469.
- Barlow, R.E., Proschan, F., 1975. *Statistical Theory of Reliability and Life Testing*. Holt, New York.
- Bassok, Y., R. Anupindi. 2008. Analysis of Supply Contracts with Commitments and Flexibility. *Naval Research Logistics* **55**(5) 459–477.
- Bather, J. 1966. A continuous time inventory model. *J. Appl. Probab.* **3** 538–549.
- Beim, D. 1996. The Determinants of Bank Loan Pricing. *Working paper* Columbia University.

- Berger, A. and G. Udell. 1990. Collateral, Loan Quality and Bank Risk. *Journal of Monetary Economics* **25** 21-42.
- Berger, A. and G. Udell. 1995. Relationship Lending and Lines of Credit in Small Firm Finance. *Journal of Business* **68** 351-381.
- Biais, B., C. Gollier. 1997. Trade credit and credit rationing. *The Review of Financial Studies* **10**(4) 903-937.
- Booth, J. 1992. Contract Costs, Bank Loans, and the Cross-Monitoring Hypothesis. *Journal of Financial Economics* **31** 25-41.
- Browne, S. and P. Zipkin. 1991. Inventory models with continuous stochastic demands. *The Annals of Applied Probability*. **1**(3) 419-435.
- Burkart, M., and T. Ellingsen. 2004. In-Kind Finance: A Theory of Trade Credit. *American Economic Review* **94** 569-590.
- Buzacott, J. A., R. Q. Zhang. 2004. Inventory management with asset-based financing. *Management Sci.* **50**(9) 1274-1292.
- Cachon, G. P. 2003. Supply chain coordination with contracts. A. G. de Kok, S. Graves, eds. *Supply Chain Management: Design, Coordination and Operation* Chapter 6. Elsevier, Amsterdam, The Netherlands.
- Cachon, G. P. 2004. The allocation of inventory risk in a supply chain: Push, pull, and advance-purchase discount contracts. *Management Sci.* **50**(2) 222-238.
- Cachon, G., P. Zipkin. 1999. Competitive and cooperative inventory policies in a two stage supply chain. *Management Sci.* **45**(7) 936-953.
- Chen, X., D. Simchi-Levi. 2004. Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case. *Oper. Res.* **52**(6) 887-896.
- Chen, F., and Y.-S. Zheng. 1993. Inventory Models with General Backorder Costs. *European Journal of Operations Research*. **65**(2) 175-186.
- Cuñat, V. 2007. Trade credit: Suppliers as debt collectors and insurance providers. *The Review of Financial Studies* **20**(2) 491-527.
- Daripa, A., J. Nilsen. 2005. Subsidizing inventory: a theory of trade credit and prepayment. *working paper* Birkbeck School of Economics, mathematics and Statistics, University of London. London, UK.

- Denis, D., K. Rodgers. 2007. Chapter 11: Duration, Outcome, and Post-Reorganization Performance. *Journal of Financial and Quantitative Analysis* **42**(1) 101–118.
- Dong, L., K. Zhu. 2007. Two-wholesale-price contracts: push, pull, and advance-purchase discount contracts. *Manufacturing Service Oper. Management* **9**(3) 291–311.
- Federgruen, A., A. Heching. 1999. Combined pricing and inventory control under uncertainty. *Oper. Res.* **47**(3) 454–457.
- Federgruen, A., C.-Y. Lee, 1990. The dynamic lot size model with quantity discount. *Naval Res Logist.* **37**. 707–713.
- Federgruen, A., Y.S. Zheng. 1992. An efficient algorithm for computing an optimal (r, Q) policy in continuous review stochastic inventory systems. *Oper. Res.* **40**(4) 808–813.
- Federgruen, A. and P. Zipkin. 1986a. An Inventory Model with Limited Production Capacity and Uncertain Demands I: The Average-cost Criterion. *Math. Oper. Res.* **11**, 193–207.
- Federgruen, A. and P. Zipkin. 1986b. An Inventory Model with Limited Production Capacity and Uncertain Demands II: The Discounted-cost Criterion. *Math. Oper. Res.* **11**, 208–215.
- Fernandez, F. R., Udell, G. and Carbo Valverde, S. 2008. Bank Lending, Financing Constraints and SME Investment. *21st Australasian Finance and Banking Conference (2008) Paper*.
- Fiacco, A.V. and J. Kyparisis. 1985. Sensitivity analysis in nonlinear programming under second order assumptions. In: A. Bagchi and H.Th. Jongen, Editors, *Systems and Optimazation*, Springer, Berlin.
- Frank, M., and V. Maksimovic. 2005. Trade Credit, Collateral, and Adverse Selection. *Working Paper* University of Maryland.
- Gallego, G. 1998. New bounds and heuristics for (Q, r) policies. *Management Science.* **44**(2) 219–233.
- Geoffrion, A. 1976. The purpose of mathematical programming is insight, not numbers. *Interfaces.* **7**(1) 81-92.
- Giannetti, M., M. Burkart, T. Ellingsen. 2006. What you sell is what you lend? Explaining trade credit contracts. *The Review of Financial Studies* Advance access published November 2008.
- Gupta, D., W. Cooper. 2005. Stochastic comparisons in production yield management. *Operations Research.* **53**(2) 377–384.

