PART II

Supply Chain Coordination

A.G. de Kok and S.C. Graves, Eds., *Handbooks in OR & MS, Vol. 11*
© 2003 Elsevier B.V. All rights reserved.

Chapter 6

Supply Chain Coordination with Contracts

Gérard P. Cachon

The Wharton School of Business, University of Pennsylvania, Philadelphia, PA 19104, USA
E-mail: cachon@wharton.upenn.edu

1 Introduction

Optimal supply chain performance requires the execution of a precise set of actions. Unfortunately, those actions are not always in the best interest of the members in the supply chain, i.e., the supply chain members are primarily concerned with optimizing their own objectives, and that self-serving focus often results in poor performance. However, optimal performance is achievable if the firms coordinate by contracting on a set of transfer payments such that each firm's objective becomes aligned with the supply chain's objective.

This chapter reviews and extends the supply chain literature on the management of incentive conflicts with contracts. Numerous supply chain models are discussed, roughly presented in order of increasing complexity. In each model the supply chain optimal actions are identified. In each case the firms could implement those actions, i.e., each firm has access to the information required to determine the optimal actions and the optimal actions are feasible for each firm.[1] However, firms lack the incentive to implement those actions. To create that incentive the firms can adjust their terms of trade via a contract that establishes a transfer payment scheme. A number of different contract types are identified and their benefits and drawbacks are illustrated.

[1] Even in the asymmetric information models there is an assumption that the firms can share information so that all firms are able to evaluate the optimal policies. Nevertheless, firms are not required to share information. See Anand and Mendelson (1997) for a model in which firms are unable to share information even though they have the incentive to do so.

The first model has a single supplier selling to a single retailer that faces the newsvendor problem. In that model the retailer orders a single product from the supplier well in advance of a selling season with stochastic demand. The supplier produces after receiving the retailer's order and delivers her production to the retailer at the start of the selling season.[2] The retailer has no additional replenishment opportunity. How much the retailer chooses to order depends on the terms of trade, i.e., the contract, between the retailer and the supplier.

The newsvendor model is not complex, but it is sufficiently rich to study three important questions in supply chain coordination. First, which contracts coordinate the supply chain? A contract is said to coordinate the supply chain if the set of supply chain optimal actions is a Nash equilibrium, i.e., no firm has a profitable unilateral deviation from the set of supply chain optimal actions. Ideally, the optimal actions should also be a unique Nash equilibrium; otherwise, the firms may 'coordinate' on a suboptimal set of actions. In the newsvendor model the action to coordinate is the retailer's order quantity (and in some cases, as is discussed later, the supplier's production quantity also needs coordination). Second, which contracts have sufficient flexibility (by adjusting parameters) to allow for any division of the supply chain's profit among the firms? If a coordinating contract can allocate rents arbitrarily, then there always exists a contract that Pareto dominates a noncoordinating contract, i.e., each firm's profit is no worse off and at least one firm is strictly better off with the coordinating contract. Third, which contracts are worth adopting? Although coordination and flexible rent allocation are desirable features, contracts with those properties tend to be costly to administer. As a result, the contract designer may actually prefer to offer a simple contract even if that contract does not optimize the supply chain's performance. A simple contract is particularly desirable if the contract's efficiency is high (the ratio of supply chain profit with the contract to the supply chain's optimal profit) and if the contract designer captures the lion's share of supply chain profit.

Section 3 extends the newsvendor model by allowing the retailer to choose his retail price in addition to his stocking quantity. Coordination is more complex in this setting because the incentives provided to align one action (e.g., the order quantity) might cause distortions with the other action (e.g., the price). Not surprising, it is shown that some of the contracts that coordinate the basic newsvendor model no longer coordinate in this setting, whereas others continue to do so.

Section 4 extends the newsvendor model by allowing the retailer to exert costly effort to increase demand. Coordination is challenging because the

[2] The author adopts the convention (first suggested to him by Martin Lariviere) that the firm offering the contract is female and the accepting firm is male. When neither firm offers the contract, then the upstream firm is female, and the downstream firm is male.

retailer's effort is noncontractible, i.e., the firms cannot write contracts based on the effort chosen (for reasons discussed later). Furthermore, as with the retail price, coordination is complicated by the fact that the incentives to align the retailer's order-quantity decision may distort the retailer's effort decision.

Section 5 studies two models, each with one supplier that sells to multiple competing retailers. Coordination requires the alignment of multiple actions implemented by multiple firms, in contrast with the price and effort models (Sections 3 and 4) that have multiple actions implemented by a single firm (the retailer). More specifically, coordination requires the tempering of downstream competition.

Section 6 has a single retailer that faces stochastic demand but two replenishment opportunities. Early production (the first replenishment) is cheaper than later production (the second replenishment), but less informative because the demand forecast is updated before the second replenishment. Coordination requires that the retailer be given the proper incentives to balance this trade-off.

Section 7 studies an infinite horizon stochastic demand model in which the retailer receives replenishments from a supplier after a constant lead time; a departure from the single-period lost-sales models of the previous sections. As in the effort model, coordination requires that the retailer chooses a 'higher action', which in this model is a larger base-stock level. The cost of this higher action is more inventory on average, but unlike in the effort model, the supplier can verify the retailer's inventory and therefore share the holding cost of carrying more inventory with the retailer.

Section 8 adds richness to the single-location base-stock model by making the supplier hold inventory, albeit at a lower holding cost than the retailer. Whereas the focus in the previous sections is primarily on coordinating the downstream actions, in this model the supplier's action also requires coordination, and that coordination is nontrivial. To be more specific, in the single-location model the only critical issue is the amount of inventory in the supply chain, but here the allocation of the supply chain's inventory between the supplier and the retailer is important as well.

Section 9 departs from the assumption that firms agree to contracts with set transfer prices. In many supply chains the firms agree to a contractual arrangement before the realization of some relevant information. The firms could specify transfer payments for every possible contingency, but those contracts are quite complex. Instead, firms could agree to set transfer prices via an internal market after the relevant information is revealed.

Section 10 endows one firm with important information that the other firm does not possess, i.e., it is private information. For example, a manufacturer may have a more accurate demand forecast for a product than the manufacturer's supplier. As in the previous models, supply chain coordination requires each firm to implement optimal actions. But since those optimal actions depend on the private information, supply chain coordination also requires the accurate sharing of information. Sharing information is challenging because there exists the incentive to provide false information in an effort to influence the actions taken, e.g., a manufacturer may wish to offer a rosy demand forecast to try to get the supplier to build more capacity.

The final section summarizes the main insights that have developed from this literature and provides some general guidance for future research.

Each section presents one or more simple models to facilitate the analysis and to highlight the potential incentive conflicts within a supply chain. The same analysis recipe is usually followed: identify the type of contracts that can coordinate the supply chain, determine for each contract type the set of parameters that achieves coordination, and evaluate for each coordinating contract type the possible range of profit allocations, i.e., what fraction of the supply chain's profit can be earned by each member in the supply chain with a coordinating contract. Implementation issues are then explored: e.g., is a contract-type compliant with legal restrictions; what are the consequences for failing to comply with the contractual terms; and what is a contract's administrative burden (e.g., what types of data need to be collected and how often must data be collected). Each section ends with a discussion of extensions and related research.

While this chapter gives a broad treatment of the supply chain contracting literature, it does not address all papers that could possibly be classified within this literature. In particular, there are (at least) six types of closely related papers that are not discussed directly. The first is the extensive literature on quantity discounts because several excellent reviews are available, see Dolan and Frey (1987) and Boyaci and Gallego (1997). The second set that is not addressed includes papers on a single firm's optimal procurement decisions given particular contractual terms. Examples include Scheller-Wolf and Tayur's (1997) study of procurement under a minimum quantity-commitment contract, Duenyas, Hopp and Bassok's (1997) study of procurement with JIT contracts, Bassok and Anupindi's (1997a) study of procurement with total minimum commitments, and Anupindi and Akella's (1993) and Moinzadeh and Nahmias' (2000) studies of procurement with standing order contracts. The third body of excluded work is research on supply chain coordination without contracts. Examples include papers on the benefit of Quick Response (Iyer & Bergen, 1997), Accurate Response (Fisher & Raman, 1996), collaborative planning and forecasting (Aviv, 2001), Vendor Managed Inventory (Aviv & Federgruen, 1998) and information sharing within a supply chain (Gavirneni, Kapuscinski & Tayur, 1999). Fourth, papers on decentralized supply chain operations which do not explicitly consider coordination are excluded: e.g., Cachon and Lariviere (1997, 1999), Corbett and Karmarkar (2001), Erhun, Keskinocak and Tayur (2000), Ha, Li and Ng (2000) and Majumder and Groenevelt (2001). Fifth, the broad literature on franchising is not directly discussed, primarily because that literature generally avoids operational detail (see Lafontaine & Slade, 2001 for a recent review of that literature) Finally, papers on vertical restraints vis-a-vis social welfare and antitrust issues are not considered, see Katz (1989).

For earlier overviews on supply chain coordination with contracts, see Whang (1995) and the three chapters in Tayur, Ganeshan and Magazine (1998) that focus on the topic: Cachon (1998), Lariviere (1998) and Tsay, Nahmias and Agrawal (1998).

2 Coordinating the newsvendor

This section studies coordination in a supply chain with one supplier and one retailer. There is one selling season with stochastic demand and a single opportunity for the retailer to order inventory from the supplier before the selling season begins. With the standard wholesale-price contract, it is shown that the retailer does not order enough inventory to maximize the supply chain's total profit because the retailer ignores the impact of his action on the supplier's profit. Hence, coordination requires that the retailer be given an incentive to increase his order.

Several different contract types are shown to coordinate this supply chain and arbitrarily divide its profit: buyback contracts, revenue-sharing contracts, quantity-flexibility contracts, sales-rebate contracts and quantity-discount contracts.

The concept of a compliance regime is introduced. The compliance regime determines the consequences for failing to adhere to a contract. For example, it is assumed that the supplier cannot force the retailer to accept more product than the retailer orders, i.e., the retailers could clearly use the courts to prevent any attempt to do so. However, it is debatable whether the supplier is required to deliver the retailer's entire order. The compliance regime matters because it influences the kinds of contracts that coordinate the supply chain: there exist contracts that coordinate with one compliance regime, but not another.

2.1 Model and analysis

In this model there are two firms, a supplier and a retailer. The retailer faces the newsvendor's problem: the retailer must choose an order quantity before the start of a single selling season that has stochastic demand. Let $D > 0$ be demand during the selling season. Let F be the distribution function of demand and f its density function: F is differentiable, strictly increasing and $F(0) = 0$. Let $\overline{F}(x) = 1 - F(x)$ and $\mu = E[D]$. The retail price is p. The supplier's production cost per unit is c_s and the retailer's marginal cost per unit is c_r, $c_s + c_r < p$. The retailer's marginal cost is incurred upon procuring a unit (rather than upon selling a unit). For each demand the retailer does not satisfy the retailer incurs a goodwill penalty cost g_r and the analogous cost for the supplier is g_s. For notational convenience, let $c = c_s + c_r$ and $g = g_s + g_r$. The retailer earns $v < c$ per unit unsold at the end of season, where v is net of any salvage expenses. Assume the supplier's net salvage value is no greater than v, so it is optimal for the supply chain to salvage leftover inventory at the retailer. The qualitative insights from the subsequent analysis do not depend on whether it is optimal for the retailer or the supplier to salvage leftover inventory. [The supply chain contracting literature generally avoids this issue by assuming the net salvage value of a unit is the same at either firm. Tsay (2001) is an exception.] For more extensive treatment of the newsvendor model, see Silver, Pyke and Peterson (1998) or Nahmias (1993).

The following sequence of events occurs in this game: the supplier offers the retailer a contract; the retailer accepts or rejects the contract; assuming the retailer accepts the contract, the retailer submits an order quantity, q, to the supplier; the supplier produces and delivers to the retailer before the selling season; season demand occurs; and finally transfer payments are made between the firms based upon the agreed contract. If the retailer rejects the contract, the game ends and each firm earns a default payoff.

The supplier is assigned to make the contract offer, rather than the retailer, only for expositional convenience, i.e., it has no impact on the subsequent analysis. The firm that offers the contract does not matter because we seek to identify the set of contracts that coordinate the supply chain and arbitrarily allocate its profit. If one firm were indeed assigned to make the only offer, then it would offer the most favorable contract in that set which the other firm will accept. Furthermore, it is unlikely in practice that either firm makes a single offer which is regarded as the final offer. Instead, firms are likely to make many offers and counter offers before they settle on some agreement. The details of this negotiation process are generally not considered in the supply chain literature, nor are they explored here.

The contract that is actually adopted at the end of the negotiation process depends on the firms' relative bargaining power, which is a concept that is easy to understand but difficult to quantify. Power, like beauty, can be in 'the eye of the beholder', or it can be more concrete. A standard approach to model power is to assume one of the firms has an exogenous reservation profit level, i.e., the firm accepts only a contract that yields that reservation level: the higher the reservation level, the higher the firm's power.[3] Ertogral and Wu (2001) are even more explicit with their bargaining process: bargaining occurs in rounds in which either firm may

[3] Webster and Weng (2000) impose a stronger condition. They require that both firms are at least as well off with the adoption, a contract as they would be with a default contract for all realization of demand.

make an offer, but if at the end of a round an offer is not accepted there is a fixed probability the negotiations fail, i.e., the firms are left with their reservation profit. However, the reservation level approach is not entirely satisfying: it is quite likely a firm's opportunity outside of the relationship being studied is not independent of the firm's opportunity within the relationship. Nor should it be expected that the value of a firm's outside opportunity is known with certainty a priori (van Mieghem, 1999; Rochet & Stole, 2002). Aside from the reservation level approach, some researchers adjust power by changing which firm makes the contract offer or by changing when actions are chosen. In general, a firm has more power when she makes the first offer, assuming it is a 'take-it or leave-it' offer, or when she chooses her actions first, assuming she is committed to her action. These choices matter when one wants to predict with precision the particular outcome of a negotiation process, which is not done here. Additional research is surely needed on this issue.

To continue with the description of the model, each firm is risk neutral, so each firm maximizes expected profit. There is full information, which means that both firms have the same information at the start of the game, i.e., each firm knows all costs, parameters and rules. Game theorists have been also concerned with higher levels of common knowledge: e.g., does firm A know that firm B knows all information and does firm B know that firm A knows that firm B knows all information, etc. The supply chain contracting literature has not explored this issue. See Rubinstein (1989) for a model with counterintuitive implications for less than complete common knowledge.

It is quite reasonable to assume the supplier cannot force the retailer to pay for units delivered in excess of the retailer's order quantity. But can the supplier deliver *less* than the amount the retailer orders? A failure to deliver the retailer's full order may occur for a number of reasons beyond the supplier's control: e.g., unforeseen production difficulties or supply shortages for key components. The shortage may also be due to self-interest. In recognition of that motivation, the retailer could assume the supplier operates under *voluntary compliance*, which means the supplier delivers the amount (not to exceed the retailer's order) that maximizes her profit given the terms of the contract. Alternatively, the retailer could believe the supplier never chooses to deliver less than the retailer's order because the consequences for doing so are sufficiently great, e.g., court action or a loss of reputation. Call that regime *forced compliance*.

The compliance regime in reality almost surely falls somewhere between those two extremes. However, in any regime other than forced compliance the supplier can be expected to fall somewhat short on her delivery *if* the terms of the contract give the supplier an incentive to do so. In other words, any contract that coordinates the supply chain with voluntary compliance surely coordinates with forced compliance, but the reverse is not true (because the contract may fail to coordinate the supplier's action). Hence, voluntary compliance is the more conservative assumption (albeit maybe too conservative).[4]

The approach taken in this section is to assume forced compliance but to check if the supplier has an incentive to deviate from the proposed contractual terms. This seemingly contradictory stance is adopted to simplify notation: voluntary compliance requires notation to keep track of two actions, the retailer's order quantity and the supplier's production quantity, whereas forced compliance requires notation only for one action. See Cachon and Lariviere (2001) for additional discussion on compliance regimes.[5]

Let $S(q)$ be expected sales, $\min(q, D)$,c

$$S(q) = q(1 - F(q)) + \int_0^q yf(y)\,dy$$
$$= q - \int_0^q F(y)\,dy$$

(The above follows from integration by parts.) Let $I(q)$ be the expected leftover inventory, $I(q) = (q - D)^+ = q - S(q)$. Let $L(q)$ be the lost-sales function, $L(q) = (D - q)^+ = \mu - S(q)$. Let T be the expected transfer payment from the retailer to the supplier. That function may depend on a number of observations (e.g., order quantity, leftover inventory), as is seen later.

The retailer's profit function is

$$\pi_r(q) = pS(q) + vI(q) - g_r L(q) - c_r q - T$$
$$= (p - v + g_r)S(q) - (c_r - v)q - g_r\mu - T,$$

the supplier's profit function is

$$\pi_s(q) = g_s S(q) - c_s q - g_s\mu + T,$$

[4] This chapter assumes the wholesale price operates with forced compliance whereas the quantity to deliver may operate with voluntary compliance. Hence, the parameters in a contract can operate under different compliance regimes, which can be justified by the differences in ease by which the courts can verify different terms. As suggested by Fangruo Chen, it is also possible to view all contracts as iron clad contracts (i.e., everything operates with forced compliance), but the kinds of contractual terms may be limited. For example, suppose the contract were written such that the retailer's order quantity is an upper bound on the supplier's delivery quantity, i.e., forced compliance of an upper bound is analogous to our voluntary compliance with a specific quantity. Additional research is needed to determine if the distinctions in these interpretations matter.

[5] See Krasa and Villamil (2000) for a model in which the contracting parties endogenously set the compliance regime. Milner and Pinker (2001) do not explicitly define a compliance regime, but it does impact their results. They show supply chain coordination is possible when one firm is able to identify any deviation by the other firm and follow through with substantial penalties. When deviations cannot be identified for sure, supply chain coordination is no longer possible. Baiman, Fischer and Rajan (2000) focus on how the compliance regime impacts a supplier's incentive to improve quality and a buyer's incentive to inspect.

and the supply chain's profit function is

$$\Pi(q) = \pi_r(q) + \pi_s(q) = (p - v + g)S(q) - (c - v)q - g\mu. \tag{2.1}$$

Given the above profit functions, it is possible to normalize some of the variables. For example, let $\hat{p} = p - v + g_r$ be the adjusted price, let $\hat{c} = c - v$ be the adjusted production cost, let $\hat{T} = T + (c_r - v)q$ be the adjusted transfer payment, let $\hat{\pi}_r(q) = \pi_r(q) + g_r\mu$ be the retailer's adjusted profit function and let $\hat{\pi}_s(q) = \pi_s(q) + g_s\mu$ be the supplier's adjusted profit function. Those adjusted profit functions simplify to $\hat{\pi}_r(q) = \hat{p}S(q) - \hat{T}$ and $\hat{\pi}_s(q) = g_s S(q) - \hat{c}q + \hat{T}$. While those functions have cleaner notation, caution is required when defining the transfer payment for a given contract because the contract's terms (e.g., the wholesale price, the buyback rate, etc.) must be given in terms of the adjusted parameters. Unfortunately, the notational clarity gained by these adjustments is often lost when the adjusted contract terms are included. As a result, this chapter works with the unadjusted profit functions.

Let q^o be a supply chain optimal order quantity, i.e., $q^o = \arg\max \Pi(q)$. To avoid uninteresting situations, assume $\Pi(q^o) > 0$. Since F is strictly increasing, Π is strictly concave and the optimal order quantity is unique. Further, q^o satisfies

$$S'(q^o) = \overline{F}(q^o) = \frac{c - v}{p - v + g}. \tag{2.2}$$

Let q_r^* be the retailer's optimal order quantity, i.e., $q_r^* = \arg\max \pi_r(q)$. The retailer's order clearly depends on the chosen transfer payment scheme, T.

A number of contract types have been applied to this model. The simplest is the wholesale-price contract: the supplier merely charges the retailer a fixed wholesale price per unit ordered. Section 2.2 studies that contract. It is shown that the wholesale-price contract generally does not coordinate the supply chain. Hence, the analysis concentrates on two questions: what is the efficiency of the wholesale-price contract (the ratio of supply chain profit to optimal profit) and what is the supplier's share of the supply chain's profit.

More complex contracts include a wholesale price plus some adjustment that typically depends on realized demand (the quantity-discount contract is an exception). As mentioned in Section 1, the analysis recipe for all of those contracts is the same: determine the set of contract parameters that coordinate the retailer's action; then evaluate the possible range of profit allocations between the firms; and then check whether the contract coordinates under voluntary compliance, i.e., whether the supplier has an incentive to deliver less than the retailer's order quantity.

2.2 The wholesale-price contract

With a wholesale-price contract the supplier charges the retailer w per unit purchased: $T_w(q,w)=wq$. See Lariviere and Porteus (2001) for a more complete analysis of this contract in the context of the newsvendor problem. Bresnahan and Reiss (1985) study the wholesale-price contract with deterministic demand.

$\pi_r(q,w)$ is strictly concave in q, so the retailer's unique optimal order quantity satisfies

$$(p-v+g_r)S'(q_r^*)-(w+c_r-v)=0. \tag{2.3}$$

Since $S'(q)$ is decreasing, $q_r^*=q^o$ only when

$$w=\left(\frac{p-v+g_r}{p-v+g}\right)(c-v)-(c_r-v).$$

It is straightforward to confirm that $w\le c_s$, i.e., the wholesale-price contract coordinates the channel only if the supplier earns a nonpositive profit. So the supplier clearly prefers a higher wholesale price. As a result, the wholesale-price contract is generally not considered a coordinating contract. [As discussed in Cho and Gerchak (2001) and Bernstein, Chen and Federgruen (2002), marginal cost pricing does not necessarily lead to zero profit for the supplier when the marginal cost is not constant.] Spengler (1950) was the first to identify the problem of 'double marginalization'; in this serial supply chain there is coordination failure because there are two margins and neither firm considers the entire supply chain's margin when making a decision.

