Coordination Mechanisms for Supply Chains under Price and Service Competition

Fernando Bernstein
Fuqua School of Business
Duke University
Durham, NC 27708

Awi Federgruen
Graduate School of Business
Columbia University
New York, NY 10027

December 2002

Last Revised April 2006

Abstract

In a decentralized supply chain, with long-term competition between independent retailers facing random demands and buying from a common supplier, how should wholesale and retail prices be specified in an attempt to maximize supply-chain wide profits? We show what types of coordination mechanisms allow the decentralized supply chain to generate aggregate expected profits equal to the optimal profits in a centralized system, and how the parameters of these (perfect) coordination schemes can be determined. We assume that the retailers face stochastic demand functions which may depend on all of the firms' prices as well as a measure of their service level, e.g., the steady-state availability of the product. We systematically compare the coordination mechanisms when retailers compete only in terms of their prices, and when they engage in simultaneous price and service competition.

1 Introduction

We consider a decentralized supply chain with long-term competition between independent retailers facing random demands and buying from a common supplier. In this setting, we investigate how wholesale and retail prices should be specified in order to maximize supply-chain wide profits. The design of effective coordination mechanisms in supply chains has recently received considerable attention in the operations management literature, following on earlier work in economics (see e.g., Tirole 1988, and Mathewson and Winter 1984) and the marketing literature on channel coordination (see e.g., Jeuland and Shugan 1983, and Moorthy 1987). Ideally, a coordination mechanism allows the decentralized supply chain to generate aggregate expected profits equal to those in the first-best solution, i.e., the optimal profits in a centralized system. We refer to such mechanisms as perfect coordination schemes. This paper develops such schemes for settings where the retailers compete in terms of their pricing strategies, as well as those where they compete simultaneously in terms of their prices and long-term service levels.

A variety of pricing structures and contractual arrangements have been discussed in the operations management literature, see e.g., the surveys by Lariviere (1999), Tsay et al. (1999) and Cachon (2002). This parallels innovations in many industries where suppliers increasingly adopt non-standard pricing schemes to influence retail prices, retail sales and supply chain profits.[1] For example, the adoption of so-called revenue sharing schemes has revolutionized the video rental industry.

However, most of the literature considers coordination mechanisms for a supply chain with a single retailer, thereby avoiding the complications which arise under any type of competition between retailers. As to the sparse literature on coordination mechanisms for supply chains with competing retailers, a few papers (in particular Padmanbhan and Png 1995, 1997, Deneckere et al. 1996, 1997, van Ryzin and Mahajan 1999, Cachon 2002, and Bernstein and Federgruen 2005) address this question in a single period setting. Papers addressing infinite horizon models typically assume that demands occur at a constant deterministic rate and that all demands are satisfied fully and immediately, an ideal service level which, under deterministic demands, can easily be guaranteed. Here, retailer competition is

[1] See e.g., Ailawadi et al. (1999) who report that, across forty packaged good categories included in the Market Fact Book, no less than 37% of retail sales were made “on deal”, i.e., on the basis of one or several such pricing schemes.

confined to price or quantity competition. (See Chen et al. 2001, Bernstein and Federgruen 2003, and the references cited therein.)

Under random demands, the competitive dynamics among the retailers are considerably more complex, when viewed in a multi-period or infinite horizon setting. First, each retailer needs to complement its pricing strategy with an efficient strategy to replenish its inventory from the supplier. Second, the distribution of the random demands faced by a retailer depends, in general, on all the retailers' prices and service levels, i.e., the (steady-state) availability of their products. We observe an increasing number of industries in which some of the competing retailers aggressively attempt to obtain larger market shares by providing higher levels of service. For example, in the fierce competition between amazon.com and barnesandnoble.com, the latter initiated a massive advertising campaign promising same business day delivery in various parts of the country. These complications result in significant challenges when designing a coordination mechanism for the chain.

To analyze the mechanism design questions, we focus on two-echelon supply chains with a single supplier servicing a network of retailers. We assume one of several systems of demand processes whose distributions are functions of all retailer prices and all announced service levels, quantified as the firms' no-stockout frequency, i.e., the fraction of time during which the firm does not run out-of-stock.[2] Bernstein and Federgruen (2004b) consider alternative service measures, e.g., the likelihood with which customers receive delivery within a given promised time limit. We analyze a periodic review, infinite horizon model, with the retailers facing a stream of demands that are independent across time but not necessarily across firms. End-of-the-period inventories are carried over to the next period. We assume that stockouts are backlogged. Each retailer may place an order with the supplier at the beginning of each period. Similarly, the supplier may, at the beginning of the period, replenish her inventory from an outside source. The supplier fills the retailers' orders from her own inventory or, in case of stockouts, from an "emergency" or "backup" source. Such emergency procurements incur additional costs. The retailers and the supplier pay facility-specific inventory carrying and variable order costs. In addition, the retailers may incur out-of-pocket backlogging costs, which are proportional with the size of the backlogs. Contrary to most standard inventory models, but more representative of actual cost/service tradeoffs experienced in practice, ours does not require that direct backlogging costs exist. Even in their absence, every firm has a proper incentive to carry appropriate safety stocks, since a large stockout

[2]This is often referred to as the type-1 service level (see e.g., Nahmias 2001, §5.4.6).

frequency reduces the retailer's average sales and increases that of the competitors. We also show that backlog penalties may need to be charged (or paid) to the retailers as part of a coordination scheme. The exact optimal strategy for the centralized chain is unknown and of a prohibitively complex structure. We therefore define the first-best solution as the optimal supply chain wide profits achievable under an optimal pricing structure and assuming all facilities adopt some base-stock policy, i.e., each facility increases its inventory position to some given base-stock level, whenever the inventory has dropped below this level.

We first consider the case where the firms' service levels are exogenously specified. Here, we assume that the retailers respond to a given wholesale pricing scheme by non-cooperatively selecting their retail price along with a dynamic inventory replenishment strategy. We initially assume that each retailer chooses a stationary retail price to be used throughout the planning horizon. This assumption is satisfied, for example, when the model addresses supply chain coordination for a given year or season, with daily or weekly replenishment opportunities, during which retail prices remain constant by managerial choice or by necessity. Blinder et al. (1998) report on a detailed and comprehensive survey of 200 firms in the U.S. selected from a large variety of industries, firm sizes, and geographical regions, to document how rigid or sticky prices are and what factors explain this stickiness. No less than 45% of the firms vary their prices only on an annual basis and 60% conduct a price review at most twice a year. Correspondingly, 49% of the firms report making at most a single price change per item per year, and 65% at most two changes. The authors identify some twelve distinct theories to explain this price stickiness. Their book complements an earlier literature of econometric studies documenting a pervasive trend of price rigidity, for example, Carlton (1986), Cechetti (1986), and Kashyap (1995).

Under a static price choice, there is a unique Nash equilibrium of prices and associated inventory strategies in response to a linear wholesale pricing scheme. Since each firm's inventory strategy only impacts its own profit, the infinite horizon stochastic game can be reduced to a single stage game, in the sense that both games share the same set of Nash equilibria. In other words, even though the base-stock levels depend on the equilibrium prices and service levels, the replenishment dynamics can be decoupled from the strategic interations. We use this to show that the system can be coordinated with constant per unit wholesale prices specified, once and for all, at the beginning of the infinite planning horizon. The results contrast with those in the single-period model in Bernstein and Federgruen (2005) in which stockouts are assumed to result in lost sales. There, perfect coordination cannot be achieved with a simple linear wholesale pricing scheme. Instead, it is essential to combine such a scheme with a guarantee by the supplier to buy back any unsold units at a given (retailer specific) buy-back rate. That paper presents an alternative coordination mechanism, the so-called Price Discount Sharing scheme, under which the supplier subsidizes the retailer for part of the dollar amount the retailer discounts its retail price from a given list price. This scheme also needs to be combined with a buy-back guarantee. Moreover, in Bernstein and Federgruen (2005)'s single period model, the stochastic demand functions do not depend on the firms' service levels. That model therefore does not analyze how the equilibria and parameters of any coordination schemes are affected by these service levels.

We proceed with the full equilibrium model, in which the retailers compete in terms of two distinct strategic instruments: (i) their retail prices (or, equivalently, their expected sales targets), and (ii) their announced service levels (no-stockout frequencies). (Each firm continues to select a dynamic infinite horizon inventory strategy along with these two choices.) Under static pricing, the infinite horizon retailer competition game can again be reduced to a single stage game, now one in which each firm selects a price and a service level.

Under combined price and service level competition, a perfect coordination scheme can be designed on the basis of a specific vector of constant per unit wholesale prices, combined with a vector of constant per unit backlogging cost penalties to be paid by the retailers to the supplier (or vice versa).[3] This type of coordination scheme was first introduced by Celikbas et al. (1997) to coordinate the marketing and production functions within a single firm and by Lariviere (1999) in the context of a single retailer, single period model in which the retail price, and hence the demand distribution, is exogenously given. The scheme is also related to the "lost sales transfer payment" scheme in Cachon (2002, §5) for a setting with a single retailer facing an exogenously specified (Poisson) demand process, with stockouts resulting in lost sales: a given fee is paid by the retailer to the supplier (or vice versa) for every unit in lost sales. The backlogging penalties are most easily implemented when they are negative, i.e., when they are to be paid by the supplier to the retailer: here, the retailer is properly incented to report any backlogs so as to recover the backlogging penalties. If the penalty is positive, a possible way for the supplier to monitor backlogs at the retailers would involve rebate coupons to be distributed to the customer (perhaps along with the warranty or service registration card) and to be sent in to the supplier or a third party.

In §6, we discuss how, and to what extent, our coordinating mechanisms continue to

[3]These wholesale prices and backlogging penalties are again specified, once and for all, at the beginning of the infinite planning horizon.

apply when firms are allowed to change their price in each period. (Since the firms' service levels are defined in terms of long-run average fill rates, it only makes sense to treat them as static choices.) While under the assumption of static pricing the competition model reduces to a single stage game, it distinguishes itself from prior models in the literature by its ability to incorporate demand and cost implications of long-term service levels. Prior models of retailer competition with stochastic demands consider a single period and either assume that firms compete exclusively in terms of their prices, or they assume that firms compete in terms of their single period inventory levels under given prices. For example, van Ryzin and Mahajan (1999) consider a single period model in which customers choose retailers based on the availability of stocks. The authors confine themselves to linear wholesale pricing schemes with a constant per unit wholesale price, and investigate how close the best such scheme comes to achieving the first best solution. In contrast to this model, Cachon (2002, §5) shows that a linear wholesale pricing scheme achieves perfect coordination, if the retailer demands are perfectly correlated and arise in proportion to their initial stocks. Deneckere et al. (1996, 1997) consider a model with perfect competition and a (uniform) market clearing price which depends on the aggregate inventory of the retailers according to one of two demand functions, corresponding to two possible states of the general economy. Once again, a linear wholesale pricing scheme fails to coordinate the chain. The authors propose to combine it with a resale price maintenance scheme instead, under which the retailers are obliged to set their price above a given threshold. Padmanabhan and Png (1995, 1997) exhibit the benefits of a full returns policy in a two-retailer price competition model, with deterministic linear demand functions. In other models, the strategic importance of the inventory level consists in its ability to attract substitute demand from competitors that have run out of stock, while the primary demand functions are independent of any measure of service or inventory availability, see, e.g., Netessine and Rudi (2003).

Finally, as reviewed above, Bernstein and Federgruen (2005) consider the general single period price competition model with lost sales. The current paper relies heavily on Bernstein and Federgruen (2004b)'s single echelon model for price and service competition between retailers. See there for a literature review on earlier inventory models with price and/or service sensitive demand processes.

The remainder of this paper is organized as follows. Section 2 presents the model and notation. Section 3 characterizes the retailers' equilibrium behavior in response to a given wholesale pricing scheme as well as the centralized solution. Section 4 derives the proposed coordination mechanisms when service levels are predetermined and Section 5 when the retailers compete simultaneously in terms of their price and service levels. Section 6 discusses how our results can be extended to settings where dynamic pricing is allowed. Section 7 reports on a numerical study. Finally, Section 8 offers conclusions and a discussion of extensions to our model. All proofs are deferred to an Appendix.

2 Model and Notation

Consider a two-echelon supply chain with a supplier selling to N independent retailers, each facing random demands. We analyze a periodic-review infinite-horizon model in which, at the beginning of each period, each retailer may replenish its inventory by placing an order with the supplier and the supplier may choose a replenishment quantity to be procured from an outside source. Each retailer $i = 1, ..., N$ positions itself in the market by selecting a retail price p_i from a given interval $[p_i^{min}, p_i^{max}]$, as well as a steady-state service level $f_i \in [0.5, 1)$, defined as its (long-run) no-stockout frequency. (See Bernstein and Federgruen 2004b for alternative service measures.) For most of the paper, we assume that the retailers are required to adopt a stationary price. In §6 we discuss how our results and analysis extend to settings in which prices may, in principle, be varied each period.