- Gupta, D., L. Wang. 2009. A stochastic inventory model with trade credit. *Manufacturing Service Oper. Management* **11**(1) 4–18.
- Hadley, G. and T. Whitin. *Analysis of Inventory Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1963.
- Huh, W. T., G. Janakiraman, A. Muharremoglu, A. Sheopuri. 2010. Optimal Policies for Inventory Systems with a Generalized Cost Model. *Working paper*. Columbia University, New York.
- Hull, J. 2008. *Options, Futures, and Other Derivatives (Seventh Edition)* Prentice Hall, Upper Saddle River, NJ.
- Iglehart, D. 1963a. Optimality of (s, S) policies in the infinite-horizon dynamic inventory problem. *Management Sci.* **9** 259–267.
- Iglehart, D. 1963b. Dynamic programming and stationary analysis of inventory problems. Chapter 1 in *Multistage Inventory Models and Techniques*, ed. H. Scarf, D. Gilford and M. Shelly. Stanford University Press.
- Jemai, Z., F. Karaesmen. 2005. The influence of demand variability on the performance of a make-to-stock queue. *Eur. J. Oper. Res.* **164** 195–205.
- Jain, N. 2001. Monitoring Costs and Trade Credit. *Quarterly Review of Economics and Finance* **41** 89–110.
- Jemai, Z., F. Karaesmen. 2005. The influence of demand variability on the performance of a make-to-stock queue. *Eur. J. Oper. Res.* **164** 195–205.
- Karlin, S., H. Scarf. 1958. Inventory models of the Arrow-Harris-Marschak type with time lag. K. Arrow, S. Karlin, H. Scarf, eds. *Studies in the Mathematical Theory of Inventory and Production*, Chapter 10. Stanford University Press, Stanford, CA.
- Koch, T. W. 1995. *Bank Management*. Dryden Press, Fort Worth, TX.
- Kouvelis, P., W. Zhao. 2009. Financing the newsvendor: supplier vs. bank, and the structure of optimal trade credit contracts. *Working paper*, Olin Business School, Washington University in St. Louis.
- Lando, D. 2004. *Credit Risk Modeling: Theory and Applications*, Princeton University Press, Princeton.
- Lariviere, M. 1999. Supply chain contracting and coordination with stochastic demand. S. Tayur, M. Magazine, R. Ganeshan, eds. *Quantitative Models for Supply Chain Management*. Kluwer, Dordrecht, The Netherlands.