Even though the wholesale-price contract does not coordinate the supply chain, the wholesale-price contract is worth studying because it is commonly observed in practice. That fact alone suggests it has redeeming qualities. For instance, the wholesale-price contract is simple to administer. As a result, a supplier may prefer the wholesale-price contract over a coordinating contract if the additional administrative burden associated with the coordinating contract exceeds the supplier's potential profit increase.

From Eq. (2.3) the retailer's optimal order quantity satisfies

$$F(q_r^*)=1-\frac{w+c_r-v}{p-v+g_r}.$$

Since F is strictly increasing and continuous there is a one-for-one mapping between w and q_r^*. Hence, let $w(q)$ be the unique wholesale price that induces the retailer to order q_r^* units,

$$w(q)=(p-v+g_r)\overline{F}(q)-(c_r-v).$$

The supplier's profit function can now be written as

$$\pi_s(q, w(q)) = g_s S(q) + (w(q) - c_s)q - g_s \mu. \tag{2.4}$$

It is immediately apparent that the compliance regime does not matter with this contract: for a fixed wholesale price no less than c_s the supplier's profit is nondecreasing in q, so the supplier surely produces and delivers whatever quantity the retailer orders.

The supplier's marginal profit is

$$\begin{aligned}\frac{\partial \pi_s(q, w(q))}{\partial q} &= g_s S'(q) + w(q) - c_s + w'(q)q \\ &= (p - v + g_r)\overline{F}(q)\left(1 + \frac{g_s}{p - v + g_r} - \frac{qf(q)}{\overline{F}(q)}\right) - (c - v).\end{aligned}$$

The supplier's profit function is unimodal if the above is decreasing. $\overline{F}(q)$ is decreasing, so $\pi_s(q, w(q))$ is decreasing in q if $qf(q)/\overline{F}(q)$ is increasing. Demand distributions with that property are called increasing generalized failure rate (IGFR) distributions.[6] Fortunately, many of the commonly applied demand distributions are IGFR: the normal, the exponential, the Weibull, the gamma and the power distribution. Thus, with an IGFR demand distribution there is a unique sales quantity, q_s^*, that maximizes the supplier's profit. (Actually, the supplier sets the wholesale price to $w(q_s^*)$ knowing quite well the retailer then orders q_s^* units.)

While $\pi_s(q_s^*, w(q_s^*))$ is the best the supplier can hope for, the retailer may actually insist on more than $\pi_r(q_s^*, w(q_s^*))$. For example, the retailer may earn more by selling some other product in his store, i.e., his opportunity cost is greater than $\pi_r(q_s^*, w(q_s^*))$. In that case the supplier needs to offer the retailer more generous terms to get the retailer to carry the product. With a wholesale-price contract the retailer's profit is increasing in q,

$$\frac{\partial \pi_r(q, w(q))}{\partial q} = -w'(q)q = (p - v + g_r)f(q)q > 0,$$

so the supplier can increase the retailer's profit by reducing her wholesale price (which should surprise no one). As long as the retailer insists on less than the supply chain optimal profit, the retailer's minimum profit requirement actually increases the total supply chain profit: the supply chain's profit is increasing in q for $q \in [q_s^*, q^o]$ and so is the retailer's profit. Hence, an

[6] The failure rate of a demand distribution is $f(x)/\overline{F}(x)$. Any demand distribution with an increasing failure rate (IFR) is clearly also IGFR. However, there are IGFR distributions that are not IFR. See Lariviere and Porteus (2001) for additional discussion.

increase in retail power can actually improve supply chain performance (which is somewhat surprising, and controversial). However, that improvement comes about at the supplier's expense. See Messinger and Narasimhan (1995), Ailawadi (2001) and Bloom and Perry (2001) for additional discussion and empirical evidence on how power has changed in several retail markets.

The two performance measures applied to the wholesale-price contract are the efficiency of the contract, $\Pi(q_s^*)/\Pi(q^o)$ and the supplier's profit share, $\pi_s(q_s^*, w(q_s^*))/\Pi(q_s^*)$. From the supplier's perspective the wholesale-price contract is an attractive option if both of those measures are high: the product of those ratios is the supplier's share of the supply chain's optimal profit,

$$\frac{\pi_s(q_s^*, w(q_s^*))}{\Pi(q^o)} = \left(\frac{\pi_s(q_s^*, w(q_s^*))}{\Pi(q_s^*)}\right)\left(\frac{\Pi(q_s^*)}{\Pi(q^o)}\right)$$

To illustrate when that is likely, suppose $g_r = g_s = 0$ and demand follows the power distribution: $F(q) = g^k$ for $k > 0$ and $q \in [0, 1]$. In that case the efficiency of the wholesale-price contract is $(k+1)^{-(1+1/k)}(k+2)$, the supplier's profit share is $(k+1)/(k+2)$ and the coefficient of variation is $(k(k+2))^{-1/2}$. (See Lariviere & Porteus, 2001 for details.) Note that the coefficient of variation is decreasing in k but both measures are increasing in k. In fact, as $k \to \infty$ the coefficient of variation approaches zero and both measures approach 1. Nevertheless, Table 1 demonstrates the supplier's share of supply chain profit increases more quickly than supply chain efficiency.

One explanation for this pattern is that the retailer's profit represents compensation for bearing risk: with the wholesale-price contract there is no variation in the supplier's profit, but the retailer's profit varies with the realization of demand. As the coefficient of variation decreases the retailer faces less demand risk and therefore his compensation is reduced. However, the retailer is not compensated due to risk aversion. (See Tsay, 2002 for a model with risk aversion.) If the retailer were risk averse, the supplier would have to provide for yet more compensation. Instead, the retailer is being compensated for the risk that demand and supply do not match. Lariviere and Porteus (2001) demonstrate this argument holds for a broad set of demand distributions.

Anupindi and Bassok (1999) study an interesting extension to this model. Suppose the supplier sells to a retailer that faces an infinite succession of identical selling seasons.[7] There is a holding cost on leftover inventory at the end of a season but inventory can be carried over to the next season. The

[7] In fact, their model has two retailers. But in one version of their model the retailers face independent demands, and so that model is qualitatively identical to the single retailer model.

Table 1
Wholesale-price contract performance when demand follows a power distribution with parameter k

Demand distribution parameter, k	0.2	0.4	0.8	1.6	3.2
Efficiency, $\Pi(q_s^*)/\Pi(q^o)$	73.7%	73.9%	74.6%	76.2%	79.1%
Supplier's share, $\pi_s(q_s^*, w(q_s^*))/\Pi(q_s^*)$	54.5%	58.3%	64.3%	72.2%	80.8%
Coefficient of variation	1.51	1.02	0.67	0.42	0.25

retailer submits orders between seasons and the supplier is able to replenish immediately. Within each season the retailer faces a newsvendor problem that makes the trade-off between lost sales and inventory holding costs. Hence, the retailer's optimal inventory policy is to order up to a fixed level that is the solution to a newsvendor problem. But since inventory carries over from season to season, the supplier's average sales per season equals the retailer's average sales per season, i.e., the supplier's profit function is $(w(q) - c)S(q)$. The analysis of the supplier's optimal wholesale price is more complex in this setting because the supplier's profit is now proportional to the retailer's sales, $S(q)$, rather than to his order quantity, q. Nevertheless, since $S(q)/S'(q) > q$, the supplier's optimal wholesale price is lower than in the single season model. (The expression $S(q)/S'(q) > q$ simplifies to $qF(q) > \int_0^q F(y)\,dy$, which clearly holds.) Thus, the efficiency of the wholesale-price contract is even better than in the single season model.

Debo (1999) studies the repeated version of the single-shot newsvendor model without inventory carrying from period to period. He demonstrates that supply chain coordination is possible with just a wholesale-price contract if the firm's discount rate is not too high, i.e., the firms care about future profit. Cooperation is achieved via the use of trigger strategies that punish a defector.

The infinite horizon extensions to the model do not have an end-of-horizon effect, i.e., inventory is not salvaged. Cachon (2002) studies a two-period version of the model which has excess inventory and demand updating. The retailer can submit an order well in advance of the selling season and pay the supplier w_1 for each unit in that order. The supplier then produces and delivers the retailer's first order before the selling season starts. During the selling season the retailer can order from the supplier additional units. If the supplier has inventory available, then the supplier delivers the units and charges the retailer w_2 per unit, $w_2 \geq w_1$. The supply chain can operate in one of three modes. The first matches the single-period model studied by Lariviere and Porteus (2001): only the initial order before the season starts is allowed. This mode of operations is called 'push', because all inventory risk is pushed upon the retailer (i.e., the retailer bears the cost of disposing leftover inventory). The other extreme is called 'pull': the retailer orders only during the selling season, so now the supplier bears all inventory risk. A combination of push and pull is created by the use of an advanced purchase discount, $w_1 < w_2$: the retailer submits an initial order to take advantage of the advanced purchase discount and the supplier produces more than the initial order in anticipation of the retailer's orders during the selling season. It is shown that supply chain efficiency is substantially higher if the firms consider both push and pull contracts rather than just push or just pull contracts. Furthermore, there exist conditions in which advanced purchase discounts coordinate the supply chain and arbitrarily allocate its profit. See Ferguson, DeCroix and Zipkin (2002), Taylor (2002b) and Yüksel and Lee (2002) for additional work on the timing of the retailer's orders. See Section 6 for a model that studies demand updating with coordinating contracts.

Dong and Rudi (2001) study the wholesale-price contract with two newsvendors and transshipment of inventory between them. They find that the supplier is generally able to capture most of the benefits of transshipments and the retailers are worse off with transshipment. This is consistent (but probably not identical) to Lariviere and Porteus' (2001) finding that the supplier is better off and the retailer worse off with less variable demand. Related to transshipment, Chod and Rudi (2002) study a supplier selling a single resource to a downstream firm that can use that resource to produce multiple products.

Gilbert and Cvsa (2000) study the wholesale-price contract with demand uncertainty and costly investment to reduce production costs. They demonstrate that a trade-off exists between the beneficial flexibility of allowing the wholesale price to adjust to market demand and the need to provide incentives to reduce production costs. Additional detail on this paper is provided in Section 4.

2.3 The buyback contract

With a buyback contract the supplier charges the retailer w per unit purchased, but pays the retailer b per unit remaining at the end of the season:

$$T_{\mathrm{b}}(q, w, b) = wq - bI(q) = bS(q) + (w - b)q.$$

A retailer should not profit from leftover inventory, so assume $b \leq w$. See Pasternack (1985) for a detailed analysis of buyback contracts in the context of the newsvendor problem.

Buyback contracts are also called returns policies, but, unfortunately, both names are somewhat misleading since they both imply the units remaining at the end of the season are physically returned to the supplier. That does occur if the supplier's net salvage value is greater than the retailer's net salvage value. However, if the retailer's salvage value is higher, the retailer salvages the units and the supplier credits the retailer for those units, which is sometime referred to as 'markdown money' (see Tsay, 2001). An important implicit assumption is that the supplier is able to verify the number of remaining units and the cost of such monitoring does not negate the benefits created by the contract.

With a buyback contract the retailer's profit is

$$\pi_r(q, w_b, b) = (p - v + g_r - b)S(q) - (w_b - b + c_r - v)q - g_r\mu.$$

Now consider the set of buyback parameters $\{w_b, b\}$ such that for $\lambda \geq 0$,

$$p - v + g_r - b = \lambda(p - v + g) \tag{2.5}$$

$$w_b - b + c_r - v = \lambda(c - v) \tag{2.6}$$

A comparison with Eq. (2.1) reveals the retailer's profit function with that set of contracts is

$$\begin{aligned}\pi_r(q, w_b, b) &= \lambda(p - v + g)S(q) - \lambda(c - v)q - g_r\mu \\ &= \lambda\Pi(q) + \mu(\lambda g - g_r).\end{aligned} \tag{2.7}$$

It follows immediately that $q_r^* = q^o$ is optimal for the retailer. The supplier's profit function is

$$\pi_s(q, w_b, b) = \Pi(q) - \pi_r(q, w_b, b) = (1 - \lambda)\Pi(q) - \mu(\lambda g_s - (1 - \lambda)g_r).$$

So the buyback contract coordinates with voluntary compliance as long as $\lambda \leq 1$. Some ambiguity arises with $\lambda = 1$ (or $\lambda = 0$) because then q^o is optimal for the supplier (or retailer), but so is every other quantity. Hence, coordination is possible, but the optimal solution is no longer the unique Nash equilibrium.

Interestingly, voluntary compliance actually increases the robustness of the supply chain. Suppose the retailer is not rational and orders $q > q^o$. Since the supplier is allowed to deliver less than the retailer's order quantity, the supplier corrects the retailer's mistake by delivering only q^o units. However, because the retailer can refuse to accept more than he orders, the supplier cannot correct the retailer's mistake if he orders less than q^o. See Chen (1999a), Porteus (2000) and Watson (2002) for further discussion on the robustness of a coordination scheme to irrational ordering.[8]

The retailer's profit is increasing in λ and the supplier's profit is decreasing in λ so the λ parameter acts to allocate the supply chain's profit

[8] While this is an intriguing idea, it is difficult to construct a theory based on irrational behavior. Maybe a better interpretation is that shocks occur in the system that are only observable to one member. For example, while q^o is the steady-state optimal order quantity, the supplier may learn some information that reveals in fact $q' < q^o$ is indeed optimal. Thus, the retailer's apparently irrational excessive order is really due to a lack of information. Some interesting research must be able to follow from these ideas. Perhaps inspiration could come from Stidham (1992). He considers the regulation of a queue when a manager sets her actions for a defined time period but actual expected demand during that period may deviate from what the manager expects. He shows there may exist unstable equilibria, i.e., a small shock to the system sends the system away from the equilibrium rather than back to it.

between the two firms. The retailer earns the entire supply chain profit, $\pi_r(q^o, w_b, b) = \Pi(q^o)$, when

$$\lambda = \frac{\Pi(q^o) + \mu g_r}{\Pi(q^o) + \mu g} \leq 1 \tag{2.8}$$

and the supplier earns the entire supply chain profit, $\pi_s(q^o, w_b, b) = \Pi(q^o)$, when

$$0 \leq \lambda = \frac{\mu g_r}{\Pi(q^o) + \mu g}. \tag{2.9}$$

So every possible profit allocation is feasible with this set of coordinating contracts, assuming $\lambda = 0$ and $\lambda = 1$ are considered feasible.

It is interesting to note that coordination of the supply chain requires the simultaneous adjustment of both the wholesale price and the buyback rate. This has implications for the bargaining process. For example, suppose the firms agree to a particular wholesale price. Given any fixed wholesale price, the coordinating buyback rate is not the buyback rate that maximizes the supplier's or the retailer's profit. In other words, both players would have an incentive to argue for a non-Pareto optimal (i.e., noncoordinating) contract. It would be a shame if the players then agreed upon an non-Pareto optimal contract because then, by definition, there would exist some coordinating contract that could make both players better off. However, that coordinating contract would have a different wholesale price. The key lesson for managers is that they should never negotiate these parameters sequentially (i.e., agree to one parameter and then consider the second parameter). Instead, negotiations should always allow simultaneous changes to both the wholesale price and the buyback rate.

From Eq. (2.7), the parameter λ can loosely be interpreted as the retailer's share of the supply chain's profit; it is precisely the retailer's share when $g_r = g_s = 0$. Note that the λ parameter is not actually part of the buyback contract. It is introduced for expositional clarity. Most of the supply chain contracting literature does not explicitly define a comparable parameter. Instead, it is more common to present one contract parameter in terms of the other, e.g.,

$$w_b(b) = b + c_s - (c - v)\left(\frac{b + g_s}{p - v + g}\right).$$

Furthermore, coordinating parameters are often identified from first-order conditions. The approach taken above is preferred because it is more general. For example, the strategy space does not need to be continuous, there does not need to be a unique optimum and the supply chain cost function does not have to be continuous, even though all of those conditions are satisfied in this model (which is why the first-order condition approach works).

In general, a contract coordinates the retailer's and the supplier's action whenever each firm's profit is an affine function of the supply chain's profit. In effect the firms end up with something that resembles a profit-sharing arrangement. Jeuland and Shugan (1983) note that profit sharing can coordinate a supply chain, but they do not offer a specific contract for achieving profit sharing. Caldentey and Wein (1999) show that profit sharing occurs when each firm receives a fixed fraction of every other firm's utility. With that approach each firm transacts with every other firm, which may lead to an administrative burden if the number of firms is large.

There is a substantial literature on buyback contracts. Padmanabhan and Png (1995) describe several motivations for return policies that are not included in the newsvendor model. A supplier may wish to offer a return policy to prevent the retailer from discounting leftover items, thereby weakening the supplier's brand image. For instance, suppliers of fashion apparel have large marketing budgets to enhance the popularity of their clothes. It is difficult to convince consumers that your clothes are popular if they can be found in the discount rack at the end of the season. Alternatively, a supplier may wish to accept returns to rebalance inventory among retailers. There are a number of papers that consider stock rebalancing in a centralized system (Lee, 1987; Tagaras & Cohen, 1992). Rudi, Kapur and Pyke (2001) and Anupindi, Bassok and Zemel (2001) consider inventory rebalancing in decentralized systems.

In Padmanabhan and Png (1997) a supplier uses a buyback contract to manipulate the competition between retailers (see Section 5.2) Emmons and Gilbert (1998) study buyback contracts with a retail price-setting newsvendor (see Section 3). Taylor (2000a) incorporates a buyback contract with a sales-rebate contract to coordinate the newsvendor with effort-dependent demand (see Section 4). Donohue (2000) studies buyback contracts in a model with multiple production opportunities and improving demand forecasts (see Section 6). Anupindi and Bassok (1999) demonstrate buyback contracts can coordinate a two-retailer supply chain in which consumers search among the retailers to find inventory.[9] Lee, Padmanabhan, Taylor and Whang (2000) model price protection policies in a way that closely resembles a buyback.[10] However, in Taylor (2001) price protection is distinct from

[9] In their model the supplier subsidizes the holding cost of leftover inventory, which is analogous to a buyback.

[10] In their first model a retailer makes a single purchase decision even though demand occurs over two periods. Price protection is modeled as a credit for each unit remaining at the end of the first period, which resembles a buyback. In their second model the retailer may purchase at the start of each period. Again, price protection is modeled as a credit for each unit not sold at the end of the first period; as with a buyback, the price protection reduces the retailer's overage cost.

buybacks. He demonstrates coordination with arbitrary allocation of profit requires price protection in addition to buybacks when the retail price declines with time.[11]

In the context of capital intensive industry, Wu, Kleindorfer and Zhang (2002) study contracts that are similar to buyback contracts. There is one supplier and one buyer. The buyer reserves Q units of capacity for a fee, s, and pays another fee for each unit of capacity utilized, g. This is analogous to a buyback contract with a wholesale price $w=s+g$ and a buyback rate $b=g$: the buyer pays $s+g$ for each unit of capacity that is reserved and receives g for each unit of capacity *not* utilized. The buyer's demand depends on the contract parameters and the uncertain spot price for additional capacity: if the spot is less than g then the buyer satisfies his demand via the spot market exclusively, but with higher spot prices the buyer uses an optimal mixture of the reserved capacity and the spot market.

2.4 The revenue-sharing contract

With a revenue-sharing contract the supplier charges w_r per unit purchased plus the retailer gives the supplier a percentage of his revenue. Assume all revenue is shared, i.e., salvage revenue is also shared between the firms. (It is also possible to design coordinating revenue-sharing contracts in which only regular revenue is shared.) Let ϕ be the fraction of supply chain revenue the retailer keeps, so $(1-\phi)$ is the fraction the supplier earns. Revenue-sharing contracts have been applied recently in the video cassette rental industry with much success. Cachon and Lariviere (2000) provide an analysis of these contracts in a more general setting.

The transfer payment with revenue sharing is

$$T_r(q, w_r, \phi) = (w_r + (1-\phi)v)q + (1-\phi)(p-v)S(q).$$

The retailer's profit function is

$$\pi_r(q, w_r, \phi) = (\phi(p-v) + g_r)S(q) - (w_r + c_r - \phi v)q - g_r\mu.$$

Now consider the set of revenue-sharing contracts, $\{w_r, \phi\}$, such that $\lambda \geq 0$ and

$$\phi(p-v) + g_r = \lambda(p-v+g)$$
$$w_r + c_r - \phi v = \lambda(c-v).$$

[11] In this model the retailer can either order additional units at the end of the first period or return units to the supplier. Price protection is now a credit for each unit retained. Therefore, price protection is a subsidy for retaining inventory whereas the buyback is a subsidy for disposing inventory.

With those terms the retailer's profit function is

$$\pi_r(q, w_r, \phi) = \lambda\Pi(q) + \mu(\lambda g - g_r). \tag{2.10}$$

Hence, q^o is the retailer's optimal order quantity. The supplier's profit is

$$\pi_s(q, w_r, \phi) = \Pi(q) - \pi_r(q, w_r, \phi) = (1-\lambda)\Pi(q) - \mu(\lambda g - g_r),$$

so q^o is the supplier's optimal production quantity as long as $\lambda \leq 1$. The retailer's profit is increasing in λ and the supplier's is decreasing in λ. It is easy to confirm that Eqs. (2.8) and (2.9) provide the parameter values for λ such that the retailer's profit equals the supply chain's profit with the former and the supplier's profit equals the supply chain's profit with the latter. Hence, those revenue-sharing contracts coordinate the supply chain and arbitrarily allocate its profit.

The similarity between Eqs. (2.10) and (2.7) suggests a close connection between revenue-sharing and buyback contracts. In fact, in this setting they are equivalent. Consider a coordinating buyback contract, $\{w_b, b\}$. With that contract the retailer pays $w_b - b$ for each unit purchased and an additional b per unit sold. (The common description for a buyback contract has the retailer paying w_b per unit purchased and receiving a credit of b per unit not sold, which is the same as paying $w_b - b$ for each unit purchased no matter the demand realization and an additional b per unit sold.) With revenue sharing the retailer pays $w_r + (1-\phi)v$ for each unit purchased and $(1-\phi)(p-v)$ for each unit sold. Therefore, revenue-sharing and a buyback contract are equivalent when

$$w_b - b = w_r + (1-\phi)v$$
$$b = (1-\phi)(p-v)$$

In other words, the revenue-sharing contract $\{w_r, \phi\}$ generates the same profits for the two firms for any realization of demand as the following buyback contract,

$$w_b = w_r + (1-\phi)p$$
$$b = (1-\phi)(p-v)$$

While these contracts are equivalent in this setting, Sections 3 and 5.2 demonstrate that their paths diverge in more complex settings.