The supplier must anticipate incoming orders with an appropriate replenishment strategy. This situation arises, for example, when the supplier's procurement mechanism is constrained by a capacity limit or when her replenishment orders fail to be filled instantaneously, but become available during or at the end of the period in which they are placed. When faced with stockouts, the supplier takes advantage of a backup or emergency source to fill the uncovered part of the retailer orders. For example, the supplier may subcontract at the last minute or schedule overtime production. Thus, all retailer orders can be filled at the requested level, albeit that procurements from the backup source are associated with significant additional costs. Subsequent to the initial price and service level choices, decisions are made in the following sequence: at the beginning of each period, all retailers simultaneously determine their order quantity for that period. Next, these orders are filled immediately (when necessary, with the help of an emergency order to clear a stockout at the supplier), after which the supplier decides on the next replenishment order.

Each retailer incurs holding costs which are proportional with the inventory it carries. Stockouts at the retailers are backlogged. In §8 we provide a discussion of the case where stockouts result in lost sales. A retailer may incur direct, out-of-pocket backlogging costs. If so, these are proportional with the backlog size. Thus, for each retailer $i = 1, ..., N$, let

h_i^+ = the per period holding cost for each unit carried in inventory,
h_i^- = the per period direct backlogging cost for any unit backlogged.

The supplier incurs variable "regular" procurement costs, as well as linear holding costs. The additional procurement cost associated with any "emergency" end of the period procurement is given by a convex function $h_0^-(\cdot)$, reflecting increasing marginal costs:

$h_0^-(x)$ = the additional cost of clearing an end of the period shortfall of x units.

Following Bernstein and Federgruen (2004b), the demand faced by each retailer i, in any period t, has a distribution which may depend on the entire vector of retail prices p as well as the entire vector of announced minimum service levels $f = (f_1, ..., f_N)$. We thus allow a firm to provide better than its announced service level. We assume, however, that customer demand for a given firm depends on its specified rather than its actual service level, similar to it being dependent on the specified technical quality of the product (e.g., the product's expected lifetime). (See Bernstein and Federgruen 2004b for a review of how information about the firms' minimum service level is available in a variety of industries, and how customers can avail themselves of estimates when service level targets fail to be publicly available.) Moreover, we show that in the absence of direct backlogging costs (i.e., when $h_i^- = 0$) a firm will always equate its actual service level to the minimum specified level, rendering the distinction between the two service level concepts a moot point. Similarly, when a backlogging penalty is charged to a retailer as part of a coordination scheme (see §5), the retailer is always induced to specify a (minimum) service level which equals its actual (long-run average) service level. Let $D_{it}(p, f)$ be the random demand faced by retailer i in period t, under the retail price vector p and the service level vector f, with general cumulative distribution function (cdf) $\tilde{G}_i(x|p, f)$. An assumption with important implications for the firms' equilibrium behavior is that the demand variables are of the multiplicative form, i.e.,

$$D_{it}(p, f) = d_i(p, f)\epsilon_{it}, \tag{1}$$

with ϵ_{it} a general continuous random variable whose distribution is stationary and independent of the vectors p and f. Thus, for all $i = 1, ..., N$, the sequence $\{\epsilon_{it}\}$ has a common general cdf $G_i(\cdot)$ with density function $g_i(\cdot)$, inverse cdf $G_i^{-1}(\cdot)$ and standard deviation s_i. Without loss of generality, we assume $E(\epsilon_{it}) = 1$, for all i and t. Thus, $ED_{it}(p, f) = d_{it}(p, f)$ and $\tilde{G}_i(x|p, f) = G_i\left(\frac{x}{d_i(p,f)}\right)$, so that $d_i(p, f)$ represents the expected demand for retailer i.

Under the multiplicative model, the absolute level of any fractile of the demand distribution $\tilde{G}_i$ may depend on the complete vector of prices p and service levels f, but the ratio of any pair of fractiles is independent of p and f. Another implication of the multiplicative model is that the coefficient of variation of any one-period demand is the exogenously given constant s_i, i.e., it is independent of p and f. See Bernstein and Federgruen (2004a) for a discussion of non-multiplicative demand models, with demand functions that depend on the price vector p, only. As is standard in virtually all inventory models, we assume that, for all $i = 1, ..., N$, the sequence of random variables $\{\epsilon_{it} : t = 1, 2, ...\}$ is independent, so that the same independence property applies to the sequence $\{D_{it}\}$. In contrast, the demands faced by the retailers in any given period may be correlated, following a general joint distribution. The mean sales functions satisfy the basic monotonicity properties:

$$\frac{\partial d_i(p,f)}{\partial p_i} \le 0, \qquad \frac{\partial d_i(p,f)}{\partial f_i} \ge 0, \qquad \frac{\partial d_i(p,f)}{\partial p_j} \ge 0, \qquad \frac{\partial d_i(p,f)}{\partial f_j} \le 0, \qquad j \neq i, \tag{2}$$

i.e., a retailer's demand volume decreases with its own price and increases with the price of any of its competitors, and it increases with its own service level and decreases with those of the competitors. We also assume that no firm's sales increase under a uniform price increase:

$$\text{(D)} \quad \sum_{j=1}^{N} \frac{\partial d_i}{\partial p_j} < 0 \quad \text{for all} \;\; i = 1, ..., N.$$

Finally, we denote by $\eta_i(p, f) = |\frac{\partial d_i(p,f)}{\partial p_i} \frac{p_i}{d_i(p,f)}|$ the absolute price elasticity for retailer i.

As in Bernstein and Federgruen (2004b), we focus on three classes of demand functions:

(I) **The Attraction Model**. Attraction models are among the most commonly used market share models, both in empirical studies and in theoretical models, see e.g., Leeflang et al. (2000). Here, we assume a fixed potential market size M, with each retailer's actual market share determined by a vector of attraction values $a = (a_1, ..., a_N)$. More specifically, retailer i's market share is given by $a_i / \sum_{j=0}^{N} a_j$ where a_0 is a constant representing the value of the no-purchase option. In our context, we assume that retailer i's attraction value a_i depends on its retail price p_i and service level f_i according to a general function $a_i = a_i(p_i, f_i)$. This gives rise to the system of expected demand functions:

$$d_i(p, f) = M \frac{a_i(p_i, f_i)}{\sum_{j=1}^{N} a_j(p_j, f_j) + a_0}. \tag{3}$$

Clearly, the attraction values are decreasing in the price and increasing in the service variable:

$$\frac{\partial a_i}{\partial p_i} \le 0 \;\; \text{and} \;\; \frac{\partial a_i}{\partial f_i} \ge 0 \;\; \text{for all} \;\; i = 1, ..., N. \tag{4}$$

We assume that each attraction function a_i is log-concave. Let $\tilde{a}_i = \log a_i$. Common specifications include the *Multinomial Logit* (MNL) Model, with $\tilde{a}_i(p_i, f_i) = b_i(f_i) - \alpha_i p_i$, where $\alpha_i \geq 0$ and the functions $b_i(\cdot)$ are twice differentiable, increasing and concave for $i = 1, ..., N$. (See Anderson et al. 1992 and Mahajan and van Ryzin 1999 for surveys of many econometric studies employing this specification.) The MNL model is a special case of the more general class of *Separable Attraction Functions*, where $\tilde{a}_i = b_i(f_i) + \alpha_i(p_i)$. A separable function $\tilde{a}$ is appropriate when the percentage increase in the attraction value of a firm due to a marginal change in its price is independent of the prevailing service level, or vice versa. Non-separable attraction functions are useful to represent increased or decreased sensitivity of the attraction values to price changes under a higher service level regime.

(II) **The Linear Model**. The average demand functions are linear in all prices and service levels, i.e., for positive constants b_i, e_{ij}, β_i and γ_{ij}:

$$d_i(p, f) = a_i - b_i p_i + \sum_{j \neq i} e_{ij} p_j + \beta_i f_i - \sum_{j \neq i} \gamma_{ij} f_j, \; i = 1, ..., N. \tag{5}$$

(III) **The Log-Separable Model**. This demand model assumes that a regular system of price-dependent demand functions $\{q_i(p)\}$ is scaled up or down, as a function of the service levels f offered by the different firms. This gives rise to the specification:

$$d_i(p, f) = \psi_i(f) q_i(p), \tag{6}$$

with q_i and ψ_i differentiable, satisfying the monotonicity properties $\frac{\partial \psi_i(f)}{\partial f_i} > 0, \frac{\partial \psi_i(f)}{\partial f_j} < 0, \frac{\partial q_i(p)}{\partial p_i} < 0, \frac{\partial q_i(p)}{\partial p_j} > 0, j \neq i$, and with the normalization $\psi_i(f, f, ..., f) = 1$. That is, if the firms choose identical service levels, their demands only depend on the prices. We assume that each function $q_i(p)$ is log-supermodular in (p_i, p_j) and each function $\psi_i(f)$ is log-supermodular in (f_i, f_j), i.e.,

$$\frac{\partial^2 \log q_i(p)}{\partial p_i \partial p_j} \geq 0 \;\text{ for all }\; i \neq j; \quad \frac{\partial^2 \log \psi_i(f)}{\partial f_i \partial f_j} \geq 0 \;\text{ for all }\; i \neq j. \tag{7}$$

Also,

$$-\frac{\partial^2 \log q_i(p)}{\partial p_i^2} \;>\; \sum_{j \neq i} \frac{\partial^2 \log q_i(p)}{\partial p_i \partial p_j}, \; i = 1, ..., N; \tag{8}$$

$$-\frac{\partial^2 \log \psi_i(f)}{\partial f_i^2} \;>\; \sum_{j \neq i} \frac{\partial^2 \log \psi_i(f)}{\partial f_i \partial f_j}, \; i = 1, ..., N. \tag{9}$$

These are standard conditions to guarantee a unique equilibrium in single-period competition models with linear costs, see Vives (2000). Milgrom and Roberts (1990) showed that virtually all standard classes of demand functions, including the Linear, Logit, Cobb-Douglas, and CES functions, satisfy (7)-(9) with minor parameter restrictions.

3 Best Response Policies and the Centralized Solution

In this section, we describe how each chain member optimally responds to chosen (infinite horizon) strategies by the other firms. In addition, we analyze the optimal solution in the centralized system.

Assume first that the supplier adopts an arbitrary vector of constant wholesale prices w. Since the supplier fills all retailer orders, π_i, the long-run average profit for retailer i, depends on the infinite horizon strategies of the competing retailers and the supplier only via (p_{-i}, f_{-i}),[4] the price and service level choices of the competing retailers. Given a choice (p_i, f_i), retailer i faces a stream of independent demands, identically distributed like $D_i(p, f)$ so that a simple base-stock policy with base-stock level y_i optimally complements the price and service level choices among all possible infinite horizon inventory strategies. Moreover, it is easily verified that under such a strategy, firm i's long-run average profit is given by:

$$\pi_i(p, f, y_i) = (p_i - w_i)d_i(p, f) - h_i^-(d_i(p, f) - y_i) - (h_i^- + h_i^+)E[y_i - D_i(p, f)]^+. \qquad (10)$$

Given (1), the base-stock level $y_i^*(p, f)$ which optimizes the profit in (10) subject to the constraint of providing a no-stockout probability at least equal to the specified level f_i, is:

$$y_i^*(p, f) = d_i(p, f)G_i^{-1}\left(\max\left\{f_i, \frac{h_i^-}{h_i^- + h_i^+}\right\}\right), \ i = 1, ..., N. \qquad (11)$$

Thus, the optimal base-stock level is a multiple of retailer i's expected single period demand. This multiple depends on the variability of these demands (via the distribution $G_i(\cdot)$), the relative magnitude of the cost rates h_i^- and h_i^+, and the specified (minimum) service level f_i. The multiple does not depend on any of the prices or the competitors' service levels. Substituting (11) into (10) and regrouping terms, we get the (reduced) profit functions

$$\tilde{\pi}_i(p, f) = [p_i - w_i - k_i(f_i)]d_i(p, f), \qquad (12)$$

where

$$k_i(f_i) = h_i^- - h_i^- G_i^{-1}\left(\max\left\{f_i, \frac{h_i^-}{h_i^- + h_i^+}\right\}\right) + (h_i^- + h_i^+)E\left[G_i^{-1}\left(\max\left\{f_i, \frac{h_i^-}{h_i^- + h_i^+}\right\}\right) - \epsilon_i\right]^+ \qquad (13)$$

[4]For any vector $x \in \Re^n$, let $x_{-i} = \{x_j : j \neq i\}, \ i = 1, ..., N$.

denotes the expected operational cost required to support one unit of sales (which in the multiplicative model is independent of the sales volume).