- Lariviere, M. and E. Porteus. 2001. Selling to the newsvendor: An analysis of price-only contracts. *Manufacturing Service Oper. Management* **3**(4) 293–305.
- Leland, H. 2006. Princeton Lecture Series in Finance: Lecture 1: Pros and Cons of Structural Models: An Introduction. *Lecture Notes* University of California, Berkeley.
- Levi, R., R. O. Roundy, D. B. Shmoys. 2006. Primal-dual algorithms for deterministic inventory problems. *Math. Oper. Res.* **31** 267–284.
- Levi, R., R. Roundy, D. Shmoys and M. Sviridenko, 2008. A constant approximation algorithm for the one-warehouse multi-retailer problem, *Management Sci.* **54** 763–776.
- Levi, R., G. Perakis, J. Uichanco. 2011. The Data-Driven Newsvendor Problem: New Bounds and Insights. *Working Paper*. Massachusetts Institute of Technology.
- Maddah, B. S., M. Y. Jaber, N. E. Abboud. 2004. Periodic review (s, S) inventory model with permissible delay in payments. *J. Oper. Res. Soc.* **55** 147–159.
- Mian, S. L., and C. W. Smith. 1992. Accounts Receivable Management Policy: Theory and Evidence. *Journal of Finance* **47** 169–200.
- Morton, T. 1969. Bounds on the solution of the lagged optimal inventory equation with no demand backlogging and proportional costs. *SIAM Rev.* **11** 572–576.
- Muharremoglu, A., J. Tsitsiklis. 2008. A single-unit decomposition approach to multi-echelon inventory systems. *Oper. Res.* **56**(5) 1089–1103.
- Naddor, E., 1966. *Inventory Systems*, Wiley, New York.
- Nadiri, M. I. 1969. The Determinants of Trade Credit in the U.S. Total Manufacturing Sector. *Econometrica* **37** 408–423.
- Nagarajan, M., and S. Rajago0palan. 2008. Contracting under vendor managed inventory systems using holding cost subsidies. *Production and Operations Management* **17**(2) 200–210.
- Nahmias, S. 1993. *Production and Operations Analysis*. IRWIN, Homewood, IL.
- Narayanan, V. G., A. Raman, and J. Singh. 2004. Agency Costs in a Supply Chain with Demand Uncertainty and Price Competition. *Management Sci.* **51**(1) 120–132.
- Pasternack, B. A. 1985. Optimal pricing and return policies for perishable commodities. *Marketing Sci.* **4**(2) 166–176.
- Petersen, M. and R. Rajan. 1994. The Benefits of Lending Relationships: Evidence from Small Business Data. *Journal of Finance.* **49** 3-35.

- Petersen, M., and R. Rajan. 1997. Trade credit: theories and evidence. *The Review of Financial Studies* **10**(3) 661–691.
- Porteus, E. L. 2002. Foundations of Stochastic Inventory Theory. Stanford University Press, Palo Alto, CA.
- Puterman, M. 1975. A diffusion process model for a storage system. In *Logistics. Volume 1. TIMS Studies in the Management Sciences*. (M. Geisler, ed.) 143–159. North-Holland, Amsterdam.
- Rajan R.G., and L. Zingales. 1997. What do we know about capital structure? Some evidence from international data. *Journal of Finance* **50**(5) 1421–60.
- Rao, U. 2003. Properties of the periodic review (R, T) inventory control policy for stationary, stochastic demand. *Manufacturing Service Oper. Management* **5** 3753.
- Robb, D. J., E. A. Silver. 2004. Inventory management under date-terms supplier trade credit with stochastic demand and leadtime. *J. Oper. Res. Soc.* **55**(6) 692–702.
- Rosling, K. 1999. The square-root algorithm for single-item inventory optimization. *Working paper*, Department of Industrial Engineering, Växjö University, SE-351 95 Växjö, Sweden.
- Rosling, K. 2002. Inventory cost rate functions with nonlinear shortage costs. *Oper. Res.* **50**(6) 1007–1017.
- Scarf, H., 1960. The optimality of (s, S) policies in dynamic inventory problem. In *Mathematical Models in Social Sciences*, ed. K. Arrow, S. Karlin and P. Suppes, Stanford University Press.
- Sahin, I. 1979. On the stationary analysis of continuous review (s, S) inventory systems with constant lead times. *Oper. Res.* **27**(4) 717–729.
- Serfozo, R. and S. Stidham. 1978. Semi-stationary clearing processes. *Stochastic Process. Appl.* **6** 165–178.
- Sethi, S. P. and F. Cheng. 1997. Optimality of (s, S) Policies in Inventory Models with Markovian Demand. *Oper. Res.* **45**, 931–939.
- Shaked, M. and J. G. Shanthikumar. 2007, Stochastic Orders, Springer, New York.
- Shang, K. and S. Zhou. 2010. Optimal and Heuristic Echelon (r, nQ, T) Policies in Serial Inventory Systems with Fixed Costs. *Oper. Res.* **58**(2) 414–427.
- Shang, K., S. Zhou. 2010. A simple heuristic for echelon (r, nQ, T) policies in serial supply chains. *Oper. Res. Lett.* **37** 433–437.