There are several other papers that investigate revenue-sharing contracts. Mortimer (2000) provides a detailed econometric study of the impact of revenue-sharing contracts in the video rental industry. She finds that the adoption of these contracts increased supply chain profits by 7%. Dana and Spier (2001) study these contracts in the context of a perfectly competitive retail market. Pasternack (1999) studies a single retailer newsvendor model in which the retailer can purchase some units with revenue sharing and other units with a wholesale-price contract. He does not consider supply chain coordination in his model. Gerchak, Cho and Ray (2001) consider a video retailer that decides how many tapes to purchase and how much time to keep them. Revenue sharing coordinates their supply chain, but only provides one division of profit. They redistribute profits with the addition of a licensing fee. Wang, Jiang and Shen (2001) consider revenue sharing with consignment (i.e., $w_r = 0$).

2.5 The quantity-flexibility contract

With a quantity-flexibility contract the supplier charges w_q per unit purchased but then compensates the retailer for his losses on unsold units. To be specific, the retailer receives a credit from the supplier at the end of the season equal to $(w_q + c_r - v)\min(I, \delta q)$, where I is the amount of leftover inventory, q is the number of units purchased and $\delta \in [0, 1]$ is a contract parameter. [See Yüksel and Lee (2002) for a model in which the return threshold is an absolute quantity instead of a percentage of the retailer's order.] Hence, the quantity-flexibility contract fully protects the retailer on a portion of the retailer's order whereas the buyback contract gives partial protection on the retailer's entire order. (The retailer continues to salvage leftover inventory, which is why the salvage value is not included in each unit's credit.) If the supplier did not compensate the retailer for the c_r cost per unit then the retailer would receive only partial compensation on a limited number of units, which is called a backup agreement. Those contracts are studied by Pasternack (1985) and Eppen and Iyer (1997) and Barnes-Schuster, Bassok and Anupindi (1998).[12]

Tsay (1999) studies supply chain coordination with quantity-flexibility contracts in a model that resembles this one. In Tsay (1999) the retailer receives an imperfect demand signal before submitting his final order (i.e., just before deciding how much to return), whereas in this model the retailer receives a perfect signal, i.e., the retailer observes demand. Nevertheless, since production is done before any demand information is learned, the centralized solution in Tsay (1999) is also a newsvendor problem. The demand signal does not matter to the analysis or to the outcome if the retailer returns units only at the end of the season: by then the demand signal is no longer relevant. However, if the retailer is able to return units after observing the demand signal and before the selling

[12] Eppen and Iyer (1997) do not consider channel coordination. Instead, they consider the retailer's order quantity decision. However, their model is more complex: e.g., it includes demand updates, holding costs and customer returns.

season starts, then the demand signal does matter. Because the inventory that is produced is sunk, the supply chain optimal solution is to keep all inventory at the retailer no matter what signal is received. Allowing the retailer to return inventory (alternatively, allowing the retailer to cancel a portion of the initial order) creates a 'stranded inventory problem': inventory could be stranded at the supplier, unable to be used to satisfy demand. In that situation, as shown in Tsay (1999), a quantity-flexibility contract may actually prevent supply chain coordination. On another issue, Tsay (1999) assumes forced compliance, which does have some significance.[13]

With the quantity-flexibility contract the transfer payment is $T_q(q, w_q, \delta)$

$$T_q(q, w_q, \delta) = w_q q - (w + c_r - v) \int_{(1-\delta)q}^{q} F(y)\,dy,$$

where the last term is the retailer's compensation for unsold units, up to the limit of δq units. The retailer's profit function is

$$\begin{aligned}\pi_r(q, w_q, \delta) &= (p - v + g_r)S(q) - (c_r - v)q - T_q(q, w_q, \delta) - \mu g_r \\ &= (p - v + g_r)S(q) - (w_q + c_r - v)q \\ &\quad + (w_q + c_r - v)\int_{(1-\delta)q}^{q} F(y)\,dy - \mu g_r\end{aligned}$$

To achieve supply chain coordination it is necessary (but not sufficient) that the retailer's first-order condition holds at q^o:

$$(p - v + g_r)S'(q^o) - (w_q + c_r - v)(1 - F(q^o) + (1-\delta)F((1-\delta)q^o)) = 0 \tag{2.11}$$

Let $w_q(\delta)$ be the wholesale price that satisfies Eq. (2.11):

$$w_q(\delta) = \frac{(p - v + g_r)(1 - F(q^o))}{1 - F(q^o) + (1-\delta)F((1-\delta)q^o)} - c_r + v.$$

[13] There are some other minor differences that do not appear to be important qualitatively. He assumes demand is normally distributed. In addition, the retailer's final order must be in the range $[q(1+\alpha), q(1-w)]$, where q is the initial forecast and α and w are contract parameters. In this model the retailer's final order must be in the range $[\delta q, q]$, where q is the initial order and δ is a contract parameter. He does not include a supplier goodwill cost, nor a retailer marginal cost, c_r.

$w_q(\delta)$ is indeed a coordinating wholesale price if the retailer's profit function is concave:

$$\frac{\partial^2 \pi_r(q, w_q(\delta), \delta)}{\partial q^2} = -(p+g_r-w_q(\delta)-c_r)f(q)-(w_q(\delta)+c_r-v)(1+(1-\delta)^2 f((1-\delta)q)) \leq 0$$

which holds when $v - c_r \leq w_q(\delta) \leq p + g_r - c_r$. That range is satisfied with $\delta \in [0, 1]$ because

$$w_q(0) = (p - v + g_r)\overline{F}(q^o) + v - c_r,$$
$$w_q(1) = p + g_r - c_r,$$

and $w_q(\delta)$ is increasing in δ.

For supply chain coordination the supplier must also wish to deliver q^o to the retailer. The supplier's profit function is

$$\pi_s(q, w_q(\delta), \delta) = g_s S(q) + (w_q(\delta) - c_s)q - (w_q(\delta) + c_r - v) \int_{(1-\delta)q}^{q} F(y)\,dy - \mu g_s$$

and

$$\begin{aligned}\frac{\partial \pi_s(q, w_q(\delta), \delta)}{\partial q} &= g_s(1 - F(q)) + (w_q(\delta) - c_s) - (w_q(\delta) + c_r - v)(F(q) \\ &\quad - (1-\delta)F((1-\delta)q)) \\ &= g_s(1 - F(q)) - c + v + (w_q(\delta) + c_r - v)(1 - F(q) \\ &\quad + (1-\delta)F((1-\delta)q))\end{aligned}$$

The supplier's first-order condition at q^o is satisfied:

$$\frac{\partial \pi_s(q^o, w_q(\delta), \delta)}{\partial q} = g_s(1 - F(q^o)) - c + v + (p - v + g_r)(1 - F(q^o)) = 0$$

However, the sign of the second-order condition at q^o is ambiguous,

$$\frac{\partial^2 \pi_s(q, w_q(\delta), \delta)}{\partial q^2} = -w_q(\delta)(f(q) - (1-\delta)^2 f((1-\delta)q)) - g_s f(q).$$

In fact, q^o may be a local minimum (i.e., the above is positive). That occurs when $g_s = 0$ and $(1-\delta)^2 f((1-\delta)q^o)$ is greater than $f(q^o)$, which is possible when δ is small, $f((1-\delta)q^{o})$ is large and $f(q^{o})$ is small. The second condition occurs when $(1-\delta)q^{o} \approx \mu$ and there is little variation in demand (i.e., so most of the density function is concentrated near the mean). The third condition occurs when $f(q^{o})$ is in the tail of the distribution, i.e., when the critical fractile is large. For example, q^{o} is a local minimum for the following parameters: D is normally distributed, $\mu=10$, $\sigma=1$, $p=10$, $c_s=1$, $c_r=0$, $g_r=g_s=v=0$ and $\delta=0.1$. Hence, supply chain coordination under voluntary compliance is not assured with a quantity-flexibility contract even if the wholesale price is $w_q(\delta)$. Channel coordination is achieved with forced compliance since then the supplier's action is not relevant.

Assuming a $(w_q(\delta), \delta)$ quantity-flexibility contract coordinates the channel, now consider how it allocates profit. When $\delta=0$, the retailer earns at least the supply chain optimal profit:

$$\begin{aligned}\pi_r(q, w_q(0), 0) &= (p - v + g_r)S(q) - \left(\frac{p-v+g_r}{p-v+g}\right)(c-v)q^{o} - \mu g_r \\ &= \Pi(q^{o}) + g_s\left(\mu - S(q^{o}) + \overline{F}(q^{o})q^{o}\right) \\ &\geq \Pi(q^{o})\end{aligned}$$

When $\delta=1$, the supplier earns at least the supply chain's optimal profit:

$$\begin{aligned}\pi_s(q, w_q(1), 1) &= g_s S(q^{o}) + (p + g_r - c)q^{o} - (p + g_r - v)\int_0^q F(y)\,\mathrm{d}y - \mu g_s \\ &= \Pi(q^{o}) + \mu g_r \\ &\geq \Pi(q^{o})\end{aligned}$$

Given that the profit functions are continuous in δ, it follows that all possible allocations of $\Pi(q^{o})$ are possible.

There are a number of other papers that study the quantity-flexibility contract, or a closely related contract. Tsay and Lovejoy (1999) study quantity-flexibility contracts in a more complex setting than the one considered here: they have multiple locations, multiple demand periods, lead times and demand forecast updates. Bassok and Anupindi (1997b) provide an in-depth analysis of these contracts for a single-stage system with more general assumptions than in Tsay and Lovejoy (1999). (They refer to their contract as a rolling horizon-flexibility contract.) In multiple-period models it is observed that these contracts dampen supply chain order variability, which is a potentially beneficial feature that the single-period model does not capture.

Cachon and Lariviere (2001) and Lariviere (2002) study the interaction between quantity-flexibility contracts and forecast sharing. In Cachon and Lariviere (2001) a downstream firm has a better demand forecast than the upstream supplier, but needs to convince the upstream supplier that her forecast is genuine. The minimum commitment in a quantity-flexibility contract is a very effective solution for this problem (see Section 10). In Lariviere (2002) the upstream firm wishes to encourage the downstream firm to exert the proper amount of effort to improve his demand forecast.

Plambeck and Taylor (2002) study quantity-flexibility contracts with more than one downstream firm and ex-post renegotiation. With multiple retailers it is possible that one retailer needs more than its initial order, q, and the other retailer needs less than its minimum commitment, δq. This creates an opportunity to renegotiate the contracts, which influences the initial contracts signed and actions taken.

2.6 The sales-rebate contract

With a sales-rebate contract the supplier charges w_s per unit purchased but then gives the retailer an r rebate per unit sold above a threshold t. This contract form is studied by Taylor (2002a) and Krishnan, Kapuscinski and Butz (2001), where the latter refers to it as a 'markdown allowance'. Both models are more complex than the one considered here. In particular, both papers allow the retail to exert effort to increase demand: in Taylor (2002a) effort is chosen simultaneously with the order quantity, whereas Krishnan et al. (2001) focus on the case in which the retailer chooses an order quantity, a signal of demand is observed and then effort is exerted. Hence, if the demand signal is strong relative to the order quantity, then the retailer does not need to exert much effort. See Section 4 for additional discussion of coordination in the presence of retail effort.

The transfer payment with the sales-rebate contract is

$$T_s(q, w_s, r, t) = \begin{cases} w_s q & q < t \\ (w_s - r)q + r\left(t + \int_t^q F(y)\,dy\right) & q \geq t \end{cases}:$$

when $q \geq t$ the retailer pays $w_s - r$ for every unit purchased, an additional r per unit for the first t units purchased and an additional r per unit for the units *not* sold above the t threshold. The retailer's profit function is then

$$\pi_r(q, w_s, r, t) = (p - v + g_r)S(q) - (c_r - v)q - g_r\mu - T_s(q, w_s, r, t)$$

For this contract to achieve supply chain coordination, q^o must at least be a local maximum:

$$\frac{\partial \pi_r(q^o, w_s, r, t)}{\partial q} = (p - v + g_r)\overline{F}(q^o) - (c_r - v) - \frac{\partial T_s(q^o, w_s, r, t)}{\partial q} = 0. \tag{2.12}$$

If $t \geq q^{o}$ the above condition is only satisfied with $w_{s} = c_{s} - g_{s}\overline{F}(q^{o})$, which is clearly not acceptable to the supplier. But this contract is interesting only if it achieves supply chain coordination for $t < q^{o}$. So assume $t < q^{o}$.

Define $w_{s}(r)$ as the wholesale price that satisfies Eq. (2.12):

$$w_{s}(r) = (p - v + g_{r} + r)\overline{F}(q^{o}) - c_{r} + v \tag{2.13}$$

Given that wholesale price, the second-order condition confirms $\pi_{r}(q, w_{s}(r), r, t)$ is strictly concave in q for $q > t$. So q^{o} is a local maximum. But $\pi_{r}(q, w_{s}(r), r, t)$ is strictly concave in q for $q \leq t$ and, due to a 'kink' at $q = t$, $\pi_{r}(q, w_{s}(r), r, t)$ need not be unimodal in q. Let $\overline{q} = \arg\max_{q \leq t} \pi_{r}(q, w_{s}(r), r, t)$. Hence, it is necessary to demonstrate that there exist coordinating contracts such that q^{o} is preferred by the retailer over $\overline{q}$. Substitute $w_{s}(r)$ into the retailer's profit function:

$$\pi_{r}(q, w_{s}(r), r, t) = \Pi(q) + g_{s}\big(\mu - S(q) + q\overline{F}(q^{o})\big) - rq\overline{F}(q^{o}) + \begin{cases} 0 & q < t \\ rq - r\big(t + \int_{t}^{q} F(y)\,dy\big) & q \geq t \end{cases}$$

and

$$\pi_{r}(q^{o}, w_{s}(r), r, t) = \Pi(q^{o}) + g_{s}\big(\mu - S(q^{o}) + q^{o}\overline{F}(q^{o})\big) + r\left(q^{o}F(q^{o}) - t - \int_{t}^{q^{o}} F(y)\,dy\right).$$

With $t = 0$ the retailer earns more than $\Pi(q^{o})$, so q^{o} is surely optimal. With $t = q^{o}$, the retailer's profit function is decreasing for $t \geq q^{o}$; $\overline{q}$ is at least as good for the retailer as q^{o}. Given that $\pi_{r}(q^{o}, w_{s}(r), r, t)$ is decreasing in t, there must exist some t in the range $[0, q^{o}]$ such that $\pi_{r}(q^{o}, w_{s}(r), r, t) = \pi_{r}(\overline{q}, w_{s}(r), r, t)$.

Now consider the allocation of profit. We have already established that with $t = 0$ the retailer earns more than $\Pi(q^{o})$. Hence, there must be a t such that $\pi_{r}(q^{o}, w_{s}(r), r, t) = \Pi(q^{o})$, i.e., the retailer earns the supply chain's profit. When $t = q^{o}$, the retailer earns $\pi_{r}(\overline{q}, w_{s}(r), r, t)$, and with a sufficiently large r such that profit is zero; the supplier earns the supply chain's profit. In fact, there is generally a set of contracts that generate any profit allocation because the sales-rebate contract is parameter rich: these three parameters are more than sufficient to coordinate one action and to redistribute rents.

Now consider the supplier's production decision. The supplier's profit function given a coordinating sales-rebate contract is

$$\pi_{s}(q, w_{s}(r), r, t) = -g_{s}(\mu - S(q)) - c_{s}q + T_{s}(q, w_{s}(r), r, t).$$

For $q > t$,

$$\frac{\partial \pi_s(q, w_s(r), r, t)}{\partial q} = g_s \overline{F}(q) - c_s + w_s(r) - r + rF(q)$$
$$= (r - g_s)(F(q) - F(q^o))$$

The above is positive for $q \leq q^o$ only if $r < g_s$. But if $r \leq g_s$, then $w_s(r) \leq c_s$; the supplier cannot earn a positive profit with $r < g_s$. As a result, it must be that $r > g_s$, which implies the supplier loses money on each unit delivered to the retailer above t: the retailer effectively pays the supplier $w_s(r) - r$ for each unit sold above the threshold t and from Eq. (2.13),

$$w_s(r) - r = c_s - v - g_s \overline{F}(q^o) - rF(q^o) < c_s.$$

So the sales-rebate contract does not coordinate the supply chain with voluntary compliance.

2.7 The quantity-discount contract

There are many types of quantity discounts.[14] This section considers an 'all unit' quantity discount, i.e., the transfer payment is $T_d(q) = w_d(q)q$, where $w_d(q)$ is the per unit wholesale price that is decreasing in q. The retailer's profit function is then

$$\pi_r(q, w_d(q)) = (p - v + g_r)S(q) - (w_d(q) + c_r - v)q - g_r \mu.$$

One technique to obtain coordination is to choose the payment schedule such that the retailer's profit equals a constant fraction of the supply chain's profit. To be specific, let

$$w_d(q) = ((1 - \lambda)(p - v + g) - g_s)\left(\frac{S(q)}{q}\right) + \lambda(c - v) - c_r + v.$$

The above is decreasing in q if $\lambda \leq \overline{\lambda}$, where

$$\overline{\lambda} = \frac{p - v + g_r}{p - v + g},$$

[14] Roughly speaking, the quantity-discount contract achieves coordination by manipulating the retailer's marginal cost curve, while leaving the retailer's marginal revenue curve untouched. Coordination is achieved if the marginal revenue and marginal cost curves intersect at the optimal quantity. Hence, there is an infinite number of marginal cost curves that intersect the marginal revenue curve at a single point. See Moorthy (1987) for a more detailed explanation for why many coordinating quantity discount schedules exist. See Kolay and Shaffer (2002) for a discussion on different types of quantity discounts. See Wilson (1993) for a much broader discussion of nonlinear pricing.

since $S(q)/q$, i.e., sales per unit ordered, is decreasing in q. The retailer's profit function is now

$$\begin{aligned}\pi_{\rm r}(q, w_{\rm d}(q)) &= \lambda(p - v + g)S(q) - \lambda(c - v)q - g_{\rm r}\mu \\ &= \lambda(\Pi(q) + g\mu) - g_{\rm r}\mu.\end{aligned}$$

Hence, $q^{\rm o}$ is optimal for the retailer and the supplier. As with the buyback and revenue-sharing contracts, the parameter λ acts to allocate the supply chain's profit between the two firms. However, the upper bound on λ prevents too much profit from being allocated to the retailer with a quantity discount. Technically, the $w_{\rm d}(q)$ schedule continues to coordinate even if $\lambda > \bar{\lambda}$, but then $w_{\rm d}(q)$ is increasing in q. In that case the retailer pays a quantity premium. See Tomlin (2000) for a model with both quantity-discount and quantity-premium contracts.[15]

2.8 Discussion

This section studies five contracts, two of which are equivalent (revenue-sharing and buyback contracts), to coordinate the newsvendor and to divide the supply chain's profit. Each contract coordinates by inducing the retailer to order more than he would with just a wholesale-price contract. Revenue-sharing and quantity-flexibility contracts do this by giving the retailer some downside protection: if demand is lower than q, the retailer gets some refund. The sales-rebate contract does this by giving the retailer upside incentive: if demand is greater than t, the retailer effectively purchases the units sold above t for less than their cost of production. The quantity discount coordinates by adjusting the retailer's marginal cost curve so that the supplier earns progressively less on each unit. However, an argument has not yet been made for why one contract form should be observed over another.

The various coordinating contracts may not be equally costly to administer. The wholesale-price contract is easy to describe and requires a single transaction between the firms. The quantity discount also requires only a single transaction, but it is more complex to describe. The other coordinating contracts are more costly to administer: the supplier must monitor the number of units the retailer has left at the end of the season, or the remaining units must be transported back to the supplier, depending on where the units are salvaged. Hence, the administrative cost argument does not explain the selection among buyback, revenue-sharing and quantity-flexibility contracts, but may explain the selection of a quantity-discount or a wholesale-price contract.

[15] Tomlin (2000) studies a supplier–manufacturer supply chain in which both firms incur costs to install capacity and both firms incur costs to convert capacity into units.

The risk neutrality assumption notwithstanding, the contracts do differ with respect to risk. With the exception of the quantity-discount contract, each of the coordinating contracts shifts risk between the two firms: as the retailer's share of profit decreases, his risk decreases and the supplier's risk increases. Hence, these contracts could provide some insurance to a risk averse retailer, but would be costly to a risk averse supplier. See Eeckhoudt, Gollier and Schlesinger (1995), Schweitzer and Cachon (2000) and Chen and Federgruen (2000) for a discussion of risk in the single-firm newsvendor model. Agrawal and Seshadri (2000) do study the influence of risk aversion in supply chain contracting. They argue that risk aversion among retailers provides an incentive for a distributor to provide risk intermediation services. In their model the distributor offers a contract with a fixed fee, a wholesale price, a return rate and premium fee for units ordered on an emergency basis to cover demand in excess of the retailer's order quantity. Finally, Plambeck and Zenios (2000) provide a principle-agent model that does incorporate risk aversion.

The supplier's exposure to demand uncertainty with some of the coordinating contracts could matter to the supplier if the retailer chose an order quantity other than q^o. For example, if the supplier offers a generous buyback to the retailer, then the supplier does not want the retailer to order too much products. Under voluntary compliance the supplier can avoid this excessive ordering error by shipping only q^o. But with forced compliance the supplier bears the full risk of an irrational retailer, a risk that even a risk neutral supplier may choose to avoid. However, with voluntary compliance the supplier may ship less than the retailer's order even if everyone is quite rational: revenue sharing and quantity discounts always coordinate the supplier's action with voluntary compliance, quantity-flexibility contracts generally, but not always, coordinate the supplier's action and sales-rebate contracts never do.