Thus, under an exogenously specified vector of service levels f^0, the retailers face an infinite-horizon sequential game which can be reduced to a single-stage game in which each firm i only chooses its price p_i and with profit functions $\{\tilde{\pi}_i(p, f^0)\}$. Specifically, if p^* is a Nash equilibrium in the single stage game, the price vector p^* combined with the N-tuple of base-stock policies with base stock levels $\{y_i^*(p^*, f^0), i = 1, ..., N\}$ is an equilibrium of Nash strategies in the infinite horizon game, and vice versa. Similarly, if the vector of service levels f is endogenously determined, the retailers face an infinite horizon game which can be reduced to a single stage game in which each firm i chooses a price p_i and a service level f_i, with profit function $\tilde{\pi}_i(p, f)$. Thus, we can restrict attention to the single-stage game, which we refer to as the retailer game.

From (12), note that $\tilde{\pi}_i$ is the product of the expected demand volume $d_i(p, f)$ of firm i and its profit margin $[p_i - w_i - k_i(f_i)]$. All service levels in f impact on a firm's expected demand, but only the firm's own service level impacts its profit margin. It is easily verified that the $k_i(\cdot)$ functions are differentiable and increasing. We assume, without loss of generality, that $\lim_{p \nearrow \infty} \tilde{\pi}_i(p) = 0$, and $\lim_{p \nearrow \infty} \frac{\partial \tilde{\pi}_i}{\partial p_i}(p) < 0$. In particular, we assume that the vector of upper limits $p^{max} = (p_1^{max}, ..., p_N^{max})$ is chosen sufficiently large that

$$\frac{\partial \tilde{\pi}_i}{\partial p_i}(p^{max}) < 0. \tag{14}$$

We now turn our attention to the supplier. As shown above, it is optimal, in a decentralized system, for all retailers to adopt some base-stock policy. Note that the specific choices of the base-stock levels have no impact on the revenues or costs incurred by the supplier. Assuming the retailers choose (p, f) as their prices and service levels, the supplier faces an i.i.d. stream of aggregate orders, the common distribution of which equals that of the aggregate single period consumer demand in the system. Let $D_0(p, f)$ denote a random variable with this distribution. (The cdf $\tilde{G}_0(x|p, f)$ of $D_0(p, f)$ can be obtained from the joint distribution of the $\{\epsilon_i\}$ variables. It depends on p and f only via the vector of mean demands d, i.e., $\tilde{G}_0(x|p, f) = \tilde{G}_0(x|d)$.) It is thus optimal for the supplier to adopt a (modified) base-stock policy, see e.g., Federgruen and Zipkin (1986). Under a modified base-stock policy, the supplier increases her inventory position, each period, to a level as close as possible to a base-stock level. If no capacity limit prevails, the base stock level can be achieved in every period. In the presence of a capacity limit, a full capacity order is placed when the difference between the base stock level and the period's beginning inventory is equal to or larger than the capacity limit.

Similar to those of the retailers, it is easy to compute the supplier's optimal base-stock level $y_0(p,f)$, see Zipkin (2000). Let $C_0(p,f)$ denote the expected holding and emergency procurement costs incurred by the supplier under an optimal (modified) base-stock policy. Since C_0 depends on (p,f) only via the cdf $\tilde{G}_0$ and since the latter depends on (p,f) only via the vector of mean demands d, it is possible to write $C_0(p,f) = \tilde{C}_0(d)$. In some cases, $\tilde{C}_0$ can be obtained in closed form. Consider, for example, the case where the supplier's orders are uncapacitated and her expediting costs are linear with cost rate h_0^-. If the random demand components $\{\epsilon_i : i = 1, ..., N\}$ follow a general multivariate Normal distribution, $D_0(p,f)$ is Normal itself and its mean $\mu_0(p,f)$ and standard deviation $\tilde{s}_0(p,f)$ can be obtained as a closed form function of the functions $\{d_i(p,f)\}$ and the variance-covariance matrix of $\{\epsilon_i\}$. Moreover, because $D_0(p,f)$ is Normal, $C_0(p,f) = (h_0^- + h_0^+)\phi\left(\Phi^{-1}\left(\frac{h_0^-}{h_0^- + h_0^+}\right)\right)\tilde{s}_0(p,f)$, see Zipkin (2000, Chapter 6), with $\phi(\cdot)$ and $\Phi(\cdot)$ the pdf and cdf of the standard Normal. For example, when the ϵ_i-variables are independent,

$$C_0(p,f) = \tilde{C}_0(d) = (h_0^- + h_0^+)\phi\left(\Phi^{-1}\left(\frac{h_0^-}{h_0^- + h_0^+}\right)\right)\sqrt{\sum_{i=1}^{N} d_i^2(p,f)s_i^2}. \tag{15}$$

In this case, $\partial\tilde{C}_0/\partial d_i = (\tilde{s}_i/\tilde{s}_0)c_0^i$, where c_0^i represents the supplier's per unit cost if she serves retailer i exclusively and $\tilde{s}_i = d_i s_i$, the standard deviation of retailer i's demand.

When designing coordination mechanisms, it is essential to characterize the first-best solution which arises when optimizing the centralized system. Such a system would adopt a price and service level vector (p^I, f^I) but the accompanying fully optimal supply chain wide replenishment strategy is unknown, see e.g., Federgruen and Zipkin (1986) and Zipkin (2000). In defining the first best solution, we therefore restrict attention to replenishment strategies under which each retailer's inventory is governed by some base-stock policy.[5] As shown above, the same (modified) base-stock rule for the supplier's inventory optimally complements such retailer base-stock policies, regardless of whether one considers the supplier's profit or the aggregate profit in the supply chain. With this restriction, let $\tilde{\pi}_I(p,f)$ denote the optimal system wide long-run average profit under the price vector p and vector of service levels f.

[5]Many supply chain models with exogenously specified demands are based on the assumption that all facilities adopt a base stock policy, see e.g., Graves and Willems (2000) and Ettl et al. (2000) and the references therein. These models have been implemented successfully at various product divisions of Eastman Kodak, IBM and other companies.

Note that

$$\begin{aligned}\tilde{\pi}_I(p,f) &= \sum_{j=1}^{N}(p_j - c_j - k_j(f_j))d_j(p,f) - C_0(p,f) \\ &= \sum_{j=1}^{N}\tilde{\pi}_j(p,f) + \left\{\sum_{j=1}^{N}(w_j - c_j)d_j(p,f) - C_0(p,f)\right\},\end{aligned}$$

where the expression within curled brackets represents the expected profit earned by the supplier. For any $f \in [0.5,1)^N$, let $p^I(f)$ denote a maximum of the continuous function $\tilde{\pi}_I(\cdot, f)$ on the compact cube $X_{i=1}^N[p_i^{min}, p_i^{max}]$ and let (p^I, f^I) denote a global maximum of the (continuous) function $\tilde{\pi}_I(\cdot,\cdot)$, which exists because $\lim_{f_i \nearrow 1}\tilde{\pi}_I(p,f) = -\infty,\ i = 1, ..., N$.

4 Coordination Under Price Competition

In this section, we show that under exogenously specified service levels f^0, perfect coordination can be achieved with simple constant per unit wholesale prices. This coordinating vector of wholesale prices is, in fact, unique. We also characterize its dependence on f^0.

Theorem 1 *Assume that the vector of service levels f^0 is exogenously specified. (a) There exists a vector $w^*(f^0)$, with*

$$w_i^*(f^0) = p_i^I(f^0) - k_i(f_i^0) - \frac{p_i^I(f^0)}{\eta_i(p^I(f^0), f^0)} \leq p_i^I(f^0) - k_i(f_i^0), \tag{16}$$

such that p^I arises as the unique price equilibrium in the retailer game induced by this vector of wholesale prices $w^*(f^0)$. Moreover, $w^*(f^0)$ is the only vector of constant wholesale prices under which $p^I(f^0)$ arises as a price equilibrium. (b) Assume, in addition, that $p_i^I(f^0) > c_i + k_i(f_i^0) + \frac{\partial \tilde{C}_0(p^I(f^0), f^0)}{\partial d_i}, i = 1, ..., N$. Then, $c_i + \frac{\partial \tilde{C}_0(p^I(f^0), f^0)}{\partial d_i} < w_i^*(f^0)$.

In our basic model, holding costs are independent of the wholesale prices. Since capital costs usually are a major component of inventory carrying costs, one may wish to assume that each holding cost rate h_i^+ increases with the wholesale price w_i, as follows: $h_i^+(w_i) = \rho_i w_i + h_i^0,\ i = 1, ..., N$. The above result also holds under this assumption.

To compute the coordinating wholesale price w_i^* for any given retailer i, it suffices to know its cost structure and demand function along with the optimal price vector $p^I(f^0)$: k_i

is easily computable, merely knowing retailer i's distribution $G_i(\cdot)$, cost parameters h_i^+ and h_i^- and its service level f_i^0, while the price elasticity η_i can be determined from the shape of this retailer's expected demand function $d_i(p, f^0)$ and the vector $p^I(f^0)$, alone. Note that k_i depends on w_i if h_i^+ does. In general, computation of w_i^* requires the determination of the unique value of w_i for which $w_i + k_i(f_i^0(w_i))$ crosses the value $p_i^I\left(1 - \frac{1}{\eta_i(p^I)}\right)$. Only in the "basic" model, where h_i^+ is independent of the wholesale price w_i (i.e., when $\rho_i = 0$), is the identity for w_i^* in the theorem a closed form expression for w_i^*. This identity also shows that under the coordinating wholesale pricing scheme, each retailer i incurs a total per unit expected cost equal to its retail price p_i^I, multiplied with a "discount" factor $\left(1 - \frac{1}{\eta_i(p^I)}\right)$ which is an increasing function of the retailer's price elasticity, another manifestation of the "inverse elasticity rule," see Tirole (1988, p. 66).

The coordinating wholesale prices $w^*(f^0)$ characterized in Theorem 1 can be complemented with the use of fixed periodic payments or franchise fees to allow for an arbitrary allocation of supply chain profits among the supplier and the retailers.

The proof of Theorem 1 shows that the retailers can be induced to adopt any desired vector of retail prices, and this with a unique vector of wholesale prices w^*. Assuming that a fixed vector of retail prices p^0 is targeted, it is of interest to investigate how the coordinating wholesale prices will change in response to a change in one of the firms' service levels f^0:

Proposition 1 *Assume a specific retail price vector p^0 is targeted. (a) If the service level f_i^0 of some retailer i is increased, the corresponding wholesale price w_j^* is increased for all of its competitors $j \neq i$, in each of the demand models (I), (II) and (III). (b) If the service level f_i^0 of some retailer i is increased, this will result in a decrease of its coordinating wholesale price w_i^* under the Linear and Log-Separable demand models (II) and (III). Under the Attraction model, w_i^* will decrease (increase) if*

$$k_i'(f_i^0) + [p_i^0 - w_i^* - k_i(f_i^0)]^2 \frac{\partial^2 d_i}{\partial p_i \partial f_i} \geq (\leq) 0. \tag{17}$$

If the wholesale price is set to achieve perfect coordination (as opposed to targeting a fixed retail price vector), these monotonicities are less clear, since the optimal price vector p^I in a centralized system has a complex dependence on the given service level f^0.

5 Coordination Under Price and Service Competition

Consider now the case where the retailers simultaneously compete in terms of their prices and service levels. Here, service levels are endogenously determined as part of the equilibrium strategies of the retailers, as opposed to being specified as exogenous input parameters. In this setting, a simple linear wholesale pricing scheme no longer suffices to coordinate the supply chain. However, perfect coordination can, in general, be achieved if a linear wholesale pricing scheme is combined with a backlogging penalty scheme under which each retailer pays the supplier a given (possibly negative) penalty for each unit backlogged, in each period.

To guarantee perfect coordination under a combined wholesale price / backlog penalty scheme, we need a restriction on the distributions of the random factors $\{\epsilon_i\}$. In particular, we need to ensure that the functions $k_i(f_i)$ be convex. Recall that $k_i(f)$ denotes retailer i's expected inventory and backlogging costs per unit of sales, when guaranteeing a service level f. The following Lemma can be found in Bernstein and Federgruen (2004b).