- Song J. S. 1994a. The effect of lead time uncertainty in a simple stochastic inventory model. *Management Sci.* **40**(5) 603–612.
- Song, J. S., H. Zhang, Y. Hou, M. Wang. 2010. The effect of lead time and demand uncertainties in (r, q) inventory system. *Oper. Res.* **58**(1) 68–80.
- Song, J. S., and P. Zipkin. 1993. Inventory control in a fluctuating demand environment. *Oper. Res.* **41**(2) 351–370.
- Song, J. S., P. Zipkin. 1996a. Managing inventory with the prospect of obsolescence. *Oper. Res.* **44**(1) 215–222.
- Song, J. S., P. Zipkin. 1996b. The joint effect of lead time variance and lot size in a parallel processing environment. *Management Sci.* **42**(9) 1352–1363.
- Stauffer, G., G. Massonnet, C. Rapine and J.-P. Gayon. 2011. A simple and fast 2-approximation for deterministic lot-sizing in one warehouse multi-retailer systems. *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 67–79.
- Veinott, A. 1965. The optimal inventory policy for batch ordering. *Operations Research*. **13**(3) 424–432.
- Veinott, A. F., Wagner, H. 1965. Computing optimal (s, S) inventory policies. *Management Sci.*, **11** 525–552.
- Wang, Y., Y. Gerchak. 2001 Supply Chain Coordination When Demand Is Shelf-Space Dependent. *Manufacturing Service Oper. Management* **3**(1) 82–87.
- Whitin, T., 1953. *The Theory of Inventory Mngement*, Princeton University, Princeton, NJ.
- Wilner, B. S. 2000. The Exploitation of Relationships in Financial Distress: The Case of Trade Credit. *Journal of Finance* **55** 153–178.
- Wilson, N., B. Summers. 2002. Trade credit firms offered by small firms: Survey evidence and empirical analysis. *J. Bus. Finance Accounting* **29**(3) 317–351.
- Yang, A.S., J.R. Birge. 2011. How inventory is (should be) financed? Trade credit in supply chains with financial constraints and distress cost. *Working paper*, Booth School of Business, University of Chicago
- Zhang, H., 1998. A note on the convexity of service-level measures of the (r, q) system, *Management Science*, **44**(3) 431-432
- Zheng, Y.S., 1992. On properties of stochastic inventory systems. *Management Science*. **38**(1) 87–103.

- Zheng, Y.S., A. Federgruen. 1991. Finding optimal (s, S) policies is about as simple as evaluating a single policy. *Operations Research*. **39**(4) 654–665.
- Zhou, J., H. Groenevelt. 2008. Impacts of financial collaboration in a three-party supply chain. *Working paper*, Simon Graduate School of Business, University of Rochester.
- Zipkin, P., 1986. Inventory service-level measures: convexity and approximation, *Management Sci.*, **32**(8), 975–981.
- Zipkin, P., 2000, *Foundations of Inventory Management*, McGraw Hill-Irwin, New York.
- Zipkin, P., 2008a. Old and new methods for lost-sales inventory systems. *Oper. Res.* **56**(5) 1256–1263.
- Zipkin, P., 2008b. On the Structure of Lost-Sales Inventory Models. *Oper. Res.* **56**(4) 937–944.