Now consider the application of these contracts in a setting with heterogenous retailers that do not compete, i.e., the action of one retailer has no impact on any other retailer, probably because of geographic dispersion. In general, suppliers are legally obligated to offer the same contractual terms to their retailers; hence, it is desirable for the supplier to offer the same contract to all of her retailers, or at the very least, the same menu of contracts.[16] If only one contract is offered, then it coordinates all of the retailers as long as the set of coordinating contracts does not depend on something that varies across the retailers. For example, the coordinating revenue-sharing contracts do not depend on the demand distribution, but do depend on the retailer's marginal cost. Hence, a single revenue-sharing contract can coordinate retailers with heterogenous demands, but not necessarily retailers with different marginal costs. However, in some cases heterogeneity can be accommodated with a single contract. Consider the

[16] Actually, a supplier can offer different contracts to retailers that do not compete.

quantity-flexibility contract, which depends on the demand distribution, and two retailers that have demands that differ by a scale factor; let retailer i's demand distribution be $F_i(x \mid \theta_i) = F(x/\theta_i)$, where θ_i is the scale parameter. Hence, the same wholesale price coordinates different retailers, $w_q(\delta \mid \theta_i) = w_q(\delta \mid \theta_j)$.

The independence of a contract to some parameter is also advantageous if the supplier lacks information regarding that parameter. For example, a supplier does not need to know a retailer's demand distribution to coordinate the supply chain with a revenue-sharing contract, but would need to know the retailer's demand distribution with a quantity-flexibility, sales-rebate or quantity-discount contract.

However, there may also be situations in which the supplier wishes to divide the retailers by offering a menu of contracts. For example, Lariviere (2002) studies a model with one supplier selling to a retailer that may exert effort to improve his demand forecast. He considers whether it is useful to offer two types of contracts, one for a retailer that exerts effort and one for a retailer that does not. Since coordinating buyback contracts are independent of the demand distribution, this separation requires the supplier to offer noncoordinating buyback contracts, i.e., supply chain efficiency must be sacrificed to induce forecasting. Quantity-flexibility contracts do depend on the demand distribution, so a menu can be constructed with two coordinating quantity-flexibility contracts, i.e., supply chain efficiency need not be sacrificed. Surprisingly, unless forecasting is very expensive, the supplier is still better off using the menu of buyback contracts even though this sacrifices some efficiency.

To summarize, the set of coordinating contracts is quite large and it is even quite likely that there exist other types of coordinating contracts. While it is possible to identify some differences among the contracts (e.g., different administrative costs, different risk exposures, etc.) none of them is sufficiently compelling to explain why one form should be adopted over another. More theory probably will not provide the answer. We now need some data and empirical analysis.

3 Coordinating the newsvendor with price-dependent demand

In the newsvendor model the retailer impacts sales only through his stocking decision, but in reality a retailer may influence sales through many different actions. Probably the most influential one is the retailer's pricing action. This section studies coordination in the newsvendor model with price-dependent retail demand. A key question is whether the contracts that coordinate the retailer's order quantity also coordinate the retailer's pricing. It is shown that buybacks, quantity-flexibility and sales-rebate contracts do not coordinate in this setting. Those contracts run into trouble because the incentive they provide to coordinate the retailer's quantity action distorts the retailer's pricing decision. Revenue sharing coordinates if there are no goodwill penalties, $g_s = g_r = 0$. With goodwill penalties there exists a single coordinating revenue-sharing contract that provides only a single allocation of the supply chain's profit. The quantity discount does better: it coordinates and allocates profit even if $g_r \geq 0$, but $g_s = 0$ is required. Another contract is introduced, the price-discount contract, which is shown to coordinate and arbitrarily allocate profit. It is essentially a buyback contract with price contingent parameters, i.e., it is a buyback contract with parameters that are set only after the retailer chooses his price. The idea of contingent contracts can also be applied with revenue-sharing contracts when there are goodwill penalties.

3.1 Model and analysis

This model is identical to the one in Section 2 except now the retailer chooses his price in addition to his order quantity. Let $F(q \mid p)$ be the distribution function of demand, where p is the retail price. It is natural to assume demand decreases stochastically in price, i.e., $\partial F(q \mid p)/\partial p > 0$. In a realistic model the retailer would be able to adjust his price throughout the season, possibly for a fee for each adjustment. Such a dynamic pricing strategy would allow the retailer to adjust his price to reflect demand conditions: e.g., if demand were less than expected the retailer could accelerate price discounts. This dynamic pricing problem is quite complex even when supply chain coordination is not considered. Hence, to obtain initial insights, assume the retailer sets his price at the same time as his stocking decision and the price is fixed throughout the season.[17]

The integrated channel's profit is

$$\Pi(q, p) = (p - v + g)S(q, p) - (c - v)q - g\mu$$

where $S(q, p)$ is expected sales given the stocking quantity q and the price p,

$$S(q, p) = q - \int_0^q F(y \mid p)\,\mathrm{d}y.$$

The integrated channel profit function need not be concave nor unimodal (see Petruzzi & Dada, 1999). Assume there exists a finite (but not necessarily unique) optimal quantity-price pair, $\{q^o, p^o\}$.

[17] A hybrid model may be more tractable. For example, suppose the retailer chooses q, then observes a demand signal and then chooses price. van Mieghem and Dada (1999) study a related model in the context of a single firm. The multiretailer model in Section 5.2 is also closely related: order quantities are chosen first, then demand is observed and then price is set to clear the market, i.e., price is variable but not a decision the firms have direct control over.

Let $p^o(q)$ be the supply chain optimal price for a given q. The following first-order condition is necessary for coordination (but not sufficient),

$$\frac{\partial \Pi(q, p^o(q))}{\partial p} = S(q, p^o(q)) + (p^o(q) - v + g)\frac{\partial S(q, p^o(q))}{\partial p} = 0. \tag{3.1}$$

A contract fails to coordinate if it is unable to satisfy the first-order condition at $p^o(q)$, or it is able to satisfy the first-order condition at $p^o(q)$ only with parameters that fail to coordinate the quantity decision.

Consider the quantity-flexibility contract. The retailer's profit function is

$$\begin{aligned}\pi_r(q, p, w_q, \delta) = (p - v + g_r)S(q, p) - (w_q + c_r - v)q \\ + (w_q + c_r - v)\int_{(1-\delta)q}^{q} F(y \mid p)\,dy - \mu g_r\end{aligned}$$

For price coordination the first-order condition must hold,

$$\begin{aligned}\frac{\partial \pi_r(q, p^o(q), w_q, \delta)}{\partial p} &= S(q, p^o(q)) + (p^o(q) - v + g_r)\frac{\partial S(q, p^o(q))}{\partial p} \\ &\quad + (w_q + c_r - v)\int_{(1-\delta)q}^{q} \frac{\partial F(y \mid p^o(q))}{\partial p}\,dy \\ &= 0. \end{aligned} \tag{3.2}$$

The second term in Eq. (3.2) is no smaller than the second term in Eq. (3.1), so the above holds only if the third term is nonpositive. But the third term is nonnegative with a coordinating w_q, so coordination can only occur if $g_s = 0$ and either $w_q = v - c_r$ or $\delta = 0$. Neither is desirable. With $w_q = v - c_r$ the supplier earns a negative profit ($w_q < c_s$), so the supplier certainly cannot be better off with that coordinating contract. With $\delta = 0$ the quantity-flexibility contract is just a wholesale-price contract, so the retailer's quantity action is not optimal (assuming the supplier desires a positive profit). Hence, the quantity-flexibility contract does not coordinate the newsvendor with price-dependent demand.

The sales-rebate contract does not fare better in this setting. With that contract

$$\begin{aligned}\frac{\partial \pi_r(q, p^o(q), w_s, r, t)}{\partial p} &= S(q, p^o(q)) + (p^o(q) - v + g_r)\frac{\partial S(q, p^o(q))}{\partial p} \\ &\quad - r\int_{t}^{q} \frac{\partial F(y \mid p^o(q))}{\partial p}\,dy.\end{aligned}$$

Since the last term is negative when $r > 0$ and $t < q$, the retailer prices below the optimal price. Coordination might be achieved if something is added to the sales-rebate contract to induce the retailer to a higher price. A buyback could provide that counterbalance: a buyback reduces the cost of leftover inventory, so a retailer need not price as aggressively to generate sales.

Now consider a buyback contract on its own. The retailer's profit function is

$$\pi_r(q, p, w_b, b) = (p - v + g_r - b)S(q, p) - (w_b - b + c_r - v)q - g_r\mu.$$

For coordination the supply chain optimal price must satisfy the first-order condition,

$$\begin{aligned}\frac{\partial \pi_r(q, p^o(q), w_b, b)}{\partial p} &= S(q, p^o(q)) + (p^o(q) - v + g_r - b)\frac{\partial S(q, p^o(q))}{\partial p} \\ &= 0. \end{aligned} \tag{3.3}$$

But a comparison of Eqs. (3.3) and (3.1) reveals Eq. (3.3) holds only if $b = -g_s$, which may violate the $b \geq 0$ constraint. In addition, with $b = -g_s$ the coordinating wholesale price is not acceptable to the supplier, $w_b(-g_s) = c_s - g_s$. Therefore, a buyback contract does not coordinate the newsvendor with price-dependent demand.[18] That result is also demonstrated by Marvel and Peck (1995) and Bernstein and Federgruen (2000). While Emmons and Gilbert (1998) recognize that the buyback contract does not coordinate this model, they nevertheless demonstrate a buyback contract with $b > 0$ may still perform better than a wholesale-price contract.

The buyback contract fails to coordinate in this setting because the parameters of the coordinating contracts depend on the price: from Eqs. (2.5) and (2.6), the coordinating parameters are

$$\begin{aligned} b &= (1 - \lambda)(p - v + g) - g_s, \\ w_b &= \lambda c_s + (1 - \lambda)(p + g - c_r) - g_s. \end{aligned}$$

For a fixed λ, the coordinating buyback rate and wholesale price are linear in p. Hence, the buyback contract coordinates the newsvendor with price-dependent demand if b and w_b are made contingent on the retail price chosen, or if b and w_b are chosen after the retailer commits to a price (but before the retailer chooses q). This is the price-discount-sharing contract studied by Bernstein and Federgruen (2000), which is also called a 'bill back' in practice. To understand the name for the contract, notice that the retailer

[18] If $g_s = 0$, then there is one buyback contract that coordinates, $w_b = c_s$ and $b_s = 0$. But that contract does not leave the supplier with a positive profit.

gets a lower wholesale price if the retailer reduces his price, i.e., the supplier shares in the cost of a price discount with the retailer. To confirm that this contract coordinates the supply chain, substitute the above contract parameters into the retailer profit function:

$$\begin{aligned}\pi_r(q,p,w_b,b) &= \lambda(p-v+g)S(q,p)-\lambda(c-v)q-g_r\mu \\ &= \lambda(\Pi(q,p)+g\mu)-g_r\mu\end{aligned}$$

Hence, for the retailer as well as the supplier, $\{q^o, p^o\}$ is optimal for $\lambda \in [0,1]$.

Now consider the revenue-sharing contract. With revenue sharing the retailer's profit is

$$\pi_r(q,p,w_r,\phi) = (\phi(p-v)+g_r)S(q,p)-(w_r+c_r-\phi v)q-g_r\mu.$$

Coordination requires

$$\begin{aligned}\frac{\partial \pi_r(q,p^o(q),w_r,\phi)}{\partial p} &= S(q,p^o(q))+(p^o(q)-v+g_r/\phi)\frac{\partial S(q,p^o(q))}{\partial p} \\ &= 0. \qquad (3.4)\end{aligned}$$

There are two important cases to consider: the first has $g_r = g_s = 0$, and the second has at least one positive goodwill cost. Begin with the first case, $g_r = g_s = 0$. In this situation,

$$\frac{\partial \pi_r(q,p,w_r,\phi)}{\partial p} = \frac{\partial \Pi(q,p)}{\partial p}$$

with any revenue-sharing contract. Thus, the retailer chooses $p^o(q)$ no matter which revenue-sharing contract is chosen. With full freedom to choose the ϕ and w_r parameters, revenue sharing is able to coordinate the retailer's quantity decision with precisely the same set of contracts used when the retailer price is fixed.

Recall that with the fixed price newsvendor revenue sharing and buybacks are equivalent: for every coordinating revenue-sharing contract there exists a buyback contract that generates the same profit allocation for all realizations of demand. Here, the contracts produce different outcomes. The difference occurs because with a buyback the retailer's share of revenue $(1-b/p)$ depends on the price, whereas with revenue sharing it is independent of the price, by definition. However, the price contingent buyback contract (which is also known as the price-discount contract) is again equivalent to revenue sharing: if $g_r = g_s = 0$, the coordinating revenue-sharing contracts yield

$$\pi_r(q,p,w_r,\phi) = \lambda \Pi(q,p)$$

and the price contingent buyback contracts yield the same profit for any quantity and price,

$$\pi_r(q, p, b(p), w_b(p)) = \lambda\Pi(q, p).$$

With the second case (either $g_r > 0$ or $g_s > 0$ revenue sharing is less successful. Now, according to Eq. (3.4), coordination is achieved only if $\phi = g_r/g$. With only one coordinating contract, revenue sharing is able to provide only one profit allocation, albeit both firms may enjoy a positive profit with this outcome, which contrasts with the single coordination outcome of the buyback contract. Again, the difficulty with coordination occurs because the coordinating parameters generally depend on the retail price

$$\phi = \lambda + \frac{\lambda g - g_r}{p - v},$$
$$w_r = \lambda(c - v) - c_r + \phi v.$$

The dependence on the retail price is eliminated only in the special case $\phi = \lambda = g_r/g$.

Coordination for all profit allocations is restored even in this case if, like with the buyback contract, the parameters of the revenue-sharing contract are made contingent on the retailer's price. In that case revenue sharing is again equivalent to the price-discount contract: price discounts are contingent buybacks and contingent buybacks are equivalent to contingent revenue sharing.

The final contract to investigate is the quantity discount. With the quantity discount the retailer keeps all revenue, so only the retailer's marginal cost curve is adjusted. As a result, the quantity discount does not distort the retailer's pricing decision. In many cases, this is ideal. To explain, the retailer's profit function with a quantity discount is

$$\pi_r(q, w_d(q), p) = (p - v + g_r)S(q, p) - (w_d(q) + c_r - v)q - g_r\mu.$$

If $g_s = 0$, then

$$\frac{\partial\pi_r(q, w_d(q), p)}{\partial p} = \frac{\partial S(q, p)}{\partial p} + (p - v + g_r)S(q, p) = \frac{\partial\Pi(q, p)}{\partial p}$$

and so $p^o(q)$ is optimal for the retailer. On the other hand, if $g_s > 0$, then the retailer's pricing decision needs to be distorted for coordination, which the quantity discount does not do. (It is possible that a quantity discount could be designed to correct this distortion when $g_s > 0$, but more careful analysis is required which is left to future research.)

Assuming $g_s = 0$, it remains to ensure that the quantity decision is coordinated. The same schedule can be used as with a fixed retail price, but now the schedule is designed assuming the optimal price is chosen:

$$w_d(q) = ((1-\lambda)(p^o - v + g) - g_s)\left(\frac{S(q,p^o)}{q}\right) + \lambda(c - v) - c_r + v,$$

where $p^o = p^o(q^o)$. It follows that

$$\begin{aligned}\pi_r(q, w_d(q), p) &= (p - v + g_r)S(q,p) - \lambda(c - v)q - g_r\mu \\ &\quad - ((1-\lambda)(p^o - v + g) - g_s)S(q,p^o)\end{aligned}$$

and so p^o is optimal for the retailer,

$$\frac{\partial \pi_r(q, w_d(q), p)}{\partial p} = \frac{\partial \Pi(q,p)}{\partial p}.$$

Given p^o is chosen,

$$\begin{aligned}\pi_r(q, w_d(q), p^o) &= \lambda(p^o - v + g)S(q,p^o) - \lambda(c - v)q - g_r\mu \\ &= (\Pi(q,p^o) + g\mu) - g_r\mu\end{aligned}$$

and so q^o is optimal for the retailer and the supplier. Coordination occurs because the retailer's pricing decision is not distorted, and the retailer's quantity decision is adjusted contingent that p^o is chosen.

3.2 Discussion

There are surely many situations in which a retailer has some control over his pricing. However, incentives to coordinate the retailer's quantity decision may distort the retailer's price decision. This occurs with the buyback, quantity-flexibility and the sales-rebate contracts. Since the quantity discount leaves all revenue with the retailer, it does not create such a distortion, which is an asset when the retailer's pricing decision should not be distorted, i.e., when $g_s = 0$. Revenue sharing does not distort the retailer's pricing decision when $g_r = 0 = g_s$. In those situations the set of revenue-sharing contracts to coordinate the quantity decision with a fixed price continue to coordinate the quantity decision with a variable price. However, when there are goodwill costs, then the coordinating revenue-sharing parameters generally depend on the retail price. The dependence is removed with only a single revenue-sharing contract; hence coordination is only achieved with a single profit allocation. Coordination is restored with arbitrary profit allocation by making the parameters contingent on the retail price chosen, e.g., a menu of revenue-sharing contracts is offered that depend on the price selected. This technique also applies to the buyback contract: the price contingent buyback contract, which is also called a price-discount-sharing contract, coordinates the price-setting newsvendor. In fact, just as buybacks and revenue sharing are equivalent with a fixed retail price, the price contingent buyback and revenue sharing are equivalent when there are no goodwill costs. When there are goodwill costs then the price contingent buyback is equivalent to the price contingent revenue-sharing contract.

4 Coordinating the newsvendor with effort-dependent demand

A retailer can increase a product's demand by lowering his price, but the retailer can take other actions to spur demand: the retailer can hire more sales people, improve their training, increase advertising, better maintain the attractiveness of the product's display, enhance the ambiance of the store interior (e.g., richer materials, wider aisles) and he can give the product a better stocking location within the store. All of those activities are costly. As a result, a conflict exists between the supplier and the retailer: no matter what level of effort the retailer dedicates toward those activities, the supplier prefers that the retailer exert even more effort. The problem is that those activities benefit both firms, but are costly to only one.

Sharing the cost of effort is one solution to the effort coordination problem. For example, the supplier could pay some of the retailer's advertising expenses, or she could compensate the retailer for a portion of his training cost. Several conditions are needed for cost sharing to be an effective strategy: the supplier must be able to observe (without much hassle) that the retailer actually engaged in the costly activity (so the supplier knows how much to compensate the retailer), the retailer's effort must be verifiable to the courts (so that any cost sharing is enforceable) and the activity must directly benefit the supplier.[19] In many cases those conditions are met. For example, the supplier generally can observe and verify whether or not a retailer purchased advertising in a local newspaper. Furthermore, if the ad primarily features the supplier's product, then the benefit of the ad is directed primarily at the supplier. Netessine and Rudi (2000a) present a coordinating contract which involves sharing advertising costs in a model that closely resembles the one in this section. In Wang and Gerchak (2001) the retailer's shelf space can be considered an effort variable. They also allow the supplier to compensate the retailer for his effort, which in their model takes

[19] If the firms interact over a long horizon it may be sufficient that the action is observable even if it is not verifiable, i.e., enforcement can be due to the threat to leave the relationship rather than the threat of court action.

the form of an inventory subsidy. Gilbert and Cvsa (2000) study a model in which effort is observable but not verifiable.

There are also many situations in which cost sharing is not as effective. For example, a supplier probably will not pay for an ad that merely promotes the retailer's brand image. In that case the ad enhances the demand for all of the retailer's products, not just the supplier's product. Also, there are many demand-improving activities that are too costly for the supplier to observe. For example, it may be too costly to visit a store on a frequent basis to ensure the presentation of the supplier's product is maintained to the supplier's satisfaction.[20]

This section studies the challenge of coordinating an action for which there is no direct transfer payment. It is shown that most of the coordinating contracts with the standard newsvendor model no longer coordinate in this setting because the incentives they provide to coordinate the retailer's quantity decision distort the retailer's effort decision. Only the quantity-discount contract continues to coordinate the supply chain. In fact, the quantity-discount contract can coordinate a retailer that chooses quantity, price and effort.

4.1 Model and analysis

To model retail effort, suppose a single effort level, e, summarizes the retailer's activities and let $g(e)$ be the retailer's cost of exerting effort level e, where $g(0)=0$, $g'(e)>0$ and $g''(e)>0$. To help avoid confusion and to simplify the notation, assume there are no goodwill costs, $g_r=g_s=0$, $v=0$ and $c_r=0$. Let $F(q\,|\,e)$ be the distribution of demand given the effort level e, where demand is stochastically increasing in effort, i.e., $\partial F(q\,|\,e)/\partial e<0$. Suppose the retailer chooses his effort level at the same time as his order quantity. Finally, assume the supplier cannot verify the retailer's effort level, which implies the retailer cannot sign a contract binding the retailer to choose a particular effort level. This approach to retail effort has been adopted in a number of marketing papers. For example, see Chu and Desai (1995), Desai and Srinivasan (1995), Desiraju and Moorthy (1997), Gallini and Lutz (1992), Lal (1990) and Lariviere and Padmanabhan (1997).

The integrated channel's profit is

$$\Pi(q,e)=pS(q,e)-cq-g(e),$$

where $S(q,e)$ is expected sales given the effort level e,

$$S(q,e)=q-\int_0^q F(y\,|\,e)\,\mathrm{d}y.$$

[20] However, in some cases it is too costly *not* to visit a store. For example, in the salty snack food category is it common for suppliers to replenish their retailers' shelves.

The integrated channel's profit function need not be concave nor unimodal. For tractability, assume the integrated channel solution is well behaved, i.e., $\Pi(q, e)$ is unimodal and maximized with finite arguments. (For instance, if $S(q, e)$ increases sufficiently quickly with e and $g(e)$ is not sufficiently convex, then infinite effort could be optimal, which is rather unrealistic.) Let q^{o} and e^{o} be the optimal order quantity and effort.