Lemma 1 *(a) $k_i(\cdot)$ is increasing and differentiable, with*

$$k_i'(f_i) = 0,\ \textit{for } f_i < \frac{h_i^-}{h_i^- + h_i^+} \ \textit{and } k_i'(f_i) = \frac{(h_i^+ + h_i^-)f_i - h_i^-}{g_i(G_i^{-1}(f_i))},\ \textit{for } f_i \geq \frac{h_i^-}{h_i^- + h_i^+}.$$

(b) $k_i(\cdot)$ is convex and $\lim_{f_i \uparrow 1} k_i'(f_i) = \infty$, for all distributions G_i such that:

(PF_2) *G_i is log-concave or, equivalently, is a Polya Frequency function of order 2 (PF_2) for all $x \geq G_i^{-1}(0.5)$, and g_i has infinite support, where $G_i^{-1}(\cdot)$ denotes the inverse of the G_i-distribution.*

Condition (PF_2) is satisfied for all distributions whose density function decreases beyond the median, e.g., the Normal and Exponential distributions and many specifications of the Gamma and Weibull distributions. In the remainder of this section, we assume that (PF_2) holds.

We now demonstrate that perfect coordination can be achieved with a linear wholesale pricing scheme, combined with a constant set of backlogging penalties, per unit backlogged in each period. (These penalties, when positive, are paid by the retailer to the supplier and, when negative, by the supplier to the retailer.) Let $k_i^*(f_i)$ denote firm i's service level cost function when the backlogging cost rate h_i^- is replaced by h_i^{-*}.

Theorem 2 *Assume that the following condition holds:*

$$h_i^+ \geq \gamma_i(p^I, f^I)\ g_i(G_i^{-1}(f_i^I)),\ \ i = 1, ..., N, \tag{18}$$

and let

$$w_i^* = p_i^I - k_i^*(f_i^I) - \frac{p_i^I}{\eta_i(p^I, f^I)} \leq p_i^I - k_i^*(f_i^I), \tag{19}$$

$$h_i^{-*} = \frac{h_i^+ f_i^I - \gamma_i(p^I, f^I)\ g_i(G_i^{-1}(f_i^I))}{1 - f_i^I},\ \ i = 1, ..., N, \tag{20}$$

where the $\gamma_i(p^I, f^I)$-factors are given by:

$$\begin{aligned}
\textit{Attraction Model:} \qquad \gamma_i(p^I, f^I) &= -\frac{\partial \tilde{a}_i(p_i^I, f_i^I)}{\partial f_i} \Big/ \frac{\partial \tilde{a}_i(p_i^I, f_i^I)}{\partial p_i}, \\
\textit{Linear Model:} \qquad \gamma_i(p^I, f^I) &= \frac{\beta_i}{b_i}, \\
\textit{Log-Separable Model:} \qquad \gamma_i(p^I, f^I) &= -\frac{\partial \ln \psi_i(f^I)}{\partial f_i} \Big/ \frac{\partial \ln q_i(p^I)}{\partial p_i}.
\end{aligned}$$

Then, the centralized solution (p^I, f^I) arises as a price and service level equilibrium in the retailer game under the (constant) wholesale prices w^* and backlogging penalties h^{-*}. In other words, the scheme where retailer i is charged the constant wholesale price w_i^* and a penalty $h_i^{-*} - h_i^-$ for each unit backlogged in each period, induces perfect coordination.

The coordinating backlogging penalties are specified as affine functions of the holding cost values $\{h_i^+\}$, where the coefficient in the linear term is given by $f_i^I/(1 - f_i^I)$. Note that the simplistic choice $h_i^{-*} = h_i^+ \left(\frac{f_i^I}{1-f_i^I}\right)$ which equates the critical fractile $\frac{h_i^{-*}}{h_i^{-*}+h_i^+}$ to f_i^I, would result in "too high" a backlogging penalty and would therefore likely result in the retailer adopting a service level $f_i > f_i^I$. The simplistic choice is coordinating when the demand functions do not depend on the service levels. Since a firm's demand increases when it increases its service level, a smaller backlogging penalty suffices to induce the firm to adopt the service level in f^I. Note that the amount by which the simplistic choice is reduced, is proportional to the marginal demand sensitivity with respect to the firm's own service level. Finally, since in general $k_i^*(\cdot) \neq k_i(\cdot)$, the coordinating wholesale prices w^* are different from those arising when the service levels are fixed a priori as $f = f^I$, i.e., when firms compete in terms of their prices only.

The conditions in Theorem 2 thus guarantee that the centralized solution (p^I, f^I) is an equilibrium set of prices and service levels. In general, it is hard to guarantee that this price and service level vector arise as the only possible equilibrium choices. One case where uniqueness can be guaranteed is the important class of MNL Models. For this class of demand functions, Theorem 4 in Bernstein and Federgruen (2004b) establishes that the single stage game employed in the proof of Theorem 2 has a unique Nash equilibrium. This equilibrium must be an interior point of the feasible action space and must therefore satisfy the first-order conditions as its unique solution as well. Note from (12) that $\partial \log \tilde{\pi}_i / \partial p_i \nearrow \infty$ as $p_i \searrow w_i + k_i(f_i)$, while $\frac{\partial \log \tilde{\pi}_i}{\partial p_i}(p_i^{max}, p_{-i}) \leq \frac{\partial \log \tilde{\pi}_i}{\partial p_i}(p_i^{max}, p_{-i}^{max}) < 0$, where the first inequality follows from $\frac{\partial^2 \log \tilde{\pi}_i}{\partial p_i \partial p_j} \geq 0$ (as is easily verified) and the second one from (14). Moreover, the first-order condition (27) takes the simple form: $-\alpha_i k_i'(f_i) + b_i'(f_i) = 0$. By the proof of Theorem 2, this implies that under the proposed coordinating scheme, (p^I, f^I) arises as the only possible equilibrium. For other types of Attraction models, it is harder to guarantee that a unique Nash equilibrium exists. Based on similar arguments as for the MNL Model, one can show that in the Linear model (p^I, f^I) arises as the unique price and service level equilibrium, see Theorem 8 in Bernstein and Federgruen (2004b).

In the Log-Separable model, if the functions $q_i(p)$ and $\psi_i(f)$ are Linear, Logit, Cobb-Douglas or CES (with minor parameter restrictions), the single-stage game in the proof of Theorem 2 is supermodular. Thus, even if multiple equilibria exist, there is an equilibrium $(\overline{p}, \overline{f})$ which is component-wise largest among all Nash equilibria and an equilibrium $(\underline{p}, \underline{f})$ which is component-wise smallest. Moreover, the following tatônnement scheme converges to $(\underline{p}, \underline{f})$ when starting in (p^{min}, f^{min}), with $f_i^{min} = 0.5$, and to $(\overline{p}, \overline{f})$ when starting in (p^{max}, f^{max}), with $f_i^{max} = 0.9999$. In the k-th iteration, (p^k, f^k) is obtained from (p^{k-1}, f^{k-1}) by determining $(p^k, f^k) = \arg\max_{p_i, f_i} \tilde{\pi}_i(p_i, f_i, p_{-i}^{k-1}, f_{-i}^{k-1})$. Convergence of the scheme to the same point when started in (p^{min}, f^{min}) or in (p^{max}, f^{max}) thus provides an easy numerical test that the equilibrium is unique.

In the MNL and Linear Models, the equilibrium service level of any retailer i is entirely independent of any of the wholesale prices, as well as any of the cost rates pertaining to its competitors. This is immediate from the first order conditions (26) and (27). For $i = 1, ..., N$, let f_i^* denote retailer i's equilibrium service level in the absence of any backlogging penalty imposed by the supplier. The following proposition shows that the coordinating penalty for any retailer i, $h_i^{-*} - h_i^-$, is positive [negative] when f_i^* is lower [greater] than its optimal centralized service level f_i^I. The additional positive [negative] penalty increases [decreases] the cost of backlogs, inducing the retailer to adopt a higher [lower] service level. This equivalence fails to hold for other types of demand functions.

Proposition 2 *Assume that the demand functions are of the MNL or Linear Models, and* *satisfy condition* (18)*. Then,* $h_i^{-*} - h_i^- \geq 0$ *if and only if* $f_i^* \leq f_i^I$*,* $i = 1, ..., N$*.*

Finally, the existence of a coordinating scheme is based on condition (18), which implies that $h_i^+ + h_i^{-*} \geq 0$, a condition necessary to guarantee that the functions $k_i^*(\cdot)$ be convex. Condition (18) is generally satisfied for sufficiently large service levels $\{f_i^I\}$, provided that $\gamma_i(p^I, f^I)$ remains bounded. Indeed, note that $\lim_{f_i \uparrow 1} g_i(G_i^{-1}(f_i)) = \lim_{x \to \infty} g_i(x) = 0$ for any distribution with unbounded support, since $E(\epsilon_i) = \int_0^\infty x g_i(x) dx = 1$. Under the Attraction Model, the $\gamma_i(p^I, f^I)$-factors remain bounded as $f_i \uparrow 1$ when, for example, $\tilde{a}_i$ is a separable function (as in the MNL-model), or when it is submodular, i.e., $\frac{\partial^2 \tilde{a}_i}{\partial p_i \partial f_i} \leq 0$, since in these cases the factor is positive and decreasing in f_i. For the case of the Linear Model, the $\gamma_i(p^I, f^I)$-factors are constants independent of f^I. Finally, under the Separable Model, $\gamma_i(p^I, f^I)$ is decreasing in f_i^I since $\psi_i(\cdot)$ is log-concave by (7) and (9).

Observe that the coordinating wholesale prices w^* are always specified to provide the retailers a positive margin, i.e., $p_i^I - w_i^* - k_i(f_i^I) > 0$. In addition, they are decreasing in the retailer's expected sales volume. (In the case of the Attraction Model, retailer i's margin is also an increasing function of the retailer's market share d_i/M.) Thus, wholesale prices are discounted on the basis of expected sales volumes. Our model thus provides an economic rationale for this most prevalent type of discounting, even though the cost structure may fail to exhibit any economies of scale with respect to the retailers' sales volumes. A similar observation was made in Bernstein and Federgruen (2003) for an infinite horizon model with deterministic demands.

6 Dynamic Pricing

Thus far, we have assumed that each of the retail firms selects a price at the beginning of the planning horizon, and maintains that price thereafter. As explained in §2, the up-front restriction to a constant price is often a necessity or it is managerially desirable, as is evidenced by the fact that close to half of the firms in the U.S. conduct a price review no more often than annualy, while 61% review their prices at most twice a year. In other settings, the price may, in principle, be varied in each period, along with the firm's order quantity. It can be shown that the same N-tuples of infinite horizon strategies, identified in Theorems 1 and 2, continue to represent Nash equilibria in this relaxed strategy space. For example, in the case of simultaneous price and service competition, assume that all of firm i's competitors adopt the service vector f^*_{-i} and stationary pricing strategies p^*_{-i}, along with infinite horizon base-stock policies with base stock levels y^*_{-i}. It is then optimal for firm i to adopt the service level f^*_i, along with a constant price p^*_i and the infinite-horizon base-stock policy with base stock level $y^*_i(p^*, f^*)$, even though the price in each period may be changed. This result follows from Fedegruen and Heching (1999, Theorem 7) and Chen and Simchi-Levi (2004). Given the choices of firm i's competitors, firm i faces a combined pricing and inventory planning Markov Decision Problem in which it is optimal to select the constant price p^*_i, as well as the above base-stock policy. Thus, while Federgruen and Heching (1999) have shown that significant profit can be gained by lowering (increasing) the retail price when inventories are high (low), these benefits arise only in non-stationary settings, or in finite horizon settings where the end-of-the-horizon truncation of the planning process generates a type of non-stationarity. Moreover, this N-tuple of strategies is a subgame perfect Nash equilibrium: even if the firms deviate from their policies for a finite number of periods, resuming the N-tuple of stationary strategies continues to be a Nash equilibrium thereafter.

In particular, even though price variations are allowed, these equilibrium strategies continue to employ a constant, stationary price for each firm under any given (stationary) pricing scheme by the supplier. However, it is in general hard to preclude the existence of alternative, more complex, equilibrium strategies, even if the associated single stage games (see the proofs of Theorems 1 and 2) can be guaranteed to have a unique Nash equilibrium. Tirole (1988) and Fudenberg and Tirole (1991) discuss the possibility of infinitely many alternative Nash strategies based on tacit collusion among the retailers. Thus, the existence of those alternative Nash equilibria makes it less certain that the proposed pricing schemes in Theorems 1 and 2 result in perfect coordination, i.e., induce the retailers to adopt the centralized optimal solution. (Even so, much has been written about the extent to which these alternative stratgies have practical relevance, see e.g., Shapiro 1989.)