The optimal effort for a given order quantity, $e^{\mathrm{o}}(q)$, maximizes the supply chain's revenue net effort cost. That occurs when

$$\frac{\partial \Pi(q, e^{\mathrm{o}}(q))}{\partial e} = p\,\frac{\partial S(q, e^{\mathrm{o}}(q))}{\partial e} - g'(e^{\mathrm{o}}(q)) = 0. \tag{4.1}$$

With a buyback contract the retailer's profit function is

$$\pi_{\mathrm{r}}(q, e, w_{\mathrm{b}}, b) = (p - b)S(q, e) - (w_{\mathrm{b}} - b)q - g(e).$$

For all $b > 0$ it holds that

$$\frac{\partial \pi_{\mathrm{r}}(q, e, w_{\mathrm{b}}, b)}{\partial e} < \frac{\partial \Pi(q, e)}{\partial e}. \tag{4.2}$$

Thus, e^{o} cannot be the retailer's optimal effort level when $b > 0$. But $b > 0$ is required to coordinate the retailer's order quantity, so it follows that the buyback contract cannot coordinate in this setting.

With a quantity-flexibility contract the retailer's profit function is

$$\pi_{\mathrm{r}}(q, e, w_{\mathrm{q}}, \delta) = pS(q, e) - w_{\mathrm{q}}\left(q - \int_{(1-\delta)q}^{q} F(y \mid e)\,\mathrm{d}y\right) - g(e).$$

For all $\delta > 0$ (which is required to coordinate the retailer's quantity decision)

$$\frac{\partial \pi_{\mathrm{r}}(q, e, w_{\mathrm{q}}, \delta)}{\partial e} < \frac{\partial \Pi(q, e)}{\partial e}.$$

As a result, the retailer chooses a lower effort than optimal, i.e., the quantity-flexibility contract also does not coordinate the supply chain in this setting.

The revenue-sharing and sales-rebate contracts fare no better. It can be shown with $\phi < 1$,

$$\frac{\partial \pi_{\mathrm{r}}(q, e, w_{\mathrm{r}}, \phi)}{\partial e} < \frac{\partial \Pi(q, e)}{\partial e},$$

and so the retailer's optimal effort is lower than the supply chain's. With the sales-rebate contract it can be shown for $r > 0$ and $q > t$,

$$\frac{\partial \pi_r(q, e, w_s, r, t)}{\partial e} > \frac{\partial \Pi(q, e)}{\partial e},$$

which means the retailer exerts too much effort. Although the sales-rebate contract does not coordinate on its own, Taylor (2000) demonstrates it can coordinate the channel if it is combined with a buyback contract: the buyback reduces the retailer's incentive to exert effort, which counteracts the retailer's excessive incentive to exert effort with a sales rebate alone. However, four parameters make for a complex contract. Krishnan et al. (2001) also study the combination of a sales-rebate contract with buybacks. However, they allow the retailer to choose effort after observing demand.

The last contract to consider is the quantity discount. As in the price-setting model, coordination can be achieved in the effort model by letting the retailer earn the entire reward for exerting effort, which is the revenue function, because the retailer is charged the entire cost of effort. Therefore, the quantity discount should let the retailer retain the revenues but charge a marginal cost based on expected revenue conditional on the optimal effort. To explain, suppose the transfer payment is $T_d(q) = w_d(q)q$, where

$$w_d(q) = (1 - \lambda)p\left(\frac{S(q, e^o)}{q}\right) + \lambda c + (1 - \lambda)\frac{g(e^o)}{q}$$

and $\lambda \in [0, 1]$. Given that $S(q, e^o)/q$ is decreasing in q, this is indeed a quantity-discount schedule. As already mentioned, it is almost surely not the only coordinating quantity discount.

The retailer's profit function with the quantity-discount contract is

$$\pi_r(q, e) = pS(q, e) - (1 - \lambda)pS(q, e^o) - \lambda cq - g(e) + (1 - \lambda)g(e^o).$$

As in the price-dependent newsvendor, the retailer chooses the supply chain optimal effort because the retailer keeps all realized revenue. Given the optimal effort e^o, the retailer's profit function is

$$\pi_r(q, e^o) = \lambda pS(q, e^o) - \lambda cq - \lambda g(e^o) = \lambda \Pi(q, e^o),$$

and so the retailer's optimal order quantity is q^o, any allocation of profit is feasible and the supplier's optimal production is q^o.

This approach is sufficiently powerful that it is quite easy to design a quantity-discount contract that coordinates the newsvendor with demand-dependent on price and effort:

$$w_{\mathrm{d}}(q) = (1-\lambda)p^{\mathrm{o}}\left(\frac{S(q,e^{\mathrm{o}})}{q}\right) + \lambda c + (1-\lambda)\frac{g(e^{\mathrm{o}})}{q}.$$

Again, the retailer retains all revenue and so optimizes price and effort. Furthermore, the quantity decision is not distorted because the quantity-discount schedule is contingent on the optimal price and effort, and not on the chosen price and effort.

4.2 Discussion

Coordination with the effort-dependent demand model is complex when the firms are not allowed to contract on the retailer's effort level directly, i.e., any contract that specifies an effort level for the retailer is either unverifiable or unenforceable. Buybacks, revenue-sharing, quantity-flexibility and sales-rebate contracts all fail to coordinate the retailer's action because they all distort the retailer's marginal incentive to exert effort. [That distortion occurs even if the retailer chose its effort after observing a demand signal, as in Krishnan et al. (2001).] The quantity-discount contract does coordinate this system because the retailer incurs the entire cost of effort but also receives the entire benefit of effort.

A number of papers in the marketing and franchising literatures elaborate on the basic retail effort model. For example, in Chu and Desai (1995) the supplier can also exert costly effort to increase demand, e.g., brand building advertising, but the impact of effort occurs only with a lag: they have a two-period model and period 1 effort by the supplier increases only period 2 demand. They also enrich the retailer's effort model to include two types of effort, effort to increase short-term (i.e., current period) sales and long-term effort to increase long-term customer satisfaction and demand (i.e., period 2 sales). They allow the supplier to compensate the retailer by paying a portion of his effort cost and/or by paying the retailer based on the outcome of his effort, i.e., a bonus for high customer satisfaction scores. The issue is the appropriate mix between the two types of compensation. Lal (1990) also includes supplier effort, but, effort again is nonenforceable. Although revenue sharing (in the form of a royalty payment) continues to distort the retailer's effort decision, it provides a useful incentive for the supplier to exert effort: the supplier will not exert effort if the supplier's profit does not depend directly on retail sales. Lal (1990) also considers a model with multiple retailers and horizontal spillovers: the demand-enhancing effort at one retailer may increase the demand at other retailers. These spillovers can lead to free riding, i.e., one retailer enjoys higher demand due to the efforts of others without exerting his own effort. He suggests that the franchisor can control the problem of free riding by exerting costly monitoring effort and penalizing franchisees that fail to exert sufficient effort.

While the models mentioned so far have effort-increasing demand, effort can make other supply chain improvements. Two are discussed: effort to reduce hazardous material consumption and effort to improve product quality.

Corbett and DeCroix (2001) study shared-savings contracts between a supplier of a hazardous material and a manufacturer that uses the material in his output. They assume the product is an indirect material, i.e., the manufacturer's revenue is not correlated with the amount of the product used. For example, an automobile manufacturer does not earn more revenue if it increases the amount of paint used on its vehicle (assuming the increased amount of paint provides no perceived quality improvement). However, with a traditional contract the supplier's revenue does depend on the amount of material used, e.g., the paint supplier's revenue is proportional to the amount of paint the manufacturer purchases. They also assume both the manufacturer and the supplier can exert costly effort to reduce the needed amount of material to produce each unit of the manufacturer's output. The manufacturer clearly has an incentive to exert some effort, since using less material reduces his procurement cost. But the supplier certainly does not have an incentive to reduce the manufacturer's consumption if the supplier's revenue is increasing in consumption. However, it is also quite plausible the supplier's effort would reduce consumption, and further, the supplier may even be more efficient at reducing consumption than the manufacturer (i.e., the supplier's effort cost to reduce consumption by a fixed amount is lower than the manufacturer's effort cost). Thus the supply chain optimal effort levels may very well have both firms exerting effort to reduce consumption.

The Corbett and DeCroix (2001) model adds several twists to the newsvendor model with effort-dependent demand: both firms can exert effort, as opposed to just one firm; and effort hurts one firm and helps the other, whereas in the newsvendor model both firms benefit from effort. Given that structure it is no longer possible to assign all of the costs and all of the benefits of effort to one firm (as the quantity-discount contract does in the newsvendor model). Hence, they show shared-savings contracts (which are related to revenue-sharing contracts) do not coordinate their supply chain, and unfortunately, they are unable to identify which contract would coordinate their model.

Several papers study how effort in a supply chain influences quality. In Reyniers and Tapiero (1995) there is one supplier and one buyer. The supplier can choose between two production processes, one that is costly but produces high quality (in the form of a low-defect probability) and one that is inexpensive but produces low quality (a high-defect probability). The choice of production process can be taken as a proxy for effort in this model. The buyer can test each unit the supplier delivers, but testing is costly. Defective units that are discovered via testing are repaired for an additional cost incurred by the supplier, i.e., an internal failure cost. If the buyer does not test and the unit is defective, then an external failure cost is incurred by the buyer. They allow a contract that includes a wholesale-price rebate for internal failures and an external failure compensation, i.e., the supplier pays the buyer a portion of the buyer's external failure cost. Internal failures are less costly to the supplier (repair cost plus rebate cost) than external failures (compensation to the buyer), so the supplier benefits if the buyer tests a higher fraction of units.

In Baiman, Fischer and Rajan (2000) there is a supplier that can exert costly effort to improve quality and a buyer that exerts testing effort that yields an imperfect signal of quality. Both effort levels are continuous variables, as opposed to the discrete effort levels in Reyniers and Tapiero (1995). If testing suggests the product is defective, the buyer incurs an internal failure cost. If the testing suggests the product is not defective (and hence the buyer accepts the product) then an external failure cost is incurred if the product is in fact defective. They show that optimal supply chain performance is achievable when both effort levels are contractible. Optimal performance is also possible if the firms can verify the external and internal failures and therefore commit to transfer payments based on those failures. Baiman, Fischer and Rajan (2001) extend their model to include the issue of product architecture. With modular design the firms can attribute external failures to a particular firm: either the supplier made a defective component or the supplier made a good component but the buyer caused a defect by poor handling or assembly. However, with an integrated design it is not possible to attribute blame for a product's failure. Hence, the product architecture influences the contract design and supply chain performance. See Novak and Tayur (2002) for another model and empirical work on the issue of effort and attributing responsibility for quality failures.

The quality literature suggests that firms that cannot contract on effort directly can contract on a proxy for effort (the frequency of internal or external failures), which is one solution around potential observability or verifiability problems. See Holmstrom (1979) for another model with moral hazard and effort signals.

Gilbert and Cvsa (2000) study a model with costly effort that is observable but not verifiable, i.e., the firms in the supply chain can observe the amount of effort taken, but the amount of effort taken is not verifiable to the courts, and therefore not contractible. This distinction can be important, as it is in their model. They have a supplier that sets a wholesale price and a buyer that can invest to reduce his marginal cost. The investment to reduce the marginal cost is observed by both firms before the supplier chooses the wholesale price. The buyer cannot fully capture the benefit of cost reduction because the supplier will adjust her wholesale price based on the observed effort. Hence, the buyer invests in less effort to reduce cost than optimal. The supplier can do better if the supplier commits to a wholesale price before observing the buyer's cost reduction. However, demand is random in their model (and observed at the same time the buyer's cost reduction is observed) and so it is beneficial to choose the wholesale price after observing demand, which is in conflict with the incentive benefits of a committed wholesale price. They demonstrate that a hybrid solution works well: the supplier commits to a wholesale-price ceiling before observing the buyer's effort and the demand realization, and after the observations the supplier chooses a wholesale price that is not greater than her wholesale-price ceiling, i.e., there is partial wholesale-price commitment and partial flexibility.

5 Coordination with multiple newsvendors

This section considers two models with one supplier and multiple competing retailers. The first model has a fixed retail price and competition occurs by allocating demand among the retailers proportional to their inventory. It is shown the retailers are biased toward ordering more inventory than optimal because of a demand-stealing effect: each retailer fails to account for the decrease in the other retailer's demands when the retailer increases his order quantity. As a result, with just a wholesale-price contract the supplier can coordinate the supply chain and earn a positive profit. Nevertheless, there are limitations to that coordinating contract: it provides for only one division of supply chain profit; and it is not even the supplier's optimal wholesale-price contract. A buyback contract does not share those limitations: with a buyback contract the supplier can coordinate the supply chain and earn more than with the optimal wholesale-price contract.

The second model with retail competition yields qualitatively very different results. In that model the following sequence of events occurs: the retailers order inventory, market demand is observed and then the market-clearing price is set. The market-clearing price is the price at which consumers are willing to purchase all of the retailers' inventory. Hence, retailers might incur a loss on each unit when the market demand realization is low. The retailers anticipate that possibility and respond by ordering less than the optimal amount. As a result, in contrast to the quantity-allocation competition of the first model, now the supplier needs an instrument to increase retail inventory. Two are considered: resale price maintenance and buyback contracts. With either one the supplier can coordinate the supply chain and extract all of the supply chain's profit.

5.1 Competing newsvendors with a fixed retail price

Take the single retailer newsvendor model (described in Section 2) and make the following modifications: set $c_r = g_r = g_s = v = 0$, increase the number of retailers to $n>1$; interpret D as the total retail demand, let F continue to be the distribution function for D; and let retail demand be divided between the n firms proportional to their stocking quantity, i.e., retailer i's demand, D_i, is

$$D_i = \left(\frac{q_i}{q}\right)D,$$

where $q=\sum_{i=1}^{n} q_i$ and $q_{-i}=q-q_i$. See Wang and Gerchak (2001) for a model that implements proportional allocation with deterministic demand.

Demands at the retailers are perfectly correlated with the proportional allocation model. Hence, either every firm has excess demand (when $D>q$) or every firm has excess inventory (when $D<q$). That could be a reasonable model when customers have low search costs; a customer that desires a unit finds a unit if there is a unit in the system. That search need not involve an actual physical inspection of each store by every customer. For example, information regarding availability could be exchanged among customers through incidental social interactions that naturally occur with daily activities. The model also presumes consumers do not care from which retailer they make their purchases, i.e., there are no retail brand preferences.

There are other demand-allocation models that maintain a constant retail price. Parlar (1988), Karjalainen (1992) and Anupindi and Bassok (1999) and Anupindi, Bassok and Zemel (1999) assign independent random demands to each retailer and then redistributed the retailer's excess demand. Netessine and Shumsky (2001) also redistribute each retailer's excess demand, but they add a twist. In their model the retailers are actually airlines and they have two fare classes. Lippman and McCardle (1997) adopt a more general approach to demand allocation which includes the independent random demand model. They represent aggregate demand as a single random variable and then allocate demand using a splitting rule that depends on the realization of demand (and not on the retailer's order quantities). The retailer's excess demand is then redistributed. In those models some of the redistributed demand may be lost, i.e., some customers may not be willing to continue their shopping if the first retailer they visit has no stock. As a result, total sales depend on the retailers' total inventory and how inventory is distributed among the retailers. With the proportional allocation model industry sales does not depend on the distribution of inventory among the retailers. Thus, the proportional allocation model is simpler to analyze.[21] However, the proportional allocation model is not a special case of the allocation model adopted in Lippman and McCardle (1995). Nevertheless, the qualitative insights from the models are consistent.

[21] The allocation of demand across multiple retailers is analogous to the allocation of demand across a set of products, which is known as the assortment problem. That problem is quite complex. See Mahajan and van Ryzin (1999) for a review of that literature.

There are also models that allocate demand dynamically. In Gans (2002) customers search among retailers without having perfect knowledge of the retailers' stocking levels.[22] As with Lippman and McCardle (1995), Gans (2002) does not consider channel coordination. van Ryzin and Mahajan (1999) assume customers may have different preferences for the retailers and choose to purchase from their most preferred retailer that has stock available.

Given the proportional allocation rule, the integrated supply chain faces a single newsvendor problem. Hence the optimal order quantity is defined by the familiar

$$F(q^{o}) = \frac{p-c}{p}. \tag{5.1}$$

Because the integrated solution remains a single-location newsvendor problem, the multiple retailer model with proportional allocation is a nice generalization of the single retailer model.

In the decentralized system we want to investigate retail behavior with either a wholesale-price contract or a buyback contract. (Since the retail price is fixed, in this case there exists a revenue-sharing contract that is equivalent to the buyback contract.) Retailer i's profit function with a buyback contract is

$$\pi_i(q_i, q_{-i}) = (p-w)q_i - (p-b)\left(\frac{q_i}{q}\right)\int_0^q F(x)\,\mathrm{d}x.$$

The above also provides the retailer's profit with a wholesale-price contract (i.e., set $b=0$). The second-order condition confirms each retailer's profit function is strictly concave in his order quantity. Hence, there exists an optimal order quantity for retailer i for each q_{-i}. In game theory parlance, retailer i has a unique optimal response to the other retailers' strategies (i.e., their order quantities). Let $q_i(q_{-i})$ be retailer i's response function, i.e., the mapping between q_{-i} and retailer i's optimal response. Since the retailers have symmetric profit functions, $q_j(q_{-j}) = q_i(q_{-i})$, $i \neq j$.

A set of order quantities, $\{q_1^*, \ldots, q_n^*\}$, is a Nash equilibrium of the decentralized system if each retailer's order quantity is a best response, i.e., for all i, $q_i^* = q_i(q_{-i}^*)$, where $q_{-i}^* = q^* - q_i^*$ and $q^* = \sum_{j=1}^{n} q_j^*$. There may not exist an Nash equilibrium, or there may be multiple Nash equilibria. If there is a unique Nash equilibrium then that is taken to be the predicted outcome of the decentralized game.

[22] Gans (2002) presents his model more generally. He has multiple suppliers of a service that compete on some dimension of quality, say their fill rate. Nevertheless, customers have less than perfect information about the suppliers' service levels and so they must develop a search strategy. The suppliers compete knowing customers have limited information.

Any Nash equilibrium must satisfy each retailer's first-order condition:

$$\frac{\partial \pi_i(q_i, q_j)}{q_i} = q^* \left(\frac{p-w}{p-b}\right) - q_i^* F(q^*) - q_{-i}^* \left(\frac{1}{q^*} \int_0^{q^*} F(x)\,\mathrm{d}x\right) = 0.$$

Substitute $q_{-i}^* = q^* - q_i^*$ into the above equation and solve for q_i^* given a fixed q^*:

$$q_i^* = q^* \frac{\left((p-w)/(p-b) - (1/q^*)\int_0^{q^*} F(x)\,\mathrm{d}x\right)}{F(q^*) - 1/q^* \int_0^{q^*} F(x)\,\mathrm{d}x}. \tag{5.2}$$

The above gives each retailer's equilibrium order conditional on q^* being the equilibrium total order quantity. Hence, the above describes an equilibrium only if $q^* = nq_i^*$. Substitute Eq. (5.2) into $q^* = nq_i^*$ and simplify:

$$\frac{1}{n} F(q^*) + \left(\frac{n-1}{n}\right)\left(\frac{1}{q^*}\int_0^{q^*} F(x)\,\mathrm{d}x\right) = \frac{p-w}{p-b}. \tag{5.3}$$

The left-hand side of Eq. (5.3) is increasing in q^* from 0 (when $q^*=0$) to 1 (when $q^*=\infty$). Hence, when $b<w<p$, there exists a unique q^* that satisfies Eq. (5.3). In other words, in this game there exists a unique Nash equilibrium in which the total order quantity, q^*, is implicitly given by Eq. (5.3) and each retailer's order quantity equals $q_i^* = q^*/n$.

Consider how the equilibrium order quantity changes in n. The left-hand side of Eq. (5.3) is decreasing in n. Hence, q^* is increasing in n for fixed contractual terms: a single retailer that faces market demand D purchases less than multiple retailers facing the same demand (with proportional allocation). [This effect generalizes beyond just the proportional allocation model, as demonstrated by Lippman and McCardle (1995).] Competition makes the retailers order more inventory because of the demand-stealing effect: each retailer ignores the fact that ordering more means the other retailers' demands stochastically decrease. Anupindi and Bassok (1999) and Mahajan and van Ryzin (2001) also noticed this effect. However, the effect does not apply universally: Netessine and Rudi (2000b) find that competition may lead some retailers to understock when there are more than two retailers and demands are not symmetric. Furthermore, if retailers sell complements, rather than substitutes, then the demand-stealing effect is reversed: each retailer tends to understock because it ignores the additional demand it creates for other retailers.

Due to the demand-stealing effect the supplier can coordinate the supply chain and earn a positive profit with just a wholesale-price contract. To explain, let $\hat{w}(q)$ be the wholesale price that induces the retailers to order q units with a wholesale-price contract (i.e., with $b=0$). From Eq. (5.3),

$$\hat{w}(q) = p\left(1 - \left(\frac{1}{n}\right)F(q) - \left(\frac{n-1}{n}\right)\left(\frac{1}{q}\int_0^q F(x)\,\mathrm{d}x\right)\right).$$

By definition $\hat{w}(q^o)$ is the coordinating wholesale price. Given that $F(q^o) = (p-c)/c$ and

$$\frac{1}{q}\int_0^q F(x)\,\mathrm{d}x < F(q),$$

it can be shown that $\hat{w}(q^o) > c$ when $n>1$. Hence, the supplier earns a positive profit with that coordinating contract. With the single retailer model channel coordination is only achieved when the supplier earns zero profit, i.e., marginal cost pricing, $\hat{w}(q^o) = c$.

Although the supplier can use the wholesale-price contract to coordinate the supply chain, that contract is not optimal for the supplier. The supplier's profit function with a wholesale-price contract is

$$\pi_s(q, \hat{w}(q)) = q(\hat{w}(q) - c).$$

Assuming $n>1$, differentiate $\pi_s(q, \hat{w}(q))$ with respect to q and evaluate at $\hat{w}(q^o)$, the coordinating wholesale price,

$$\frac{\partial \pi_s(q^o, \hat{w}(q^o))}{\partial q} = -\frac{q^o p f(q^o)}{n} < 0.$$

Hence, rather than coordinating the supply chain with the wholesale price $\hat{w}(q^o)$, the supplier prefers to charge a higher wholesale price and sell less than q^o when $n>1$.

Although the supplier does not wish to use a wholesale-price contract to coordinate the supply chain, it is possible the supplier's profit with a coordinating buyback contract may exceed her profit with the optimal wholesale-price contract. Let $w_b(b)$ be the wholesale price that coordinates the supply chain given the buyback rate. Since the buyback rate provides an incentive to the retailers to increase their order quantity, it must be that $w_b(b) > \hat{w}(q^o)$, i.e., to coordinate the supply chain the supplier must use a wholesale price that is higher than the coordinating wholesale-price contract, which, recall, is lower than the supplier's optimal wholesale price. From Eqs. (5.1) and (5.3)

$$w_b(b) = p - (p-b)\left[\frac{1}{n}\left(\frac{p-c}{p}\right) + \left(\frac{n-1}{n}\right)\left(\frac{1}{q^o}\int_0^{q^o} F(x)\,dx\right)\right].$$

Given that $q_i^* = q^*/n$, retailer i's profit with a coordinating buyback contract is

$$\begin{aligned}\pi_i(q_i^*, q_{-i}^*) &= (p - w(b))q^o/n - (p-b)\left(\frac{1}{n}\right)\int_0^{q^o} F(x)\,dx \\ &= \left(\frac{p-b}{pn^2}\right)q^o\left[p - c - \frac{p}{q^o}\int_0^{q^o} F(x)\,dx\right] \\ &= \left(\frac{p-b}{pn^2}\right)\Pi(q^o)\end{aligned}$$

The supplier's profit with the coordinating contract is

$$\begin{aligned}\pi_s(q^o, w_b(b), b) &= \Pi(q^o) - n\pi_i(q_i^*, q_{-i}^*) \\ &= \left(\frac{p(n-1)+b}{pn}\right)\Pi(q^o).\end{aligned}$$

Hence, the supplier can extract all supply chain profit with $b = p$. As shown earlier, the coordinating contract with $b = 0$ provides a lower bound for the supplier's profit (because the supplier could do even better with a higher wholesale price than $w_b(0)$). The ratio of the supplier's lower bound to the supplier's maximum profit, $\Pi(q^o)$, provides a measure of how much improvement is possible by using a coordinating buyback contract:

$$\frac{\pi_s(q^o, w_b(0), 0)}{\Pi(q^o)} = \frac{n-1}{n}.$$

Hence, as n increases the supplier's potential gain decreases from using a coordinating buyback contract rather than her optimal wholesale-price contract. In fact, the supplier can use a wholesale-price contract to capture most of the supply chain's optimal profit with a relatively few number of retailers: for $n = 5$ the supplier captures at least 80% of the optimal profit and for $n = 10$ the supplier captures at least 90%. Mahajan and van Ryzin (2001) also observe that downstream competition can mitigate the need for coordinating contracts.

5.2 Competing newsvendors with market-clearing prices

In the previous model retail competition influences the allocation of demand. In this model, first analyzed by Deneckere, Marvel and Peck (1997), competition influences the retail price. Specifically, the market price depends on the realization of demand and the amount of inventory purchased.

Suppose industry demand can take on one of two states, high or low. Let q be the retailers' total order quantity. When demand is low, the market-clearing price is

$$p_l(q) = (1 - q)^+$$

and when demand is high the market-clearing price is

$$p_h(q) = \left(1 - \frac{q}{\theta}\right)^+,$$

for $\theta > 1$. Suppose either demand state is equally likely.

There is a continuum of retailers, indexed on the interval $[0, 1]$. Retailers must order inventory from a single supplier before the realization of the demand is observed. After demand is observed the market-clearing price is determined. Perfect competition is assumed, which means the retailers continue to order inventory until their expected profit is zero. Leftover inventory has no salvage value and the supplier's production cost is zero. Deneckere et al. (1997) show the qualitative insights from this model continue to hold with a continuous demand state space, a general supplier cost function and a general demand function. In another paper, Deneckere, Marvel and Peck (1996) show the qualitative insights also hold if the retailers choose their prices before the realization of demand. (In that model all demand is allocated to the retailer with the lowest price, and any residual demand is subsequently allocated to the retailer with the second lowest price, etc.)

To set a benchmark, suppose a single monopolist controls the entire system. In this situation the monopolist can choose how much of her inventory to sell on the market after demand is observed. At that point the cost of inventory is sunk, so the monopolist maximizes revenue: in the low-demand state the monopolist sells $q = 1/2$ at price $p_l(1/2) = 1/2$; and in the high-demand state $q = \theta/2$ with the same price, $p_h(\theta/2) = 1/2$. So the inventory order should be one of those two quantities. Given the production cost is zero, ordering $\theta/2$ units is optimal.[23] Furthermore, the monopolist sells her

[23] The monopolist is actually indifferent between $\theta/2$ and a greater amount. A positive production cost would eliminate that result. But that is not an interesting issue. The important result is that the supply chain optimal order quantity is the greater amount.

entire stock in the high-demand state, but in the low-demand state the monopolist does not sell some of her inventory. The monopolist's expected profit is

$$\Pi^{o} = (1/2)p_l(1/2)(1/2) + (1/2)p_h(\theta/2)(\theta/2) = \frac{1+\theta}{8}.$$

Now consider the system in which the supplier sells to the perfectly competitive retailers with just a wholesale-price contract. The retailer's expected profit is

$$\frac{1}{2}p_l(q)q + \frac{1}{2}p_{\mathrm{h}}(q)q - wq = \begin{cases} \frac{1}{2}q(2-q-q/\theta) - wq & q \le 1 \\ \frac{1}{2}q(1-q/\theta) - wq & q > 1 \end{cases}.$$

Let $q_1(w)$ be the quantity that sets the above profit to zero when $q \le 1$, which is the equilibrium outcome due to perfect competition:

$$q_1(w) = \frac{2\theta}{1+\theta}(1-w).$$

For $q_1(w) \le 1$ to hold, it must be that $w \ge (1/2) - 1/(2\theta)$. Let $q_2(w)$ be the quantity that sets the above profit to zero when $q > 1$,

$$q_2(w) = \theta(1-2w).$$

For $q_2(w) > 1$ to hold it must be that $w < (1/2) - 1/(2\theta)$.

Let $\pi_s(w)$ be the supplier's profit. From the above results,

$$\pi_s(w) = \begin{cases} q_1(w)w & w \ge (1/2) - 1/(2\theta) \\ q_2(w)w & \text{otherwise} \end{cases}.$$

Let $w^*(\theta)$ be the supplier's optimal wholesale price:

$$w^*(\theta) = \begin{cases} \frac{1}{2} & \theta \le 3 \\ \frac{1}{4} & \text{otherwise} \end{cases}$$

and

$$\pi_s(w^*(\theta)) = \begin{cases} \dfrac{\theta}{2(1+\theta)} & \theta \le 3 \\ \frac{1}{8}\theta & \text{otherwise} \end{cases}.$$

So when $\theta \leq 3$ the retailers order

$$q_1(w^*(\theta)) = \frac{\theta}{1+\theta}$$

and the market-clearing prices are

$$p_l(q_1(w^*(\theta))) = \frac{1}{1+\theta}, \qquad p_h(q_1(w^*(\theta))) = \frac{\theta}{1+\theta}.$$

When $\theta > 3$ the retailers order

$$q_2(w^*(\theta)) = \frac{\theta}{2}$$

and the market-clearing prices are

$$p_l(q_2(w^*(\theta))) = 0, \qquad p_h(q_2(w^*(\theta))) = \frac{1}{2}.$$

No matter the value of θ $\pi_s(w^*(\theta)) < \Pi^o$, so the supplier does not capture the maximum possible profit with a wholesale-price contract. When $\theta \leq 3$ the supplier falls short because the retailers sell too much in the low-demand state. To mitigate those losses the retailers order less than the optimal quantity, but then they are unable to sell enough in the high-demand state. When $\theta > 3$ the supplier falls short because the retailers sell too much in the low-demand state even though they sell the optimal amount in the high-demand state. Hence, in either case the problem is that competition leads the retailers to sell too much in the low-demand state. Recall, the monopolist does not sell all of her inventory in the low-demand state, but the perfectly competitive retailers cannot be so restrained.

To earn a higher profit the supplier must devise a mechanism to prevent the low-demand state market-clearing price from falling below 1/2. In short, the supplier must curtail the destructive competition that results from having more inventory than the system needs. Deneckere et al. (1997) propose the supplier implements resale price maintenance: the retailers may not sell below a stipulated price. [For other research on resale price maintenance, see Ippolito (1991), Shaffer (1991) and Chen (1999b).] Let $\bar{p}$ be that price. When $\bar{p}$ is above the market-clearing price the retailers have unsold inventory, so demand is allocated among the retailers. Assume demand is allocated so that each retailer sells a constant fraction of his order quantity, i.e., proportional allocation.

Given the optimal market-clearing price is always 1/2, the search for the optimal resale price maintenance contract should begin with $\overline{p} = 1/2$.[24] Let $q(t)$ be the order quantity of the tth retailer and let $\pi_r(t)$ be the tth retailer's expected profit. Assume the retailers' total order quantity equals $\theta/2$, i.e.,

$$\int_0^1 q(t)\,\mathrm{d}t = \frac{\theta}{2}. \tag{5.4}$$

Hence, the market price in either demand state is 1/2. We later confirm that the retailers indeed order $\theta/2$ in equilibrium. Evaluate the tth retailer's expected profit:

$$\pi_r(t) = -q(t)w + \frac{1}{2}\left(\frac{1/2}{\theta/2}\, q(t)\right)\overline{p} + \frac{1}{2}\, q(t)\overline{p} :$$

the retailer sells $(1/2)q(t)/(\theta/2)$ in the low-demand state and sells $q(t)$ in the high-demand state. Simplify the above profit:

$$\pi_r(t) = q(t)\left(\frac{1+\theta}{4\theta} - w\right).$$

so the supplier can charge

$$\overline{w} = \frac{1+\theta}{4\theta}.$$

We must now confirm the retailers indeed order $\theta/2$ given that wholesale price. Say the retailers order $1/2 < q < \theta/2$, so the tth retailer's expected profit is

$$-q(t)w + \frac{1}{2}\left(\frac{1/2}{q}\, q(t)\right)\overline{p} + \frac{1}{2}\, q(t)\left(1 - \frac{q}{\theta}\right).$$

The above is decreasing in the relevant interval and equals 0 when the wholesale price is $\overline{w}$. So with the $(\overline{p}, \overline{w})$ resale price maintenance contract the retailers order $q = \theta/2$, the optimal quantity is sold in either state and the retailers' expected profit is zero.[25] Hence, the supplier earns Π^o with that contract.

---

[24] The optimal market-clearing price is independent of the demand realization because the demand model, $q = \theta(1-p)$, has a multiplicative shock, θ. The optimal market-clearing price would differ across states with an additive shock, e.g., if the demand model were $q = \theta - p$. In that case resale price maintenance could only coordinate the supply chain if the resale price were state dependent.

[25] Given $(\overline{p}, \overline{w})$, any $q(t)$ that satisfies Eq. (5.4) is an equilibrium, i.e., there are an infinite number of equilibria. The authors do not specify how a particular equilibrium would be chosen.

Resale price maintenance prevents destructive competition in the low-demand state, but there is another approach to achieve the same objective. Suppose the supplier offers a buyback contract with $b = 1/2$. Since retailers can earn $b = 1/2$ on each unit of inventory, the market price cannot fall below 1/2: for the market price to fall below 1/2 it must be that some retailers are willing to sell below 1/2, but that is not rational if the supplier is willing to give 1/2 on all unsold units. Therefore, the retailers sell at most 1/2 in the low-demand state and $\theta/2$ in the high-demand state.

The retailers' profit with a buyback contract is

$$\frac{1}{2}(p_l(1/2)(1/2) + b(q - 1/2)) + \frac{1}{2}(p_h(q)q) - qw,$$

which simplifies to

$$q\left(\frac{3}{4} - w - \frac{q}{2\theta}\right).$$

(That profit assumes $1/2 < q < \theta/2$.) The supplier wants the retailers to order $q = \theta/2$. From the above equation the retailers earn a zero profit with $q = \theta/2$ when $w = 1/2$. Hence, the supplier maximizes the system's profit with a buyback that offers a full refund on returns.

Although resale price maintenance and the buyback contract achieve the same objective, the supplier sets a higher wholesale price with the buyback contract, i.e., $1/2 > (1 + \theta)/4\theta$: retailers do not incur the cost of excess inventory in the low-demand states with a buyback contract, but they do with resale price maintenance. A buyback contract is also not the same as a revenue-sharing contract in this situation. (Section 2 demonstrates the two contracts are equivalent in the single newsvendor model.) The buyback contract prevents the market-clearing price from falling below 1/2 in the low-demand state, but revenue sharing does not prevent destructive competition: in the low-demand state the retailers have no alternative use for their inventory, so they still attempt to sell all of it in the market. [However, revenue sharing does prevent destructive price competition in Dana and Spier (2001) because in their model the retailers incur a marginal cost for each sale rather than each unit purchased.] Those contracts are also different in the single newsvendor model with price-dependent demand (see Section 3). However, in that model the revenue-sharing contract coordinates the supply chain and the buyback contract does not. The key distinction is that in the single newsvendor model the retailer controls the market price, whereas in this competitive model the retailers do not.

It is interesting that a buyback contract coordinates the supply chain in either competitive model even though in the first one the supplier must discourage the retailers from ordering too much and in the second one the supplier must encourage the retailers to order more. To explain this apparent contradiction, the wholesale-price component of the contract always reduces the retailer's order quantity and the buyback component always increases the retailer's order quantity. Thus, depending on the relative strength of those two components, the buyback contract can either increase or decrease the retailers' order quantities.

5.3 Discussion

Retail competition introduces several challenges for supply chain coordination. There may exist a demand-stealing effect which causes each retailer to order more than the supply chain optimal quantity because each retailer ignores how he reduces his competitors' demand. For coordination the supplier needs to *reduce* the retailers' order quantities, which can be done with just a wholesale-price contract above marginal cost. But that wholesale-price contract only provides for one division of the supply chain's profit, and it is not even the supplier's optimal wholesale-price contract. The supplier can do better with a buyback contract and coordinate the supply chain. However, the incremental improvement over the simpler wholesale-price contract decreases quickly as retail competition intensifies. In contrast to the demand-stealing effect, in the presence of the destructive competition effect the supplier needs to *increase* the retailers' order quantities. This occurs when demand is uncertain and the retail price is set to clear the market. When demand is high the retailers earn a profit, but when demand is low deep discounting to clear inventory leads to losses. The retailers anticipate this problem and respond by curtailing their inventory purchase. Both resale price maintenance and buyback contracts prevent deep discounting, and therefore alleviate the problem.

There are several other papers that study supply chain coordination with competing retailers. Padmanabhan and Png (1997) demonstrate a supplier can benefit by mitigating retail competition with a buyback contract even with deterministic demand and less than perfect retail competition. In their model two retailers first order stock and then choose prices. Retailer i sells $q_i = \alpha - \beta p_i - \gamma p_j$, where α, β and γ are constants and $\beta > \gamma$. With just a wholesale-price contract ($b = 0$), the retailers price to maximize revenue, because their inventory is sunk. When the supplier offers a full returns policy ($b = w$), the retailers price to maximize profit because unsold inventory can be returned for a full refund. The retailers price more aggressively when they are maximizing revenue. They anticipate this behavior when choosing their order quantity, and so order less when they expect more intense price competition.[26] Thus, for any given wholesale price the

[26] When maximizing revenue the retailer chooses a price so that marginal revenue equals zero. When maximizing profit the retailer chooses a price so that marginal revenue equals marginal cost. Thus, the retailer's optimal price is higher when maximizing profit.

retailers order more with the full returns policy. Since demand is deterministic, in neither case does the supplier actually have to accept returns. So the supplier is better off with the full returns policy when demand is deterministic.[27] When demand is stochastic the supplier may not prefer the full returns policy because that policy may induce the retailers to order too much inventory. However, one suspects the supplier could benefit in that situation from a partial return credit, i.e., $b < w$.[28] See Bernstein and Federgruen (1999) for a more complex model with deterministic demand and competing retailers. See Tsay and Agrawal (2000) and Atkins and Zhao (2002) for models with two retailers that compete on price and service.

Several authors have studied coordination when retailers face oligopolistic competition, i.e., they may earn nonzero profit in equilibrium. Even though this is a different type of competition, the demand-stealing effect remains, but establishing the existence and possibly the uniqueness of equilibrium is generally more challenging. Cachon and Lariviere (2000) demonstrate that revenue sharing can coordinate retailers that compete along a single dimension, e.g., quantity-competing retailers or price-competing retailers. Revenue sharing is not successful if the retailers compete both on quantity and price (e.g., a firm's demand depends on its price and possibly its fill rate, for which quantity is taken as a proxy.) However, Dana and Spier (2001) find that revenue sharing does coordinate perfectly competitive price-setting newsvendors. Bernstein and Federgruen (2000) show that a nonlinear form of the price-discount contract coordinates price- and quantity-competing retailers.

Rudi et al. (2001) study a model with two retailers that each face a newsvendor problem. Inventory can be shipped between the retailers for a fee. Those shipments occur after demand is observed, but before demand is lost. Hence, if retailer i has excess inventory and retailer j has excess demand, then some portion of retailer i's excess inventory can be shipped to retailer j to satisfy retailer j's excess demand. At first glance it would appear the redistribution of excess demand that occurs in the Lippman and McCardle (1995) model is qualitatively equivalent to the redistribution of inventory that occurs in this model. One difference is that the firms in the demand redistribution model do not incur an explicit fee for each demand unit moved between the retailers. A second difference is that the firms in the inventory redistribution model control the redistribution process, and so demand is only lost if total demand exceeds total

[27] Padmanabhan and Png (1997) state that the full returns policy helps the manufacturer by *increasing* retail competition. They are referring to the competition in the ordering stage, not in the pricing stage. The retailers order more precisely because they anticipate less aggressive competition in the pricing stage.

[28] In a slightly different model of imperfect competition between two retailers, Butz (1997) demonstrates a buyback contract allows the supplier to coordinate the channel.

inventory (assuming the firms set transfer prices so that Pareto improving trades always occur).

Rudi et al. (2001) demonstrate the retailers may either order too much or too little inventory in this model, depending on the transfer prices for redistributed inventory. When the receiving firm must pay the maximum fee (so he is indifferent between receiving the transfer and incurring a lost sale), the firms order too much inventory. Each firm profits from selling his inventory to the other firm, so each firm is biased toward ordering too much. When the receiving firm pays the minimum fee (so the sending firm is indifferent between salvaging excess inventory and shipping it to the other firm), the firms order too little inventory (neither firm profits from excess inventory, but can depend on some portion of the other firm's inventory). Given these two extremes, there exists a set of intermediate transfer prices such that the firms order the optimal amount of inventory.

Rudi et al. (2001) do not include a supplier in their model. The supplier could be a facilitator of the inventory redistribution. For example, in their model the price at which a firm sells excess inventory is the same as the price at which a firm buys excess inventory. But with a supplier those prices need not be the same: the supplier could buy excess inventory at one price (a buyback) and redistribute at a different price (which could differ from the initial wholesale price). The inclusion of the supplier would also change each retailer's inventory problem. In their model each retailer expects to sell only a portion of his excess inventory, because the other retailer purchases only enough to meet his excess demand. With a buyback contract the supplier stands ready to buy all excess inventory at a fixed price. [Dong and Rudi (2001) do study transshipment with a supplier but they only consider the wholesale-price contract.]

Anupindi et al. (2001) study a general inventory redistribution game with multiple locations. They adopt a 'coopetive' analysis: some decisions are analyzed with concepts from cooperative game theory, whereas others implement noncooperative game theory.

Lee and Whang (2002) have a supplier and free inventory redistribution at an intermediate point in the selling season. [With Rudi et al. (2001) the redistribution occurs at the end of the season, i.e., after a perfect demand signal is received.] In their model the redistribution transfer price is the clearing price of a secondary market rather than a price dictated by the retailers or the supplier. They find that the spot market is advantageous to the supplier for low margin items, but not for high margin items. If the supplier cannot control the spot market, then the supplier can attempt to influence the spot market via minimum order quantity requirements or return policies. For example, a return policy will remove inventory from the spot market and thereby raise its price. This is analogous to the use of buybacks to prevent destructive price competition discussed in Section 5.2.

Gerchak and Wang (1999) and Gurnani and Gerchak (1998) consider supply chains with multiple upstream firms rather than multiple downstream firms. In these assembly systems the upstream firms are different suppliers, each producing a component for the manufacturer's product (the downstream firm). Total production is constrained by the supplier with the smallest output and the excess output of the other suppliers is wasted. Hence, as with destructive competition, the suppliers are biased toward producing too little. They study several contracts that encourage the suppliers to increase their production quantities. Bernstein and DeCroix (2002) also study coordination in assembly systems. They discuss how the organization of the assembly structure influences supply chain performance.

Bernstein et al. (2002) demonstrate that coordination of competing retailers with a wholesale-price contract is easier with Vendor Managed Inventory (the supplier controls the retailers' inventory policies and the retailers choose prices) than with standard operations, i.e., the retailers choose prices and inventory policies. This represents a different approach to coordination: instead of aligning incentive via contracts, the firms transfer decision rights. The pros and cons of this approach relative to formal contracts have not been fully explored.