Ever since the seminal paper by Maskin and Tirole (1988), much of the economics literature on dynamic oligopoly models has restricted attention to Markov Perfect Equilibria (MPE) only, see, e.g., Ericson and Pakes (1995), Fershtman and Pakes (2000) and Curtat (1996). Here, each firm's strategy must prescribe actions as a function of the prevailing "payoff sensitive" system state only. The latter is defined as the minimal state specification which is sufficient to describe future payoffs to the firms.[6] In our case, the payoff sensitive

[6]Maskin and Tirole (1988, page 553) motivate their restriction as follows: "We have several reasons

state vector in each period is the vector of stationary inventory levels in this period. Under the restriction to MPEs, the uniqueness result is straightforward under the following modification of the rules of engagement between the supplier and the retailers: assume that each firm is allowed to sell (part of) its inventory at the beginning of each period back to the supplier, with full credit, before determining its price and replenishment decisions for the current period. It is then a dominant strategy for each firm to initially "sell back" all of its inventory and then face its price and replenishment decisions with a starting inventory level of zero. In other words, under the credit option, the system starts each period in the unique state where all starting inventory levels are zero.[7] Since the system returns in each period to the same state, the above N-tuple of strategies is the unique MPE in the modified system. (Assuming that the firms start with a vector of inventory levels below the vector of equilibrium base-stock levels, these strategies can be implemented without any firm ever using the sell back option.)

Finally, in settings where the retailers may be able to monitor and enforce an alternative equilibrium of strategies based on tacit price collusion, any incentive for them to do so (as opposed to adopting the coordinating strategies) may be eliminated by reallocating $\tilde{\pi}_I(p^I, f^I)$, the maximum possible aggregate supply chain wide profits, via periodic fixed transfer payments in the form of franchise fees or supplier allotments. More specifically, if firm i's long-run average profit value under some alternative equilibrium equals $\hat{\pi}_i$, since $\sum_{i=0}^{N} \hat{\pi}_i \leq \tilde{\pi}_I(p^I, f^I)$, it is possible to identify transfer payments which, in conjuction with the above pricing schemes, would induce all firms to adopt the intended equlibrium and have all be better off.

for restricting our attention to Markov strategies. Their most obvious appeal is their simplicity. Firms' strategies depend on as little as possible while still being consistent with rationality. More relevant from our perspective is that Markov strategies seem at times to accord better with the customary conception of a reaction in the informal industrial organization literature than do, say, the reactions emphasized in the repeated game (or 'supergame') tradition, the best-established formal treatment of dynamic oligopoly to date."

[7] If a firm faces a backlog at the beginning of a period, it is clearly optimal to bring the firm's inventory level up to zero by making an initial purchase to clear the backlog.

7 Numerical Study

In this section, we report on a numerical study conducted to provide a comparison of the supply chain performance under price competition, under combined price and service competition, and under an optimal centralized solution. The purpose of the study is to understand how firms operate in these settings under different market conditions (e.g., customer sensitivities to prices and service levels, variability of demand), and how the retailers' position in the market (as given by their service levels) affects the resulting equilibrium strategies relative to the centralized solution. In addition, we assess the impact of changes in the supplier's cost structure. In all cases, we compute the parameters of the coordination schemes and discuss the impacts of their implementation.

The numerical study consists of a base scenario, as well as 13 alternative scenarios obtained by varying one (set of) parameter(s) at a time. All scenarios have $N = 3$ retailers and assume linear demand functions, as in (5), and Normally distributed variables $\{\epsilon_i : i = 1, ..., N\}$ which are independent of each other and have mean one and standard deviation $s_i = 0.5$. Recall that under linear demand functions, a unique price equilibrium $p^*(f^0)$ exists for any given vector f^0 as well as a unique pair (p^*, f^*) under combined price and service competition. As mentioned in §3, the choice of Normal distributions allows for a closed form expression of the supplier's expected cost C_0. For the base scenario I, we have:

$$
\begin{aligned}
d_1(p, f) &= 50 - 18p_1 + 7p_2 + 7p_3 + 100f_1 - 20f_2 - 20f_3, \\
d_2(p, f) &= 50 + 3p_1 - 10p_2 + 3p_3 - 20f_1 + 100f_2 - 20f_3, \quad \text{and} \\
d_3(p, f) &= 50 + 3p_1 + 3p_2 - 10p_3 - 20f_1 - 20f_2 + 100f_3.
\end{aligned}
$$

That is, retailer 1 has a clientele which is significantly more sensitive to price changes by any of the three retailers than its two competitors. On the other hand, the customers of all three retailers exhibit the same sensitivity to uniform price changes in the industry. (Note that $b_i^{tot} = 4$ for all $i = 1, 2, 3$, where $b_i^{tot} = b_i - \sum_{j \neq i} e_{ij}$.)

The retailers face identical cost parameters $c_i = 10$, $h_i^+ = 4$, and $h_i^- = 0$. Finally, $h_0^+ = 0.6$ and $h_0^- = 6$, and $p_i^{min} = 0$ and $p_i^{max} = 30$, $i = 1, 2, 3$. By choosing $h_0^- = 6$ and $c = 10$, we assume that the additional per unit cost to clear a shortage at the supplier is 60% of the normal variable procurement cost. When evaluating decentralized solutions, we assume that each retailer i pays the supplier a constant per unit wholesale price $w_i = 13.5$. (The supplier's 35% markup corresponds approximately with the profit-to-revenue ratio in the centralized solution for the base instance.)

For each of the 14 scenarios, we have computed both the centralized and decentralized solution, each for all service level combinations obtained by varying the service level of each retailer on a grid from 0.5 to 0.99, with a width of 0.01. (It is very hard to determine the global optimum (p^I, f^I) of the centralized system since the centralized profit function π_I typically has many local optima.) We thus determine (p^I, f^I) as the service level vector which, along with the best corresponding retail prices, results in the largest system-wide profit value among all grid points.

Scenarios (II)-(VI) investigate the impact of different customer price sensitivities. In Scenario (II), we set $e_{12} = 5, e_{13} = 9$, maintaining the total price sensitivity b^{tot}. However, while in the base scenario retailer 1's customers are equally sensitive to the prices charged by retailers 2 and 3, in (II) they are considerably more sensitive to retailer 3's price, perhaps because its product is a closer substitute or because it is geographically closer than retailer 2. In scenario (III), we increase b_1^{tot} from 4 to 6 by setting $e_{12} = 5, e_{13} = 7$. Thus, retailer 1 now has a larger total price sensitivity as well as larger sensitivities to price changes by individual retailers. In (IV), we reduce b_1^{tot} from 4 to 3 by setting $e_{12} = 7.5, e_{13} = 7.5$, maintaining the symmetry between retailers 2 and 3. Scenarios (V) and (VI) investigate the impact of reduced and increased price elasticities for retailers 2 and 3, by setting $b_2 = b_3 = 8$ and $b_2 = b_3 = 12$, respectively.

The first six scenarios assume that all retailers have identical sensitivities with respect to their own service level. In scenarios (VII) and (VIII), we decrease and increase this coefficient for retailer 1 from its value in the base scenario, $\beta_1 = 100$, to $\beta_1 = 80$ and to $\beta_1 = 120$, respectively. The former case represents a setting where retailer 1's customers are significantly more price-sensitive but care less about service than those of the competitors. In the latter case, retailer 1's customers are more demanding with respect to both attributes. Scenario (IX) specifies $\gamma_{ij} = 30$, for all $i \neq j$, and represents a setting with increased service-level competition. Finally, in scenario (X), we assess what impact the inventory-related cost parameters at the supplier have on the centralized and decentralized solutions by setting $h_0^+ = 3, h_0^- = 30$. In the last four scenarios (XI)-(XIV), we investigate the impact of demand variability, varying the coefficient of variation of the demand distributions from 0.2 to 0.6.

Table 1 reports, for each of the 14 scenarios, the optimal service levels in a centralized system f^I, the coordinating wholesale prices under price competition and exogenous service levels set at f^I, the equilibrium service levels f^*, and the parameters of the coordination scheme under simultaneous price and service level competition.

Sc.	f^I	w^* (p$-$*only*)	f^*	w^* (*simult.*)	h_1^{-*}	h_2^{-*}	h_3^{-*}
I	(0.75,0.98,0.98)	(13.2,17.0,17.0)	(0.79,0.90,0.90)	(13.4, 16.5, 16.5)	-2.1	147.6	147.6
II	(0.75,0.96,0.99)	(13.3,16.0,18.2)	(0.79,0.90,0.90)	(13.4, 15.5, 17.6)	-2.1	52.9	342.7
III	(0.73,0.95,0.97)	(12.6,14.3,15.0)	(0.79,0.90,0.90)	(12.7, 13.9, 14.5)	-2.8	34.8	84.0
IV	(0.76,0.99,0.99)	(13.6,18.7,18.7)	(0.79,0.90,0.90)	(13.8, 18.1, 18.1)	-1.7	342.7	342.7
V	(0.73,0.97,0.97)	(15.8,22.3,22.3)	(0.79,0.92,0.92)	(16.1, 21.9, 21.9)	-2.8	72.6	72.6
VI	(0.77,0.99,0.99)	(11.6,13.0,13.0)	(0.79,0.87,0.87)	(11.7, 12.4, 12.4)	-1.3	351.6	351.6
VII	(0.50,0.97,0.97)	(13.1,16.4,16.4)	(0.73,0.90,0.90)	(13.7, 15.8, 15.8)	-3.1	84.0	84.0
VIII	(0.83,0.99,0.99)	(13.4,17.7,17.6)	(0.83,0.90,0.90)	(13.4, 17.1, 16.8)	-0.3	342.7	147.6
IX	(0.50,0.96,0.96)	(12.4,14.7,14.7)	(0.79,0.90,0.90)	N/A	N/A	N/A	N/A
X	(0.74,0.97,0.97)	(14.9,16.7,16.7)	(0.79,0.90,0.90)	(15.1, 16.2, 16.2)	-2.5	84.0	84.0
XI	(0.91,0.99,0.99)	(13.3,17.8,17.8)	(0.93,0.97,0.97)	N/A	N/A	N/A	N/A
XII	(0.86,0.99,0.99)	(13.3,17.5,17.5)	(0.89,0.94,0.94)	N/A	N/A	N/A	N/A
XIII	(0.81,0.98,0.98)	(13.3,17.2,17.2)	(0.84,0.92,0.92)	(13.4, 16.8, 16.8)	-2.8	135.5	135.5
XIV	(0.70,0.98,0.98)	(13.2,16.8,16.8)	(0.74,0.87,0.87)	(13.4, 16.1, 16.1)	-1.4	155.7	155.7

Table 1 - Equilibrium service levels and coordinating schemes

Table 2 exhibits, again for each scenario, the retailers' prices in the centralized and decentralized systems, both for $f = f^I$ and $f = f^*$. (Recall that $p^I(f^I) = p^I$ is the optimal price vector under centralization, while $p^*(f^*) = p^*$ represents the equilibrium price vector under simultaneous price and service competition.)

Sc	$p^I(f^I)$	$p^I(f^*)$	$p^*(f^I)$	$p^*(f^*)$
I	(18.5, 22.4, 22.4)	(18.2, 20.8, 20.8)	(17.6, 20.1, 20.1)	(17.3, 18.7, 18.7)
II	(18.6, 21.6, 23.2)	(18.3, 20.4, 21.3)	(17.6, 19.7, 20.4)	(17.3, 18.7, 18.7)
III	(16.1, 19.8, 20.5)	(16.2, 19.0, 19.3)	(16.1, 19.2, 19.6)	(16.2, 18.5, 18.5)
IV	(20.1, 24.0, 24.0)	(19.4, 21.9, 21.9)	(18.4, 20.6, 20.6)	(17.9, 18.8, 18.8)
V	(22.7, 30.0, 30.0)	(23.3, 30.0, 30.0)	(18.8, 23.6, 23.6)	(18.7, 22.3, 22.3)
VI	(16.0, 17.8, 17.8)	(15.6, 16.1, 16.1)	(17.0, 18.2, 18.2)	(16.5, 16.5, 16.5)
VII	(16.8, 21.7, 21.7)	(17.1, 20.4, 20.4)	(16.1, 19.9, 19.9)	(16.4, 18.6, 18.6)
VIII	(19.8, 23.2, 22.9)	(19.2, 21.3, 21.3)	(18.6, 20.5, 20.2)	(18.1, 18.7, 18.7)
IX	(15.4, 19.5, 19.5)	(16.3, 18.2, 18.2)	(15.4, 18.7, 18.7)	(16.4, 17.5, 17.5)
X	(19.2, 22.1, 22.1)	(19.0, 20.9, 20.9)	(17.4, 19.8, 19.8)	(17.3, 18.7, 18.7)
XI	(18.3, 21.3, 21.3)	(18.4, 20.9, 20.9)	(17.3, 18.6, 18.6)	(17.4, 18.3, 18.3)
XII	(18.5, 21.8, 21.8)	(18.4, 21.0, 21.0)	(17.5, 19.3, 19.3)	(17.4, 18.5, 18.5)
XIII	(18.5, 21.9, 21.9)	(18.0, 20.5, 20.5)	(17.5, 19.5, 19.5)	(17.1, 18.3, 18.3)
XIV	(18.6, 22.8, 22.8)	(18.0, 21.1, 21.1)	(17.6, 20.6, 20.6)	(17.1, 19.0, 19.0)

Table 2 - Prices in the centralized and decentralized systems

The following general conclusions can be drawn from the numerical study.