Throughout this section it has been assumed the supplier is independent of all of the retailers. However, in some markets a supplier may choose to own her own retailer or to sell directly to consumers. Tsay and Agrawal (1999) explore the channel conflicts such a move creates. It has also been assumed that the supplier simultaneously offers a contract to all retailers that is observable by all retailers. However, McAfee and Schwarz (1994) argue that a supplier has an incentive to sequentially offer contracts to retailers and to keep these contractual terms secret. Retailers anticipate this behavior, respond accordingly and thereby destroy the effectiveness of some contracts. See the following papers for additional discussion on this issue: O'Brien and Shaffer (1992), Marx and Shaffer (2001a,b, 2002).

6 Coordinating the newsvendor with demand updating

With the standard newsvendor problem the retailer has only one opportunity to order inventory. However, it is reasonable that the retailer might have a second opportunity to order inventory. Furthermore, the retailer's demand forecast may improve between the ordering epochs (Fisher & Raman, 1996). Hence, all else being equal, the retailer would prefer to delay all ordering to the second epoch. But that creates a problem for the supplier. With a longer lead time the supplier may be able to procure components more cheaply and avoid overtime labor. If the supplier always has sufficient capacity to fill the retailer's order, the supplier prefers an earlier rather than later order commitment by the retailer.[29] Thus, supply chain coordination requires the firms to balance the lower cost of early production with the better information afforded by later production. It is shown that coordination is achieved and profits divided with a buyback contract as long as the supplier is committed to a wholesale price for each order epoch.

6.1 Model and analysis

Based on Donohue (2000), consider a model in which the retailer receives a forecast update only once before the start of the selling season. Let $\xi \geq 0$ be the realization of that demand signal. Let $G(\cdot)$ be its distribution function and $g(\cdot)$ its density function. Let $F(\cdot \mid \xi)$ be the distribution function of demand after observing the demand signal. Demand is stochastically increasing in the demand signal, i.e., $F(x \mid \xi_h) < F(x \mid \xi_l)$ for all $\xi_h > \xi_l$. For convenience, let period 1 be the time before the demand signal and let period 2 be the time between the demand signal and the start of the selling season.

Let q_i be the retailer's total order as of period i, i.e., q_1 is the retailer's period 1 order and $q_2 - q_1$ is the retailer's period 2 order. The retailer's period 2 order is placed after observing the demand signal. Within each period the supplier chooses her production after receiving the retailer's order. Early production is cheaper than later production, so let c_i be the supplier's per unit production cost in period i, with $c_1 < c_2$. The supplier charges the retailer w_i per unit for units ordered in period i. In addition, the supplier offers to buyback all unsold units for b per unit. The supplier offers these terms at the start of period 1 and commits to not change the terms.[30] Let p be the retail price. Normalize to zero the salvage value of leftover inventory and any indirect costs due to lost sales.

The supplier does not have a capacity constraint in either period and delivers stock to the retailer at the end of each period. The supplier operates under voluntary compliance, so the supplier may deliver less than the retailer orders.[31] However, the supplier may also produce more in period 1 than the retailer orders. For simplicity, there is no holding cost on inventory carried from period 1 to period 2.

[29] This preference is not due to the time value of income, i.e., the supplier prefers an early order even if the retailer pays only upon delivery.

[30] There may be some incentive to alter the terms after the demand signal is received. Suppose the news is good. In that case the supplier may prefer to leave the retailer with a smaller fraction of supply chain profit (if the retailer has a constant minimum acceptable profit) or the supplier may argue that it deserves a larger fraction of the profit as a reward for producing a good product. There is a large literature in economics on renegotiation and its impact on contract design (see Tirole 1986; Demougin 1989; Holden, 1999).

[31] Donohue (2000) assumes forced compliance. She also assumes the supplier offers a buyback contract.

All information is common knowledge. For example, both firms know $F(\cdot \mid \xi)$ as well as $G(\cdot)$. In particular, both firms observe the demand signal at the start of period 2.[32] Both firms seek to maximize expected profit.

Begin with period 2. Let $\Omega_2(q_2 \mid q_1, \xi)$ be the supply chain's expected revenue minus the period 2 production cost:

$$\Omega_2(q_2 \mid q_1, \xi) = pS(q_2 \mid \xi) - c_2 q_2 + c_2 q_1. \tag{6.1}$$

Let $q_2(q_1, \xi)$ be the supply chain's optimal q_2 given q_1 and ξ. Let $q_2(\xi) = q_2(0, \xi)$, i.e., $q_2(\xi)$ is the optimal order if the retailer has no inventory at the start of period 2. Given $\Omega_2(q_2 \mid q_1, \xi)$ is strictly concave in q_2,

$$F(q_2(\xi) \mid \xi) = \frac{p - c_2}{p}. \tag{6.2}$$

$q_2(\xi)$ is increasing in ξ, so it is possible to define the function $\xi(q_1)$ such that

$$F(q_1 \mid \xi(q_1)) = \frac{p - c_2}{p}. \tag{6.3}$$

$\xi(q_1)$ partitions the demand signals into two sets: if $\xi > \xi(q_1)$ then the optimal period 2 order is positive, otherwise it is optimal to produce nothing in period 2.

The retailer also faces in period 2 a standard newsvendor problem, with the modification that the retailer may already own some stock. Let $\pi_2(q_2 \mid q_1, \xi)$ be the retailer's expected revenue minus period 2 procurement cost,

$$\pi_2(q_2 \mid q_1, \xi) = (p - b)S(q_2 \mid \xi) - (w_2 - b)q_2 + w_2 q_1,$$

where assume the supplier delivers the retailer's order in full. (To simplify notation the contract parameters are not included in the arguments of the functions considered in this section.) To coordinate the retailer's period 2 decision, choose contract parameters with $\lambda \in [0, 1]$ and

$$p - b = \lambda p$$
$$w_2 - b = \lambda c_2$$

[32] The supplier does not have to observe the signals directly if the supplier knows $F(\cdot \mid \xi)$. In that case the supplier can infer ξ from the retailer's order quantity because $F(\cdot \mid \xi)$ is strictly decreasing in ξ. See Brown (1999) for a model in which the upstream firm is not able to use the downstream's order quantity to exactly infer the downstream firm's demand signal.

Not surprisingly, those parameters are analogous to the coordinating buyback parameters in the single-period newsvendor model. With any of those contracts

$$\pi_2(q_2 \mid q_1, \xi) = \lambda(\Omega_2(q_2 \mid q_1, \xi) - c_2 q_1) + w_2 q_1.$$

Thus, $q_2(q_1, \xi)$ is also the retailer's optimal order, i.e., the contract coordinates the retailer's period 2 decision.

Now consider whether the supplier indeed fills the retailer's entire period 2 order. Let x be the total inventory in the supply chain at the start of period 2, $x \geq q_1$. The supplier's inventory at the start of period 2 is $x - q_1$. Let y be the inventory at the retailer after the supplier's delivery in period 2. The supplier completely fills the retailer's order when $y = q_2$. The supplier clearly delivers the retailer's full order when $x \geq q_2$ because there is no reason to partially fill the retailer's order and have leftover inventory. If $q_2 > x$, the supplier must produce additional units to deliver the retailer's complete order. Let $\Pi_2(y \mid x, q_1, \xi)$ be the supplier's profit, where $x \leq y \leq q_2$,

$$\begin{aligned}\Pi_2(y \mid x, q_1, q_2, \xi) &= bS(y \mid \xi) - by + w_2(y - q_1) - (y - x)c_2 \\ &= (1 - \lambda)(\Omega_2(y \mid q_1, \xi) - c_2 q_1) + c_2 x - w_2 q_1\end{aligned}$$

where the above follows from the contract terms, $w_2 = \lambda c_2 + b$. Given $q_2 > x$, the supplier fills the retailer's order entirely as long as $q_2 \leq q_2(q_1, \xi)$, i.e., the supplier does not satisfy the retailer if the retailer happens to irrationally order too much. Therefore, in period 2 the retailer orders the supply chain optimal quantity and the supplier fills the order entirely, even with voluntary compliance and no matter how much inventory the supplier carries between periods.

In period 1, assuming a coordinating $\{w_2, b\}$ pair is chosen, the retailer's expected profit is

$$\pi_1(q_1) = -(w_1 - w_2 + \lambda c_2)q_1 + \lambda E[\Omega_2(q_2(q_1, \xi) \mid q_1, \xi)].$$

The supply chain's expected profit is

$$\Omega_1(q_1) = -c_1 q_1 + E[\Omega_2(q_2(q_1, \xi) \mid q_1, \xi)].$$

Choose w_1 so that

$$w_1 - w_2 + \lambda c_2 = \lambda c_1$$

because then

$$\pi_1(q_1) = \lambda \Omega_1(q_1).$$

It follows that the retailer's optimal order quantity equals the supply chain's optimal order quantity, q_1^o, and any portion of the supply chain's profit can be allocated to the retailer. Given $\Omega_1(q_1)$ is strictly concave, q_1^o satisfies:

$$\begin{aligned}\frac{\partial\Omega_1(q_1^o)}{\partial q_1} &= -c_1 + c_2(1 - G(\xi(q_1^o))) + \int_0^{\xi(q_1^o)} pS'(q_1^o \mid \xi)g(\xi)\,d\xi \\ &= 0. \end{aligned} \tag{6.4}$$

With centralized operations it does not matter whether inventory is left at the supplier in period 1 because the supply chain moves all inventory to the retailer in period 2: inventory at the supplier has no chance of selling. With decentralized control supply chain coordination is only achieved if the supplier does not hold inventory between periods: there is no guarantee, even if the retailer orders the optimal period 2 quantity, that the retailer orders all of the supplier's inventory. However, it is quite plausible the supplier might attempt to use cheaper period 1 production to profit from a possible period 2 order.

Assuming the supplier fills the retailer's second-period order (which we earlier confirmed the supplier will do), the supplier's period 2 profit is

$$\begin{aligned}\Pi_2(x, q_1, q_2, \xi) &= bS(q_2 \mid \xi) - bq_2 - (q_2 - x)^+ c_2 \\ &= (1-\lambda)\Omega_2(q_2 \mid q_1, \xi) - w_2 q_2 + xc_2 - (x - q_2)^+ c_2.\end{aligned}$$

Given that $q_2 \geq q_1$, the above is strictly increasing in x for $x \leq q_1$. Hence, the supplier surely produces and delivers the retailer's period 1 order (as long as $q_1 \leq q_1^o$). The supplier's period 1 expected profit is

$$\begin{aligned}\Pi_1(x \mid q_1) &= -c_1 x + E[\Pi_2(x, q_1, q_2, \xi)] \\ &= -c_1 x + E[(1-\lambda)\Omega_2(q_2 \mid q_1, \xi)] - w_2 q_2 + xc_2 \\ &\quad - c_2 \int_0^{\xi(x)} (x - q_2(\xi))g(\xi)\,d\xi.\end{aligned}$$

It follows that

$$\frac{\partial\Pi_1(x \mid q_1)}{\partial x} = -c_1 + c_2(1 - G(\xi(x)))$$

and from Eq. (6.4)

$$\frac{\partial\Pi_1(q_1^o \mid q_1^o)}{\partial x} = -c_1 + c_2(1 - G(\xi(q_1^o))) < 0.$$

Hence, with a coordinating $\{w_1, w_2, b\}$ contract the supplier produces just enough inventory to cover the retailer's period 1 order. Overall, those contracts coordinate the supply chain and arbitrarily allocate profits.

Interestingly, with a coordinating contract the supplier's margin in period 2 is actually lower than in period 1:

$$w_2 - c_2 = w_1 - (\lambda c_1 + (1 - \lambda)c_2) < w_1 - c_1.$$

Intuition suggests the supplier should charge a higher margin for the later production since it offers the retailer an additional benefit over early production. Nevertheless, that intuition is incompatible with supply chain coordination (at least with a buyback contract).

6.2 Discussion

Forecast improvements present several challenges for supply chain coordination. Just as in the simpler single-period model, the retailer must be given incentives to order the correct amount of inventory given the forecast update. In addition, the supplier must correctly balance inexpensive early production against more expensive later production. Finally, the decentralized supply chain must be careful about inventory placement, since unlike with centralized operations, inventory is not necessarily moved to the optimal location in the supply chain, i.e., inventory can become 'stranded' at the supplier.

As in the single-location model, a buyback contract does coordinate this supply chain and arbitrarily allocates profit, even with voluntary compliance. Somewhat surprising, the supplier's margin with later production is smaller than her margin with early production, even though later production provides the retailer with a valuable service.

There are a number of useful extensions to this work. Consider a model with the following adjustments. Suppose the supplier at the start of period 1 picks her capacity, K, which costs c_k per unit. The supplier can produce at most K units over the two periods. Let c be the cost to convert one unit of capacity into one product. In this setting the supply chain optimal solution never produces in period 1: given that early production is no cheaper than later production, the supply chain should delay production until after it has the best demand forecast. A slight modification of the model lets the supplier produce K units in each period. In that case there is some incentive to conduct early production because then the total amount of inventory available in period 2 increases.

In a qualitatively similar model Brown and Lee (1998) study pay-to-delay contracts. With that contract the retailer reserves m units of the supplier's capacity in period 1 for a constant fee per unit. That commits the retailer to purchase at least m units in period 2. They show both firms can be better off with the pay-to-delay contract than with a contract that does not include minimum purchase agreements. However, the pay-to-delay contract cannot coordinate this supply chain. The reason is simple, minimum purchase agreements may result in more production than is optimal given the information signal: if a bad demand signal is observed it may be optimal for the supply chain to produce less than the minimum purchase agreement.

Information acquisition occurs exogenously in both the Brown and Lee (1998) and Donohue (2000), which is reasonable as long as the information is learned before the selling season starts. But suppose the firms' had a replenishment opportunity in the middle of the season. In that case early sales provides information on future sales. However, demand equals sales only if the retailer does not run out of inventory at the start of the season. See Barnes-Schuster, Bassok and Anupindi (2002) for a model in which the demand signal may be truncated due to lost sales. However, that is not the only additional complication in their model. The optimal solution has inventory held at the supplier between periods, hence coordination requires that inventory not be stranded at the supplier at the start of the second period. See Lu, Song and Regan (2002) for another model with midseason replenishment opportunity.

In Kouvelis and Gutierrez (1997) demand occurs in each period, with period 1 demand being the primary market demand and period 2 demand being the secondary market demand. Leftover inventory from period 1 can either be salvaged in the primary market or moved to the secondary market. That decision depends on the realization of the exchange rate between the two markets' currencies. Hence, the information learned between periods is not a demand signal, as in Donohue (2000), but rather the realization of period 2 effective production cost. They coordinate this supply chain (with one manager responsible for each market's decisions) using a nonlinear scheme. Kouvelis and Lariviere (2000) show an internal market can also coordinate this supply chain (see Section 10).

van Mieghem (1999) studies forecast updating with several additional twists. In his model the downstream firm is the manufacturer and the upstream firm is the subcontractor. At issue is the production of a component that is part of the manufacturer's product. The manufacturer has only one market for his product, but the subcontractor can sell her component either to the manufacturer or to an outside market. (But the subcontractor has access to the manufacturer's market only via the manufacturer, i.e., the subcontractor cannot sell directly to that market.) Both markets have random demand. Both firms choose a capacity level in period 1, where the manufacturer's capacity produces the component. At the start of period 2 the firms observe demand in their respective markets and then convert capacity into final output. Hence, like Donohue (2000), the firms receive a demand signal between their early decision (how much capacity to construct) and their later decision (how much to produce), albeit in van Mieghem (1999) it is a perfect demand signal. Unlike Donohue (2000), in van Mieghem (1999) the downstream firm has production capability, the upstream firm has a random opportunity cost for capacity not sold to the downstream firm, and buyback contracts are not considered. See Milner and Pinker (2000) for another model with early capacity decisions and later forecast adjustments.

van Mieghem (1999) and Donohue (2000) also differ on what they assume about the firms ability to commit to future actions. In Donohue (2000) the supplier commits to a period 2 wholesale price, whereas in van Mieghem (1999) the firms renegotiate their agreement between periods. Anand, Anupindi and Bassok (2001) demonstrate that a supplier's inability to commit to future prices may cause a retailer to carry inventory purely for strategic reasons. They have a two-period model with deterministic demand. The supplier's period 2 wholesale price is decreasing in the retailer's period 1 inventory, thereby providing the retailer with a motivation to carry inventory. See Gilbert and Cvsa (2000) for another model in which the ability to commit to future wholesale prices matters.

Future research should consider a model with endogenous information acquisition, i.e., the firms must exert effort to improve their demand forecasts. Should one firm exert the effort or should both firms undertake forecast improvement activities? To the best of my knowledge, that issue has not been explored. There is also the possibility the firms could have different forecasts: the firms could exert a different amount of effort toward forecasting or the firms could have different sources of forecasting information. If there are asymmetric forecasts, supply chain performance may improve via forecast sharing: see Section 10 for a discussion of that issue.

7 Coordination in the single-location base-stock model

This section considers a model with perpetual demand and many replenishment opportunities. Hence, the newsvendor model is not appropriate. Instead, the base-stock inventory policy is optimal: with a base-stock policy a firm maintains its inventory position (on-order plus in-transit plus on-hand inventory minus backorders) at a constant base-stock level. It is assumed, for tractability, that demand is backordered, i.e., there is no lost sale. As a result, expected demand is constant (i.e., it does not depend on the retailer's base-stock level). Optimal performance is now achieved by minimizing total supply chain costs: the holding cost of inventory and the backorder penalty costs. In this model the supplier incurs no holding costs, but the supplier does care about the availability of her product at the retail level. To model that preference, it is assumed that a backorder at the retailer incurs a cost on the supplier. Since the retailer does not consider that cost when choosing a base-stock level, it is shown the retailer chooses a base stock level that is lower than optimal for the supply chain, which means the retailer carries too little inventory. Coordination is achieved and costs are arbitrarily allocated by providing incentives to the retailer to carry more inventory.

This model also provides a useful building block for the two-location model considered in the next section.

7.1 Model and analysis

Suppose a supplier sells a single product to a single retailer. Let L_r be the lead time to replenish an order from the retailer. The supplier has infinite capacity, so the supplier keeps no inventory and the retailer's replenishment lead time is always L_r, no matter the retailer's order quantity. (There are two firms, but only the retailer keeps inventory, which is why this is considered a single-location model.) Let $\mu_r = E[D_r]$. Let F_r and f_r be the distribution and density functions of D_r, respectively: assume that F_r is strictly increasing, differentiable and $F_r(0) = 0$, which rules out the possibility that it is optimal to carry no inventory.

The retailer incurs inventory holding costs at rate $h_r > 0$ per unit of inventory. For analytical tractability, demand is backordered if stock is not available. [There are a few papers that consider multiple demand periods with lost sales: e.g., Moses and Seshadri (2000), Duenyas and Tsai (2001) and Anupindi and Bassok (1999).] The retailer incurs backorder penalty costs at rate $\beta_r > 0$ per unit backordered. The supplier has unlimited capacity, so the supplier need not carry inventory. However, suppose the supplier incurs backorder penalty costs at rate $\beta_s > 0$ per unit backordered at the *retailer*. In other words, the supplier incurs a cost whenever a customer wants to purchase the supplier's product from the retailer but the retailer does not have inventory. This cost reflects the supplier's preference for maintaining sufficient availability of her product at the retail level in the supply chain. Let $\beta = \beta_r + \beta_s$, so β is the total backorder cost rate incurred by the supply chain. Cachon and Zipkin (1999) adopt the same preference structure for the retailer, but in their model the supplier has limited inventory. Their model is discussed in the next section. Narayanan and Raman (1997) adopt a different preference structure: they assume a fixed fraction of consumers who experience a stockout for the supplier's product choose to purchase another product from a different supplier at the same retailer. Hence, the backorder cost for the supplier is her lost margin, whereas the backorder cost for the retailer is the difference in the margin between selling the supplier's product and selling the other product (which may actually benefit the retailer).

Sales occur at a constant rate μ_r, due to the backorder assumption, no matter how the firms manage their inventory. As a result, the firms are only concerned with their costs. Both firms are risk neutral. The retailer's objective is to minimize his average inventory holding and backorder cost per unit time. The supplier's objective is to minimize her average backorder cost per unit time.

Define the retailer's inventory level to equal inventory in-transit to the retailer plus the retailer's on-hand inventory minus the retailer's backorders. (This has also been called the effective inventory position.) The retailer's inventory position equals his inventory level plus on-order inventory (inventory ordered, but not yet shipped). Since the supplier immediately ships all orders, the retailer's inventory level and position are identical in this setting.

Let $I_r(y)$ be the retailer's expected inventory at time $t+L_r$ when the retailer's inventory level is y at time t:

$$I_r(y) = \int_0^y (y-x) f_r(x)\,dx = \int_0^y F_r(x)\,dx. \tag{7.1}$$

Let $B_r(y)$ be the analogous function that provides the retailer's expected backorders:

$$B_r(y) = \int_y^\infty (x-y) f_r(x)\,dx = \mu_r - y + I_r(y). \tag{7.2}$$

Inventory is monitored continuously, so the retailer can maintain a constant inventory position. In this environment it can be shown that a base-stock policy is optimal. With that policy the retailer continuously orders inventory so that his inventory position always equals his chosen base-stock level, s_r.