(A) Under price and service level competition, retailers differentiate themselves less than in a centralized system.

This pattern applies across all fourteen scenarios (as well as others we have evaluated but whose results we do not report in the paper). Here, we measure the degree of differentiation in the service levels by the span of the vector of service levels, i.e., $\max_i f_i - \min_i f_i$. Similarly, we measure the degree of differentiation in the prices by the span of the price vector. As far as the differentiation in the service levels is concerned, an even stronger contrast arises between the centralized and the decentralized systems: across all scenarios, the service level of the best service provider is lower and that of the worst service provider higher under competition than in the centralized solution. In other words, supply chain wide profits are maximized by providing clearer and more distinct alternatives than in a competitive setting. In the latter, retailers tend to adopt clustered price and service level profiles. This phenomenon is reminiscent of what is known to be the case in the classical Hotelling model, under exogenously given prices, where the retailers differentiate themselves by their location, as opposed to their service level. Similarly, Borenstein and Netz (1999) have substantiated why airlines differentiate their departure times less under competition, compared to the centralized solution.

(B) Prices and service levels tend to be less sensitive to parameter changes in the decentralized system than in the centralized system.

Compare first the solutions in each of the scenarios (II)-(X), with base scenario (I). The observed pattern holds throughout, except for the service levels of retailers 2 and 3 in scenarios (V) and (VI). Compared to (I), these change by one percentage point in the centralized system, while they change by 2% and 3%, respectively, in the equilibrium of the decentralized system. This occurs because the operational cost functions $k_i(\cdot)$ are convex and extremely sensitive to changes in the service level, when it is close to one. In scenarios (V) and (VI) changes in the price sensitivities of retailers 2 and 3 affect their choice of service levels. However, these changes are less pronounced in the centralized system, where the service levels are significantly closer to one.

Turning next to the remaining scenarios (XI)-(XIV), the same observation applies to the comparison of the service levels for retailers 2 and 3. In addition, in these scenarios, changes in the demand variability tend to have a somewhat stronger impact on the prices in the decentralized as opposed to the centralized system, even under fixed service levels. This can be explained as follows. System-wide profits in the centralized system depend both on the variability of the individual retailers' demands as well as that of the aggregate demand through the supplier's cost function $C_0(p, f)$. In contrast, in the decentralized system, prices result from a retailer game in which only the coefficients of variation of the individual retailer demands matter. For example, going from scenario (XI) to (XII) corresponds to a 50% increase in the coefficient of variation of demand faced by each retailer, while the coefficient of variation of the aggregate demand experienced by the supplier increases only by 22.5%.

(C) In the centralized solution, the retailers tend to charge more than in the decentralized system.

Two opposite forces influence how prices compare in both systems. On the one hand, double marginalization implies higher retail prices relative to those optimal in the centralized system. On the other hand, competition among the retailers tends to lower their prices. In most scenarios, the competitive effect dominates, confirming what is often assumed to be the case in classical oligopoly models in which firms only compete in prices. In scenario (VI), however, $p^I(f^I) < p^*(f^I)$ and $p^I(f^*) < p^*(f^*)$. In that case, the increased values of b_2 and b_3 seem to reduce the relative effect that competition has on these firms, so the effect of double marginalization dominates.

(D) Competition may result in higher or lower service levels compared with those arising in a centralized system.

Thus, while prices are generally higher in the centralized system, the same fails to be the case for the retailers' service levels. As mentioned in (A), all that can be predicted is that the lowest service provider in the centralized solution increases its service level in the decentralized system, while the highest service level provider in the centralized system decreases its service level under decentralization.

(E) The combined wholesale price/backlogging rate coordinating scheme derived in Section 5 achieves perfect coordination in almost all scenarios, i.e., in almost all cases $h^{+*} + h^{-*} \geq 0$. In those cases where $h^{+*} + h^{-*} < 0$, we can identify coordinating wholesale price/backlogging rate schemes under which aggregate supply chain profits come very close to the first-best level.

As discussed in Section 5 under simultaneous competition, the unique vector of wholesale prices w^* and backlogging penalty rates h^{-*} coordinate the supply chain if $h^+ + h^{-*} \geq 0$.

Proposition 2 shows that $h_i^{-*} \geq 0$ if and only if $f_i^I \geq f_i^*$. In all scenarios, this inequality is satisfied for retailers 2 and 3, while $f_1^I < f_1^*$, implying that $h_1^{-*} < 0$. However, $h_1^+ + h_1^{-*}$ remains positive for all scenarios, except for (IX), (XI) and (XII).

The general shape of the coordinating backlogging rate in (20) shows that, at least under linear demand functions, h_i^{-*} increases rapidly to $+\infty$ as f_i^I approaches one. This is illustrated in Table 1 by some of the large positive penalty rates charged to retailers 2 and 3. The magnitude of h_i^{-*} is also influenced by the difference between f_i^* and f_i^I. For example, the equilibrium service levels for retailer 1 in scenarios (IV), (II), and (III) differ from those in the centralized system by 1%, 2%, and 3%, respectively, and the (absolute) value of h_1^{-*} increases accordingly. Also, h_1^{-*} is very close to zero in scenario (VIII), where the centralized and decentralized service levels for retailer 1 differ only by a few decimal points, while in scenario (IX), $h_1^+ + h_1^{-*} < 0$ as f_1^* exceeds f_1^I by 29%, the largest difference among all retailers and all scenarios. Scenarios (XI) and (XII) are the only other two in which $h_1^+ + h_1^{-*} < 0$. These two scenarios have in common that (i) $f_1^* > f_1^I$ and (ii) f_1^I is itself relatively high. It follows from Proposition 2 that (i) implies that $h_1^{-*} < 0$ and the expression in (20) shows that the absolute value of h_1^{-*} can be relatively large when f_1^I is relatively high.

Even if in scenarios (IX), (XI), and (XII), it is not possible to achieve perfect coordination with the linear wholesale price/backlogging penalty scheme presented in Section 5, it is possible to come close to the first-best solution. For example, we have observed that in (IX), when charging retailer 1 a wholesale price of \$12.4 and retailers 2 and 3 wholesale prices of \$14.7, along with backlogging penalties of −\$3.5, \$52.9 and \$84, we obtain an equilibrium that leads to aggregate supply chain profits within 2.3% of the first best solution.

Finally, the difference between retailer i's coordinating wholesale price under simultaneous price and service competition and that under price-only competition is given by $k_i^*(f_i^I) - k_i(f_i^I)$, see (16) and (19). It is easy to verify that for a fixed service level f, the function $k_i(f)$ given in (13) is strictly increasing in h_i^-. Then, $k_i^*(f_i^I) - k_i(f_i^I) > (<) 0$ if and only if $h_i^{*-} > (<) h_i^-$. In the scenarios investigated in this numerical study, $h_i^- = 0$, and $h_2^{-*},\ h_3^{-*} > 0$ while $h_1^{-*} < 0$. Thus, retailers 2 and 3 have to pay (sometimes hefty) backlogging penalty fees to the supplier under simultaneous competition, but their wholesale prices are lower than in the setting where the service levels are exogenously set at f^I. At the same time, in the coordinating scheme under simultaneous competition, retailer 1 is subsidized for its consumer backlogs, thus providing an incentive to reduce its inventory levels, but is charged a higher wholesale price than it would be in a setting where the centralized optimal service level is exogenously determined.

We conclude this section with a discussion of the sensitivity of the centralized and decentralized solutions to the service levels, and of the firms' relative performance (in terms of market shares and profits) in both systems. To this end, Table 3 characterizes the centralized solution under given service levels $f^I = (0.75, 0.98, 0.98)$, $f^* = (0.79, 0.90, 0.90)$, as well as four other service level combinations. These are obtained by deviating from f^I with 2 alternative service levels for retailers 1 and 2, each. (Since in the base scenario retailers 2 and 3 are symmetric, similar results are obtained for retailer 3.) The table displays the optimal prices and corresponding sales volumes for the three retailers, and the aggregate optimal profits in the supply chain. The last three columns display the coordinating wholesale prices (under the given vector of service levels). In Table 4, we evaluate for the same six vectors of service levels, the unique equilibrium which arises in the decentralized system. After listing the equilibrium prices, we exhibit the equilibrium mean sales volumes, the equilibrium expected profit of each firm, and the percentage gap between the centralized solution and the aggregate supply chain wide profits in the decentralized uncoordinated system.

f^0	$p^I(f^0)$	$d^I(f^0)$	$\tilde{\pi}_I(f^0)$	$w^*(f^0)$
$f^I = (0.75, 0.98, 0.98)$	(18.5, 22.4, 22.4)	(65.5, 12.5, 12.5)	618	(13.2, 17.0, 17.0)
$f^* = (0.79, 0.90, 0.90)$	(18.2, 20.8, 20.8)	(56.8, 15.2, 15.2)	574	(13.2, 17.0, 17.0)
(0.50, 0.98, 0.98)	(17.4, 22.3, 22.3)	(60.7, 14.2, 14.2)	598	(13.2, 16.8, 16.8)
(0.99, 0.98, 0.98)	(20.4, 21.9, 21.9)	(49.1, 16.5, 16.5)	510	(13.0, 16.1, 16.1)
(0.75, 0.70, 0.98)	(17.8, 19.1, 21.7)	(56.4, 13.3, 12.6)	511	(13.0, 16.3, 16.3)
(0.75, 0.99, 0.98)	(18.6, 22.7, 22.4)	(66.9, 10.3, 13.0)	617	(13.2, 17.0, 17.0)

Table 3 - Centralized Solutions

f^0	$p^*(f^0)$	$d^*(f^0)$	$\tilde{\pi}^*(f^0)$	$\tilde{\pi}^*_S(f^0)$	gap
$f^I = (0.75, 0.98, 0.98)$	(17.6, 20.1, 20.1)	(50.9, 25.5, 25.5)	(123, 63, 63)	323	7.4%
$f^* = (0.79, 0.90, 0.90)$	(17.3, 18.7, 18.7)	(42.4, 27.6, 27.6)	(83, 69, 69)	310	7.5%
(0.50, 0.98, 0.98)	(16.5, 20.2, 20.2)	(46.7, 26.5, 26.5)	(102, 68, 68)	317	7.2%
(0.99, 0.98, 0.98)	(20.0, 20.2, 20.2)	(33.5, 26.9, 26.9)	(61, 70, 70)	278	6.1%
(0.75, 0.70, 0.98)	(17.1, 17.1, 19.8)	(42.2, 25.6, 23.0)	(82, 55, 51)	288	6.8%
(0.75, 0.99, 0.98)	(17.6, 20.4, 20.1)	(52.2, 23.3, 26.0)	(130, 53, 66)	322	7.5%

Table 4 - Decentralized Solutions

Both the centralized and the decentralized solutions can be quite sensitive to the service levels. As far as the former is concerned, the optimal aggregate profits vary between \$618 under f^I and \$266 under the vector $(0.99, 0.5, 0.5)$ (not shown in Table 3), where retailer 1 adopts an extremely high service level and its competitors a very low one. Even when changing the service levels from f^I to f^*, the optimal centralized profit drops by around 7%. Interestingly, an increase of one percentage point in the service level of retailer 2 in f^I has only a minor impact on aggregate profits under centralized and decentralized control, while that increase may represent significantly higher operational costs to that firm itself. In fact, under price competition retailer 2 experiences a considerable decrease in its equilibrium profit (from \$63 to \$53).

Combined price and service competition causes all retailers to reduce their prices $p^* = p^*(f^*)$ beyond their equilibrium prices $p^*(f^I)$ under the exogenously given service levels f^I, which themselves are lower than p^I. This results, for example, in prices for retailers 2 and 3 that are 17% lower than in the centralized solution. The change in prices is, as discussed above, associated with a reduction of the high-end retailers' service levels from 98% to 90% and an increase of retailer 1's service level from 75% to 79%. The equilibrium sales volumes for retailers 2 and 3 under combined price and service competition are approximately 8% higher than their values under price competition with $f^0 = f^I$, while those of retailer 1 decline by approximately 16%. (Thus, aggregate sales are close to equal in the two equilibria.) The retailers' market shares vary dramatically under the three solutions. That of retailer 1 is close to 72% under the centralized solution, but reduces to less than 50% under price competition and $f^0 = f^I$, and to 43% under combined price and service competition. Retailer 1 charges a significantly lower price than its competitors, both in the decentralized and the centralized solution. In both settings, the higher price sensitivity of retailer 1's customers results in its positioning itself as the low-price/low-service-level alternative in the market.

The coordinating wholesale prices generate significant gross profit margins $(w^* - c)$ for the supplier which are sometimes larger, but sometimes smaller than the gross retailer margins $(p^I - w^*)$. For example, under $f^0 = f^I$, the margin vectors are $(\$3.2, \$7.0, \$7.0)$ and $(\$5.3, \$5.4, \$5.4)$, respectively. Consider now the impact of a change in one of the service levels on the coordinating wholesale prices. As an example, set $f^0 = f^I$. With f_1^0 and f_3^0 fixed, as f_2^0 is increased from 0.7 to 0.99, its own coordinating wholesale price w_2^* increases from \$16.3 to \$17.0, while the retail price p_2 increases from \$17.1 to \$20.4. In contrast, Proposition 1 shows that if the targeted retail price p_2 was kept constant, w_2^* would decrease.

Finally, note that the profit gap between the centralized and decentralized systems under fixed service levels ranges, in these examples, from 6.1% to 7.5%. In addition, while price-only competition with service levels set at f^I leads to a gap of 7.4%, simultaneous competition leads to a significantly higher gap of 14.1% relative to the centralized system under f^I. Thus, the ability to pre-specify the retailers' service levels at their system-optimal values may, under decentralized control, be very beneficial for the overall performance of the supply chain.

8 Conclusions and Extensions

We have addressed a general model for two-echelon supply chains with several competing retailers served by a common supplier. Each retailer's stochastic demand function depends on its own retail price, as well as those of its competitors, but also on the service levels guaranteed by all firms. The retailers' service levels are defined as their no-stockout frequency. Most of our analysis has focused on three basic classes of stochastic demand functions which depend on the vector of retail prices p and the vector of service levels f, i.e., the Attraction Models, the Linear Models, and the Log-Separable Models.

Focusing first on the case where the firms' service levels are exogenously specified, we have shown that perfect coordination can be achieved by a simple linear wholesale pricing scheme (with constant per unit wholesale prices). When service levels are endogenously determined, i.e., when the retailers simultaneously compete in terms of their prices and their service levels, coordination can again be achieved with a linear wholesale pricing scheme, albeit that this wholesale pricing scheme needs to be combined with a set of constant per unit backlogging penalties to be paid by the retailers to the supplier, or vice versa. Finally, we derive a number of managerial insights which arise from the numerical study.

An important assumption in our paper, common to most stochastic inventory models, is that stockouts at the retailers are fully backlogged. It is of interest to consider the alternative setting where stockouts result in lost sales. Assume first that the retailers do not guarantee any particular service level, so that the mean demand functions only depend on the vector of retail prices. In this case, a firm's service level is defined as the long-run frequency with which it does not run out of stock at the end of a period. Similar to the proof of Theorem 1, it can again be shown that, under a vector of constant wholesale prices, a Nash equilibrium of infinite horizon strategies exists in which each retailer adopts a given price and a stationary base stock policy, provided a Nash equilibrium exists in a related single-stage game. This single-stage game has been analyzed in Bernstein and Federgruen

(2005). Theorem 4 there shows that this single-stage game has a unique equilibrium vector of prices which is monotone in the vector of wholesale prices, under a condition with respect to the shape of the distributions of the error factors. (The condition is satisfied, e.g., for Exponentials and Normal distributions with coefficient of variation less than or equal to one.) Under these conditions, it is possible to show that, as in the case of full backlogging, perfect coordination can be achieved with a linear wholesale pricing scheme. A situation, similar to this lost sales model, arises in our model with backlogging if there are explicit out-of-pocket backlogging costs that depend on the retail price (e.g., by being specified as a percentage of the retail price). It is easily verified that the profit functions in the related single stage game are structurally similar to the profit functions discussed above. If the retailers guarantee specific service levels, the situation is more complex, either when these service levels are exogenously specified or when the retailers engage in simultaneous price and service competition, and it is no longer possible to guarantee the existence of a coordinating wholesale pricing scheme. (The difficulty results from the fact that the retailers' profit functions no longer have increasing differences in their own retail price and the wholesale price charged by the supplier.)

References

Ailawadi, K., P. Farris and E. Shames (1999), "Trade Promotion: Essential to Selling through Resellers," *Sloan Management Review*, Fall, 83-92.

Anderson, S.P., A. de Palma and J.F. Thisse (1992), *Discrete Choice Theory of Product Differentiation*, The MIT Press, Cambridge, MA.

Bell, D., R. Keeney and J.D. Little (1975), "A Market Share Theorem," *Journal of Marketing Research*, Vol. XII, 136-141.

Bernstein, F. and A. Federgruen (2003), "Pricing and Replenishment Strategies in a Distribution System with Competing Retailers," *Operations Research*, 51, 3, 409-426.

Bernstein, F. and A. Federgruen (2004a), "Dynamic Inventory and Pricing Models for Competing Retailers," *Naval Research Logistics*, 51, 2, 258-274.

Bernstein, F. and A. Federgruen (2004b), "A General Equilibrium Model for Retail Industries with Price and Service Competition," *Operations Research*, 52, 6, 868-886.

Bernstein, F. and A. Federgruen (2005), "Decentralized Supply Chains with Competing Retailers under Demand Uncertainty," *Management Science*, 51, 1, 18-29.

Blinder, A., E. Canetti, D. Lebow and J. Rudd (1998), *Asking About Prices: A New Approach to Understanding Price Stickiness*, Russell Sage Foundation, New York, NY.

Cachon, G. (2002), "Supply Chain Coordination with Contracts," to appear in *Handbooks in Operations Research and Management Science*, S. Graves and F. de Kok (Eds.), North Holland, Amsterdam.

Carlton, D. (1986), "The Rigidity of Prices," *American Economic Review*, 76, 637-658.

Cechetti, S. (1986), "The Frequency of Price Adjustments: A study of Newsstand Prices of Magazines: 1953 to 1979," *Journal of Econometrics*, 31, 255-274.

Celikbas, M., J. Shanthikumar, and J. Swaminathan (1997), "Coordinating Production Quantities and Marketing Efforts through Penalty Schemes," UC Berkeley Working Paper.

Chen, F., A. Federgruen and Y. S. Zheng (2001), "Coordination Mechanisms for a Distribution System with One Supplier and Multiple Retailers," *Management Science*, 47, 693-708.

Chen, X. and D. Simchi-Levi (2004), "Coordinating Inventory Control and Pricing Strategies with Random Demand and Fixed Ordering Cost: The Infinite Horizon Case," *Mathematics of Operations Research*, 29, 3, 698-723.

Curtat, L. (1996), "Markov Equilibria of Stochastic Games with Complementarities," *Games and Economic Behavior*, 17, 2, 177-199.

Deneckere, R., H. Marvel and J. Peck (1996), "Demand Uncertainty, Inventories, and Resale Price Maintenance," *The Quarterly Journal of Economics*, 111, 885-914.

Deneckere, R., H. Marvel and J. Peck (1997), "Demand Uncertainty and Price Maintenance: Markdowns as Destructive Competition," *American Economic Review*, 87, 4, 619-641.

Ericson, R. and A. Pakes (1995), "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies*, 62, 1, 53-82.

Ettl, M., G. Feigin, G. Lin and D. Yao (2000), "A Supply Network Model with Base-Stock Control and Service Requirements," *Operations Research*, 48, 2, 216-232.

Federgruen, A. and A. Heching (1999), "Combined Pricing and Inventory Control Under Uncertainty," *Operations Research*, 47, 3, 454-475.

Federgruen, A. and P. Zipkin (1986), "An Inventory Model with Limited Production Capacity and Uncertain Demands," *Mathematics of Operations Research*, 11, 193-215.

Fershtman, C. and A. Pakes (2000), "A Dynamic Oligopoly with Collusion and Price Wars," *RAND Journal of Economics*, 31, 2, 207-236.

Fudenberg, D. and J. Tirole (1991), *Game Theory*, The MIT Press, Cambridge, MA.

Graves, S. and S. Willems (2000), "Optimizing Strategic Safety Stock Placement in Supply Chains," *Manufacturing and Service Operations Management*, 2, 1, 68-83.

Jeuland, A. and S. Shugan (1983), "Managing Channel Profits," *Marketing Science*, 2 (Summer), 239-272.

Kashyap, A. (1995), "Sticky Prices: New Evidence from Retail Catalogs," *Quarterly Journal of Economics*, 254-270.

Lariviere, M. (1999), "Supply Chain Contracting and Coordination with Stochastic Demand," in *Quantitative Models for Supply Chain Management*, S. Tayur, M. Magazine and R. Ganeshan (Eds.), Kluwer, Dordrecht, Netherlands.

Leeflang, P., D.Wittink, M.Wedel and P.Naert (2000), *Building models for marketing decisions*, Kluwer Academic Publishers, Dordrecht/Boston/London.

Maskin, E. and J. Tirole (1988), "A Theory of Dynamic Oligopoly, i and ii," *Econometrica*, 56, 3, 549-570.

Mathewson, G. and R. Winter (1984), "An Economic Theory of Vertical Restraints," *Rand Journal of Economics*, 15, 27-38.

Milgrom, P. and J. Roberts (1990), "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58, 1255-1277.

Moorthy, K. (1987), "Managing Channel Profits: Comment," *Marketing Science*, 6, 375-379.

Nahmias, S. (2001), *Production and Operations Analysis*, McGraw-Hill, New York, NY.

Netessine, S. and N. Rudi (2003), "Centralized and Competitive Inventory Models with Demand Substitution," *Operations Rsearch*, 51, 329-335.

Padmanabhan, V. and I.P.L. Png (1995), "Returns Policies: Make Money by Making Good," *Sloan Management Review*, Fall, 65-72.

Padmanabhan, V. and I.P.L. Png (1997), "Manufacturer's Returns Policies and Retail Competition," *Marketing Science*, 16, 81-94.

Shapiro, C. (1989), "Theories of Oligopolies Behavior," Chapter 6 in *Handbook of Industrial Organization*, Vol. 1, Eds. R. Schmalensee and R. Willig., North Holland, Amsterdam.

Tirole, J. (1988), *The Theory of Industrial Organization*, The MIT Press, Cambridge, MA.

Tsay, A., S. Nahmias and N. Agrawal (1999), "Modeling Supply Chain Contracts: A Review," in *Quantitative Models for Supply Chain Management*, S. Tayur, M. Magazine and R. Ganeshan (Eds.), Kluwer, Dordrecht, Netherlands.

van Ryzin, G. and S. Mahajan (1999), "Retail Inventories and Consumer Choice," in *Quantitative Methods in Supply Chain Management*, S. Tayur, M. Magazine and R. Ganeshan (Eds.), Kluwer, Dordrecht, Netherlands.

Vives, X. (2000), *Oligopoly Pricing: Old Ideas and New Tools*, The MIT Press, Cambridge, Massachusetts.

Zipkin, P. (2000), *Foundations of Inventory Management*, McGraw-Hill, New York.

Appendix: Proofs

Proof of Theorem 1. (a) Under any given vector of constant wholesale prices w, the retailers face an infinite horizon non-cooperative game. At the beginning of the game, the retailers simultaneously select a price vector $p \in \Pi_{j=1}^N [p_j^{min}, p_j^{max}]$. Thereafter, each retailer i selects an infinite horizon (possibly history-dependent) replenishment strategy σ_i. Let $\sigma = (\sigma_1, ..., \sigma_N)$. Firm i's average profit Π_i depends on the price vector p and the N-tuple of strategies σ, i.e., $\Pi_i = \Pi_i(p, \sigma)$. In the infinite horizon game, a Nash equilibrium is a vector p^* and an N-tuple of inventory strategies σ^*, such that $\Pi_i(p_i, \sigma_i, p^*_{-i}, \sigma^*_{-i}) \leq \Pi_i(p^*, \sigma^*)$ for all $i = 1, ..., N$. As explained in Section 3, given a specific price vector p and service level vector f^0, each retailer i optimally adopts a base-stock inventory policy with base-stock level $y_i^*(p, f^0)$, and earns a long-run average profit of $\tilde{\pi}_i(p, f^0)$, see (12). This implies that the price vectors that are part of an infinite horizon Nash equilibrium are the same as those that are Nash equilibria in the single-stage game in which retailer i only selects a price $p_i \in [p_i^{min}, p_i^{max}]$ and earns profits $\tilde{\pi}_i(p, f^0)$. It therefore follows from Theorem 2(b), Theorem 7(a), and Theorem 9(a) in Bernstein and Federgruen (2004b) for the Attraction, Linear, and Log-Separable demand models (I)-(III), respectively, that the solution to the first order conditions

$$\frac{\partial \log \tilde{\pi}_i(p, f^0)}{\partial p_i} = 0 \tag{21}$$

is the unique Nash equilibrium in the single-stage game and hence the unique equilibrium prices chosen as part of an equilibrium strategy in the infinite horizon game, as well.

Writing p^I for $p^I(f^0)$, this implies that p^I is the unique set of equilibrium prices if and only if for all $i = 1, ..., N$, p_i^I satisfies the first order condition (21), i.e.,

$$\frac{\partial \log \tilde{\pi}_i(p_i^I, p_{-i}^I, f^0 | w_i^*)}{\partial p_i} = \frac{1}{p_i^I - w_i^* - k_i(f_i^0)} + \frac{\partial d_i(p^I, f^0)/\partial p_i}{d_i(p^I, f^0)} = 0. \tag{22}$$

For the demand models (I)-(III), it is easily verified that $\partial^2 \log \tilde{\pi}_i / \partial p_i \partial w_i > 0$. In other words, the function $\partial \log \tilde{\pi}_i / \partial p_i$ is strictly increasing in w_i. To show the existence of a unique coordinating wholesale price $w_i^* \leq p_i^I - k_i(f_i^0) \stackrel{def}{=} \overline{w}_i$, under which equation (22) is satisfied, it thus suffices to verify that

$$\lim_{w_i \searrow -\infty} \frac{\partial \log \tilde{\pi}_i(p_i^I, p_{-i}^I, f^0 | w_i)}{\partial p_i} = \frac{\partial d_i(p^I, f^0)/\partial p_i}{d_i(p^I, f^0)} < 0, \text{ while } \lim_{w_i \nearrow \overline{w}_i} \frac{\partial \log \tilde{\pi}_i(p_i^I, p_{-i}^I, f^0 | w_i)}{\partial p_i} > 0.$$

Finally, the identity for $w_i^*(f^0)$ is immediate from (22).

(b) Define $w_i^0 = c_i + \partial\tilde{C}_0/\partial d_i$. Note that

$$\frac{\partial\tilde{\pi}_I(p^I, f^0)}{\partial p_i} = d_i + \left(p_i^I - c_i - k_i(f_i^0) - \frac{\partial\tilde{C}_0}{\partial d_i}\right)\frac{\partial d_i}{\partial p_i} + \sum_{j\neq i}\left(p_j - c_j - k_j(f_j^0) - \frac{\partial\tilde{C}_0}{\partial d_j}\right)\frac{\partial d_j}{\partial p_i} = 0.$$

It then follows from our assumption and $\dfrac{\partial d_j}{\partial p_i} \geq 0$, that $d_i + \left(p_i^I - c_i - k_i(f_i^0) - \dfrac{\partial\tilde{C}_0}{\partial d_i}\right)\dfrac{\partial d_i}{\partial p_i} < 0$.

Therefore, $\dfrac{\partial\tilde{\pi}_i(p^I, f^0|w_i = w_i^0)}{\partial p_i} = d_i + \left(p_i^I - c_i - k_i(f_i^0) - \dfrac{\partial\tilde{C}_0}{\partial d_i}\right)\dfrac{\partial d_i}{\partial p_i} < 0$, and hence

$$\frac{\partial \log\tilde{\pi}_i(p^I, f^0|w_i = w_i^0)}{\partial p_i} < 0.$$

The proof of part (a) shows that $w_i^*(f^0) \geq w_i^0$. ■

Proof of Proposition 1. Rewrite (22) as:

$$\frac{1}{p_i^0 - w_i^* - k_i(f_i^0)} + \frac{\partial\tilde{d}_i(p_i^0, f_i^0)}{\partial p_i} = 0,\ i = 1, ..., N, \tag{23}$$

where $\tilde{d}_i = \log d_i$. It follows from the Implicit Function Theorem that the coordinating wholesale prices are differentiable functions of the service levels, with the matrix

$$\left(\frac{\partial w_i^*}{\partial f_j}\right)_{i,j=1}^N = \text{ diag }(-[p_1^0 - w_1^* - k_1(f_1^0)]^2, ..., -[p_N^0 - w_N^* - k_N(f_N^0)]^2)\times$$

$$\times\left(\frac{k_i'(f_i^0)}{[p_i^0 - w_i^* - k_i(f_i^0)]^2}\delta_{ij} + \frac{\partial^2\tilde{d}_i(p^0, f^0)}{\partial p_i\partial f_j}\right)_{i,j=1}^N.$$

Here, diag $(\Delta_1, ..., \Delta_N)$ denotes a diagonal matrix, Δ_i the i-th diagonal element and δ_{ij} the Kronecker delta, i.e., $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. Part (a) then follows therefore directly from the observation that $\frac{\partial^2\tilde{d}_i}{\partial p_i\partial f_j} \leq 0$ in each of the models (I), (II) and (III). This is immediate in the Log-Separable model where this second-order partial derivative is zero. In the Linear model, $\frac{\partial^2\tilde{d}_i}{\partial p_i\partial f_j} = -\frac{b_i\gamma_{ij}}{d_i^2} \leq 0$ and in the Attraction model $\frac{\partial^2\tilde{d}_i}{\partial p_i\partial f_j} = \frac{\partial a_i/\partial p_i\times\partial a_j/\partial f_j}{(\sum_{l=1}^N a_l)^2} \leq 0$, by (4). For part (b), note that $\frac{\partial w_i^*}{\partial f_i^0} = -k_i'(f_i^0) - [p_i^0 - w_i^* - k_i(f_i^0)]^2\frac{\partial^2\tilde{d}_i(p^0,f^0)}{\partial p_i\partial f_i}$ and $k_i'(\cdot) \geq 0$ imply that $\frac{\partial w_i^*}{\partial f_i^0} < 0$ in the Linear and Log-Separable models where $\frac{\partial^2\tilde{d}_i}{\partial p_i\partial f_i} = \frac{b_i\beta_i}{d_i^2} \geq 0$ and $\frac{\partial^2\tilde{d}_i}{\partial p_i\partial f_i} = 0$, respectively. In the Attraction model, $\frac{\partial^2\tilde{d}_i}{\partial p_i\partial f_i} < 0$ may occur. ■

Proof of Theorem 2. Similar to the proof of Theorem 1, under any given pricing scheme by the supplier, the retailers face an infinite horizon game. In this case, the retailers select simultaneously at the beginning of the game, a vector of service levels $f \in [0, 1)^N$ along with a price vector $p \in \Pi_{j=1}^N[p_j^{min}, p_j^{max}]$. Thereafter, each retailer i again selects an infinite horizon replenishment strategy σ_i. Firm i's average profit per period, Π_i, now depends on the price vector p, the service level vector f and the N-tuple of strategies σ, i.e., $\Pi_i = \Pi_i(p, f, \sigma)$. In the infinite horizon game, a Nash equilibrium is now a triple (p^*, f^*, σ^*) such that $\Pi_i(p_i, f_i, \sigma_i, p^*_{-i}, f^*_{-i}, \sigma^*_{-i}) \leq \Pi_i(p^*, f^*, \sigma^*)$ for all $i = 1, ..., N$. Following the arguments in Theorem 1, one observes that the pairs (p^*, f^*) (of a price vector and a service level vector) that are part of an infinite horizon Nash equilibrium are the same as those that are Nash equilibria in the single stage game in which retailer i only selects a service level $f_i \in [0, 1)$ along with a price $p \in [\max\{p_i^{min}, w_i + k_i(f_i)\}, p_i^{max}]$ and in which it receives a profit $\tilde{\pi}_i(p, f)$. We therefore again first establish that a solution (p^*, f^*) to the first order conditions:

$$\frac{\partial \tilde{\pi}_i}{\partial p_i}(p_i, f_i) = 0 \text{ and } \frac{\partial \tilde{\pi}_i}{\partial f_i}(p_i, f_i) = 0, \quad i = 1, ..., N, \tag{24}$$

is a Nash equilibrium in this single stage game and is hence part of an equilibrium strategy of the infinite horizon game. For the Attraction and Linear models, this result follows from Theorem 3(a) and Theorem 8 in Bernstein and Federgruen (2004b), respectively. As for the Log-Separable demand functions, the result is obtained by showing that for all $i = 1, ..., N$, $\log \tilde{\pi}_i = \log[p_i - w_i - k_i(f_i)] + \log q_i(p) + \log \psi_i(f)$ is jointly concave in (p_i, f_i). Joint concavity of the second and third terms is immediate from (7) – (9); that of the first term follows from the fact that the margin function $(p_i - w_i - k_i(f_i))$ is jointly concave in (p_i, f_i) by Lemma 1, and therefore log-concave. Note from (20) that $f_i^I \geq \frac{h_i^{-*}}{h_i^+ + h_i^{-*}}$. Then, by Lemma 1 and the definition of h_i^{-*},

$$k_i^{*'}(f_i^I) = \frac{(h_i^+ + h_i^{-*})f_i^I - h_i^{-*}}{g_i(G_i^{-1}(f_i^I))} = \gamma_i(p^I, f^I), \ i = 1, ..., N. \tag{25}$$

Next, note that the first order conditions in (24) are equivalent to:

$$\frac{\partial \log \tilde{\pi}_i}{\partial p_i} = \frac{1}{p_i - w_i - k_i(f_i)} + \frac{\partial d_i / \partial p_i}{d_i(p, f)} = \frac{1}{p_i - w_i - k_i(f_i)} - \frac{\eta_i(p, f)}{p_i} = 0, \ i = 1, ..., N, \tag{26}$$

$$k_i'(f_i)\frac{\partial d_i}{\partial p_i} + \frac{\partial d_i}{\partial f_i} = 0, \ i = 1, ..., N \tag{27}$$

Straightforward algebra verifies that under the backlogging penalties h^{-*} and wholesale prices w^*, given by (20) and (19), the centralized solution (p^I, f^I) satisfies this pair of equations. It thus remains to show that for all $i = 1, ..., N$, $h_i^{-*} + h_i^+ \geq 0$, ensuring that the functions $\{k_i^*(\cdot)\}$ are convex and hence that each function $\log \tilde{\pi}_i(p, f)$ is jointly concave in (p_i, f_i). But $h_i^{-*} + h_i^+ \geq 0$ is immediate by adding h_i^+ to both sides of (20) and using (18). ■

Proof of Proposition 2. In the MNL case, the equilibrium service level vector f^* under a given wholesale price vector w and backlogging rate cost vector h^- satisfies the first order condition (27): $-k_i'(f_i^*)\alpha_i + b_i'(f_i^*) = 0,\ i = 1, ..., N$. From Lemma 1, $k_i'(f_i^*) = \left((h_i^+ + h_i^-)f_i^* - h_i^-\right)/g_i(G_i^{-1}(f_i^*))$. On the other hand, under the coordinating scheme in this model, retailer i is charged the wholesale price w_i^* and a penalty $h_i^{-*} - h_i^- = h_i^+ \frac{f_i^I}{1-f_i^I} - \frac{g_i(G_i^{-1}(f_i^I))}{1-f_i^I}\frac{b_i'(f_i^I)}{\alpha_i} - h_i^-$. Then, $f_i^* < [>] f_i^I$ implies that $-k_i'(f_i^I)\alpha_i + b_i'(f_i^I) < [>] - k_i'(f_i^*)\alpha_i + b_i'(f_i^*) = 0$, since $k_i(\cdot)$ is convex and $b_i(\cdot)$ is concave. This, in turn, implies that $h_i^{-*} - h_i^- > [<] \frac{h_i^+ f_i^I - g_i(G_i^{-1}(f_i^I))k_i'(f_i^I)}{1-f_i^I} - h_i^- = 0$. A similar argument applies to the Linear Model. ■