Let $c_r(s_r)$ be the retailer's average cost per unit time when the retailer implements the base-stock policy s_r:

$$\begin{aligned} c_r(s_r) &= h_r I_r(s_r) + \beta_r B_r(s_r) \\ &= \beta_r(\mu_r - s_r) + (h_r + \beta_r) I_r(s_r). \end{aligned}$$

Given the retailer's base-stock policy, the supplier's expected cost function is

$$\begin{aligned} c_s(s_r) &= \beta_s B_r(s_r) \\ &= \beta_s(\mu_r - s_r + I_r(s_r)). \end{aligned}$$

Let $c(s_r)$ be the supply chain's expected cost per unit time,

$$\begin{aligned} c(s_r) &= c_r(s_r) + c_s(s_r) \\ &= \beta(\mu_r - s_r) + (h_r + \beta) I_r(s_r). \end{aligned} \tag{7.3}$$

$c(s_r)$ is strictly convex, so there is a unique supply chain optimal base-stock level, s_r^o. It satisfies the following critical ratio equation

$$I_r'(s_r^o) = F_r(s_r^o) = \frac{\beta}{h_r + \beta}. \tag{7.4}$$

Let s_r^* be the retailer's optimal base-stock level. The retailer's cost function is also strictly convex, so s_r^* satisfies

$$F_r(s_r^*) = \frac{\beta_r}{h_r + \beta_r}.$$

Given $\beta_r < \beta$, it follows from the above two expressions that $s_r^* < s_r^o$, i.e., the retailer chooses a base-stock level that is less than optimal. Hence, channel coordination requires the supplier to provide the retailer with an incentive to raise his base-stock level.

Suppose the firms agree to a contract that transfers from the supplier to the retailer at every time t

$$t_I I_r(y) + t_B B_r(y)$$

where y is the retailer's inventory level at time t and t_I and t_B are constants. Furthermore, consider the following set of contracts parameterized by $\lambda \in (0, 1]$,

$$t_I = (1 - \lambda)h_r$$
$$t_B = \beta_r - \lambda\beta.$$

(Here we choose to rule out $\lambda = 0$ since then any base-stock level is optimal.) Given one of those contracts, the retailer's expected cost function is now

$$c_r(s_r) = (\beta_r - t_B)(\mu_r - s_r) + (h_r + \beta_r - t_I - t_B)I_r(s_r). \tag{7.5}$$

The contract parameters have been chosen so that

$$\beta_r - t_B = \lambda\beta > 0$$

and

$$h_r + \beta_r - t_I - t_B = \lambda(h_r + \beta) > 0.$$

It follows from Eqs. (7.3) and (7.5) that with these contracts

$$c_r(s_r) = \lambda c(s_r). \tag{7.6}$$

Hence, s_r^o minimizes the retailer's cost, i.e., those contracts coordinate the supply chain. In addition, those contracts arbitrarily allocate costs between the firms, with the retailer's share of the cost increasing in the parameter λ. Note, the λ parameter is not explicitly incorporated into the contract, i.e., it is merely used for expositional clarity.

Now consider the sign of the t_I and the t_B parameters. Since the contract must induce the retailer to choose a higher base-stock level, it is natural to conjecture $t_I > 0$, i.e., the supplier subsidizes the retailer's inventory holding cost. In fact, that conjecture is valid when $\lambda \in (0, 1]$. It is also natural to suppose $t_B < 0$, i.e., the supplier penalizes the retailer for backorders. But $\lambda \in (0, 1]$ implies $t_B \in [-\beta_s, \beta_r)$, i.e., with some contracts the supplier subsidizes the retailer's backorders ($t_B > 0$): in those situations the supplier encourages backorders by setting $t_B > 0$ because without that encouragement the large inventory subsidy leads the retailer to $s_r^* > s_r^o$.

The above analysis is reminiscence of the analysis with the newsvendor model and buyback contracts. This is not a coincidence, because this model is qualitatively identical to the newsvendor model. To explain, begin with the retailer's profit function in the newsvendor model (assuming $c_r = g_r = g_s = v = 0$):

$$\begin{aligned}\pi_r(q) &= pS(q) - wq \\ &= (p - w)q - pI(q).\end{aligned}$$

The retailer's profit has two terms, one that increases linearly in q and the other that depends on the demand distribution. Now let $p = h_r + \beta_r$ and $w = h_r$. In that case,

$$\pi_r(q) = \beta_r q - (h_r + \beta_r)I_r(q) = -c_r(q) + \beta_r \mu_r.$$

Hence, there is no difference between the maximization of $\pi_r(q)$ and the minimization of $c_r(s_r)$.

Now recall that the transfer payment with a buyback contract is $wq - bI(q)$, i.e., there is a parameter (i.e., w) that affects the payment linearly in the retailer's action (i.e., q), and a parameter that influences the transfer payment through a function (i.e., $I(q)$) that depends on the retailer's action and the demand distribution. In this model

$$t_I I_r(y) + t_B B_r(y) = (t_I + t_B)I_r(y) + t_B(\mu_r - y)$$

so t_B is the linear parameter and $t_I + t_B$ is the other parameter. In the buyback contract the parameters work independently. To get the same effect in the base-stock model the supplier could adopt a transfer payment that depends on the retailer's inventory position, s_r, and the retailer's inventory. That contract would yield the same results.

7.2 Discussion

Coordination with the infinite horizon base-stock model is qualitatively the same as coordination in the single-period newsvendor model. In particular, coordination via a holding cost and backorder cost transfer payment is like coordination via a buyback contract. One suspects that quantity-flexibility or sales-rebate like contracts could also coordinate in this setting. Choi, Dai and Song (2002) consider a similar model with the addition of a capacity constraint. They demonstrate that standard service level measures do not provide sufficient control over the firm managing inventory.

8 Coordination in the two-location base-stock model

The two-location base-stock model builds upon the single-location base-stock model discussed in the previous section. Now the supplier no longer has infinite capacity. Instead, she must order replenishments from her source and those replenishment always are filled within L_s time (i.e., her source has infinite capacity). So in this model the supplier enjoys reliable replenishments but the retailer's replenishment lead time depends on how the supplier manages her inventory. Only if the supplier has enough inventory to fill an order does the retailer receive that order in L_r time. Otherwise, the retailer must wait longer than L_r to receive the unfilled portion. That delay could lead to additional backorders at the retailer, which are costly to both the retailer and the supplier, or it could lead to lower inventory at the retailer, which helps the retailer.

In the single-location model the only critical issue is the amount of inventory in the supply chain. In this model the allocation of the supply chain's inventory between the supplier and the retailer is important as well. For a fixed amount of supply chain inventory the supplier always prefers that is more allocated to the retailer, because that lowers both her inventory and backorder costs. (Recall that the supplier is charged for retail backorders.) On the other hand, the retailer's preference is not so clear: less retail inventory means lower holding costs, but also higher backorder costs. There are also subtle interactions with respect to the total amount of inventory in the supply chain. The retailer is biased to carry too little inventory: the retailer bears the full cost of his inventory but only receives a portion of the benefit (i.e., he does not benefit from the reduction in the supplier's backorder cost). On the other hand, there is no clear bias for the supplier because of two effects. First, the supplier bears the cost of his inventory and does not benefit from the reduction in the retailer's backorder cost, which biases the supplier to carry too little inventory. Second, the supplier does not bear the cost of the retailer's inventory (which increases along with the supplier's inventory), which biases the supplier to carry too much inventory. Either bias can dominate, depending on the parameters of the model.

Even though it is not clear whether the decentralized supply chain will carry too much or too little inventory (however, it generally carries too little inventory), it is shown that the optimal policy is never a Nash equilibrium of the decentralized game, i.e., decentralized operation is never optimal. However, the competition penalty (the percent loss in supply chain performance due to decentralized decision making) varies considerably: in some cases the competition penalty is relatively small, e.g., less than 5%, whereas in other cases it is considerable, e.g., more than 40%. Therefore, the need for coordinating contracts is not universal.

8.1 Model

Let h_s, $0 < h_s < h_r$, be the supplier's per unit holding cost rate incurred with on-hand inventory. (When $h_s \geq h_r$ the optimal policy does not carry inventory at the supplier and when $h_s \leq 0$ the optimal policy has unlimited supplier inventory. Neither case is interesting.) The firms' operating decisions have no impact on the amount of in-transit inventory, so no holding cost is charged for either the supplier's or the retailer's pipeline inventory. Let $D_s > 0$ be demand during an interval of time with length L_s. (As in the single-location model, $D_s > 0$ ensures that the supplier carries some inventory in the optimal policy.) Let F_s and f_s be the distribution and density functions of that demand. As with the retailer, assume F_s is increasing and differentiable. Let $\mu_s = E[D_s]$. Retail orders are backordered at the supplier but there is no explicit charge for those backorders. The supplier still incurs per unit backorder costs at rate β_s for backorders at the retailer. The comparable cost for the retailer is still β_r. Even though there are no direct consequences to a supplier backorder, there are indirect consequences: lower retailer inventory and higher retailer backorders.

Both firms use base-stock policies to manage inventory. With a base-stock policy firm $i \in \{r, s\}$ orders inventory so that its inventory position remains equal to its base-stock level, s_i. (Recall that a firm's inventory level equals on-hand inventory, minus backorders plus in-transit inventory and a firm's inventory position equals the inventory level plus on-order inventory.) These base-stock policies operate only with local information, so neither firm needs to know the other firm's inventory position.

The firms choose their base-stock levels once and simultaneously. The firms attempt to minimize their average cost per unit time. (Given that one firm uses a base-stock policy, it is optimal for the other firm to use a base-stock policy.) They are both risk neutral. There exists a pair of base-stock levels, $\{s_r^o, s_s^o\}$, that minimize the supply chain's cost. [In fact, that policy is optimal among all possible policies. See Chen and Zheng (1994) for an elegant proof.] Hence, it is feasible for the firms to optimize the supply chain, but incentive conflicts may prevent them from doing so.

This model is essentially the same as the one considered by Cachon and Zipkin (1999).[33] However, the notation differs somewhat. (Caution, in some cases the notation is inconsistent.)

The first step in the analysis of this model is to evaluate each firm's average cost. The next step evaluates the Nash equilibrium base-stock levels. The third step identifies the optimal base-stock levels and compares them to the Nash equilibrium ones. The final step explores incentive structures to coordinate the supply chain. The remaining portions of this section describe alternative coordination techniques, summarize the results and discuss research in related models.

8.2 Cost functions

As in the single-location model, $c_r(y)$, $c_s(y)$ and $c(y)$ are the firms' and the supply chain's expected costs incurred at time $t+L_r$ at the retail level when the retailer's inventory level is y at time t. However, in the two-location model the retailer's inventory level does not always equal the retailer's inventory position, s_r, because the supplier may stockout. Let $c_i(s_r, s_s)$ be the average rate at which firm i incurs costs at the retail level and $c(s_r, s_s) = c_r(s_r, s_s) + c_s(s_r, s_s)$. To evaluate c_i, note that at any given time t the supplier's inventory position is s_s (because the supplier uses a base-stock policy). At time $t+L_s$ either the supplier's on-hand inventory is $(s_s - D_s)^+$ or the supplier's backorder equals $(D_s - s_s)^+$. Therefore, the retailer's inventory level at time $t+L_s$ is $s_r - (D_s - s_s)^+$. So

$$c_i(s_r, s_s) = F_s(s_s)c_i(s_r) + \int_{s_s}^{\infty} c_i(s_r + s_s - x)f_s(x)\,dx :$$

at time $t+L_s$ the supplier can raise the retailer's inventory level to s_r with probability $F_s(s_s)$, otherwise the retailer's inventory level equals $s_r + s_s - D_s$.

[33] Cachon and Zipkin (1999) assume periodic review, whereas this model assumes continuous review. That difference is inconsequential. In addition to the local inventory measure, they allow firms to use echelon inventory to measure their inventory position: a firm's echelon inventory position equals all inventory at the firm or lower in the supply chain plus on-order inventory minus backorders at the retail level. For the retailer there is no difference between the local and echelon measures of inventory position, but those measures are different for the supplier. They allow for either $\beta_s = 0$ or $\beta_r = 0$, but those special cases are not treated here. They include holding costs for pipeline inventory into their cost functions. Finally, they also study the Stackelberg version of this game (the firms choose sequentially instead of simultaneously).

Based on the analogous reasoning, let $I_r(s_r, s_s)$ and $B_r(s_r, s_s)$ be the retailer's average inventory and backorders given the base-stock levels:

$$I_r(s_r, s_s) = F_s(s_s)I_r(s_r) + \int_{s_s}^{\infty} I_r(s_r + s_s - x)f_s(x)\,dx,$$

$$B_r(s_r, s_s) = F_s(s_s)B_r(s_r) + \int_{s_s}^{\infty} B_r(s_r + s_s - x)f_s(x)\,dx.$$

Let $\pi_i(s_r, s_s)$ be firm i's total average cost rate. Since the retailer only incurs costs at the retail level,

$$\pi_r(s_r, s_s) = c_r(s_r, s_s).$$

Let $I_s(s_s)$ be the supplier's average inventory. Analogous to the retailer's functions (defined in the previous section)

$$I_s(y) = \int_0^y F_s(x)\,dx.$$

The supplier's average cost is

$$\pi_s(s_r, s_s) = h_s I_s(s_s) + c_s(s_r, s_s).$$

Let $\Pi(s_r, s_s)$ be the supply chain's total cost, $\Pi(s_r, s_s) = \pi_r(s_r, s_s) + \pi_s(s_r, s_s)$.

8.3 Behavior in the decentralized game

Let $s_i(s_j)$ be an optimal base-stock level for firm i given the base-stock level chosen by firm j, i.e., $s_i(s_j)$ is firm i's best response to firm j's strategy. Differentiation of each firm's cost function demonstrates that each firm's cost is strictly convex in its base-stock level, so each firm has a unique best response.

With a Nash equilibrium pair of base stocks, $\{s_r^*, s_s^*\}$, neither firm has a profitable unilateral deviation, i.e.,

$$s_r^* = s_r(s_s^*) \quad \text{and} \quad s_s^* = s_s(s_r^*).$$

Existence of a Nash equilibrium is not assured, but in this game existence of a Nash equilibrium follows from the convexity of the firm's cost functions (see Friedman, 1986). (Technically it is also required that the firms' strategy spaces have an upper bound. Imposing that bound has no impact on the analysis.) In fact, there exists a unique Nash equilibrium. To demonstrate uniqueness begin by bounding each player's feasible strategy space, i.e., the set of strategies a player may choose. For the retailer it is not difficult to show that $s_r(s_s) > \hat{s}_r > 0$, where $\hat{s}_r$ minimizes $c_r(y)$, i.e.,

$$F_r(\hat{s}_r) = \frac{\beta}{h_r + \beta}.$$

In other words, if the retailer were to receive perfectly reliable replenishments the retailer would choose $\hat{s}_r$, so the retailer certainly does not choose $s_r \le \hat{s}_r$ if replenishments are unreliable. (In other words, $\hat{s}_r$ is optimal for the retailer in the single-location model discussed in the previous section.) For the supplier, $s_s(s_r) > 0$ because $\partial\pi_s(s_r, s_s)/\partial s_s < 0$ given $F_s(s_s) = 0$ and $c_s'(y) < 0$.

Uniqueness of the Nash equilibrium holds if for the feasible strategies, $s_r > \hat{s}_r$ and $s_s > 0$, the best reply functions are contraction mappings (see Friedman, 1986), i.e.,

$$|s_i'(s_j)| < 1. \tag{8.1}$$

From the implicit function theorem

$$s_r'(s_s) = -\frac{\int_{s_s}^{\infty} c_r''(s_r + s_s - x) f_s(x)\,\mathrm{d}x}{F_s(s_s) c_r''(s_r) + \int_{s_s}^{\infty} c_r''(s_r + s_s - x) f_s(x)\,\mathrm{d}x}$$

and

$$s_s'(s_r) = -\frac{\int_{s_s}^{\infty} c_s''(s_r + s_s - x) f_s(x)\,\mathrm{d}x}{[h_s - c_s'(s_r)] f_s(s_s) + \int_{s_s}^{\infty} c_s''(s_r + s_s - x) f_s(x)\,\mathrm{d}x}.$$

Given $s_r > \hat{s}_r$ and $s_s > 0$, it follows that $c_r''(x) > 0$, $c_s'(y) < 0$, $c_s''(y) > 0$ and $F_s(s_s) > 0$. Hence, Eq. (8.1) holds for both the supplier and the retailer.

A unique Nash equilibrium is quite convenient, since that equilibrium is then a reasonable prediction for the outcome of the decentralized game. (With multiple equilibria it is not clear the outcome of the game would even be an equilibrium, since the players may choose strategies from different equilibria.) Hence, the competition penalty is an appropriate measure of the gap between optimal performance and decentralized performance, where the competition penalty is defined to be

$$\frac{\Pi(s_s^*, s_r^*) - \Pi(s_s^o, s_r^o)}{\Pi(s_s^o, s_r^o)}.$$

In fact, there always exists a positive competition penalty, i.e., decentralized operations always leads to suboptimal performance in this game.[34] To explain, note that the retailer's marginal cost is always greater than the supply chain's

$$\frac{\partial c_r(s_r, s_s)}{\partial s_r} > \frac{\partial c(s_r, s_s)}{\partial s_r}$$

because $c_r'(s_r) > c'(s_r)$. Since both $c_r(s_r, s_s)$ and $c(s_r, s_s)$ are strictly convex, it follows that, for any s_s, the retailer's optimal base stock is always lower than the supply chain's optimal base stock. Hence, even if the supplier chooses s_s^o, the retailer does not choose s_r^o, i.e., $s_r(s_s^o) < s_r^o$.

Although the Nash equilibrium is not optimal, Cachon and Zipkin (1999) find in a numerical study that the magnitude of the competition penalty depends on the parameters of the model. When the firms' backorder penalties are the same (i.e., $\beta_r/\beta_s = 1$) the median competition penalty for their sample is 5% and the competition penalty is no greater than 8% in 95% of their observations. However, very large competition penalties are observed when either $\beta_r/\beta_s < 1/9$ or $\beta_r/\beta_s > 9$. The retailer does not have a strong concern for customer service when $\beta_r/\beta_s < 1/9$, and so the retailer tends to carry far less inventory than optimal. Since the supplier does not have direct access to customers, the supplier can do little to prevent backorders in that situation, and so the supply chain cost is substantially higher than need be. In the other extreme, $\beta_r/\beta_s > 9$, the supplier cares little about customer service, and thus does not carry enough inventory. In that situation the retailer can still prevent backorders, but to do so requires a substantial amount of inventory at the retailer to account for the supplier's long lead time. The supply chain's cost is substantially higher than optimal if the optimal policy has the supplier carry inventory to provide reliable replenishments to the retailer. However, there are situations in which the optimal policy does not require the supplier to carry much inventory: either the supplier's holding cost is nearly as high as the retailer's (in which case keeping inventory at the supplier gives little holding cost advantage) or if the supplier's lead time is short (in which case the delay due to a lack of inventory at the supplier is negligible). In those cases the competition penalty is relatively minor.

In the single-location model decentralization always leads to less inventory than optimal for the supply chain. In this setting the interactions between the firms are more complex, and so decentralization generally leads to too little inventory, but not always. Since the retailer's backorder cost rate is lower than the supplier's backorder cost rate, for a fixed s_s the retailer always carries too little inventory, which certainly contributes to a less than optimal amount of inventory in the system. However, the retailer is only a part

[34] However, when $\beta_s = 0$ the optimal policy is a Nash equilibrium with just the right parameters. See Cachon and Zipkin (1999) for details.

of the supply chain. In fact, from Cachon and Zipkin's (1999) numerical study, the supplier's inventory may be so large that even though the retailer carries too little inventory, the total amount of inventory in the decentralized supply chain may exceed the supply chain's optimal quantity. Suppose β_r is quite small and β_s is quite large. In that case the retailer carries very little inventory. To attempt to mitigate the build up of backorders at the retail level the supplier provides the retailer with very reliable replenishments, which requires a large amount of inventory, an amount that may lead to more inventory in the supply chain than optimal.

The main conclusion from the analysis of the decentralized game is that the competition penalty is always positive, but only in some circumstances is it very large. It is precisely in those circumstances that the firms could benefit from an incentive scheme to coordinate their actions.

8.4 Coordination with linear transfer payments

Supply chain coordination in this setting is achieved when $\{s_r^o, s_s^o\}$ is a Nash equilibrium. Cachon and Zipkin (1999) propose a set of contracts to achieve that goal, but they do not answer two important questions: do the contracts allow for an arbitrary division of the supply chain's cost and is the optimal solution a unique Nash equilibrium? This section studies their contracts and answers both of those questions.

In the single-location model the supplier coordinates the supply chain with a contract that has linear transfer payments based on the retailer's inventory and backorders. Suppose the supplier offers the same arrangement in this model with the addition of a transfer payment based on the supplier's backorders:

$$t_I I_r(s_r, s_s) + t_B^r B_r(s_r, s_s) + t_B^s B_s(s_s),$$

where t_I, t_B^r and t_B^s are constants and $B_s(s_s)$ is the supplier's average backorder:

$$B_s(y) = \mu_s - y + I_s(y).$$

Recall that a positive value for the above expression represents a payment from the supplier to the retailer and a negative value represents a payment from the retailer to the supplier. While both firms can easily observe $B_s(s_s)$, an information system is needed for the supplier to verify the retailer's inventory and backorder.

The first step in the analysis provides some results for the optimal solution. The second step defines a set of contracts and confirms those contracts coordinate the supply chain. Then the allocation of costs is considered. Finally, it is shown that the optimal solution is a unique Nash equilibrium.

The traditional approach to obtain the optimal solution involves reallocating costs so that all costs are preserved. Base stock policies are then optimal and easily evaluated.[35] However, to facilitate the comparison of the optimal policy to the Nash equilibrium of the decentralized game, it is useful to evaluate the optimal base-stock policy without that traditional cost reallocation.[36]

Given $\Pi(s_r, s_s)$ is continuous, any optimal policy with $s_s > 0$ must set the following two marginals to zero

$$\frac{\partial \Pi(s_r, s_s)}{\partial s_r} = F_s(s_s)c'(s_r) + \int_{s_s}^{\infty} c'(s_r + s_s - x)f_s(x)\,dx \tag{8.2}$$

and

$$\frac{\partial \Pi(s_r, s_s)}{\partial s_s} = F_s(s_s)h_s + \int_{s_s}^{\infty} c'(s_r + s_s - x)f_s(x)\,dx. \tag{8.3}$$

Since $F_s(s_s) > 0$, there is only one possible optimal policy with $s_s > 0$, $\{\tilde{s}_r^1, \tilde{s}_s^1\}$, where $\tilde{s}_r^1$ satisfies

$$c'(\tilde{s}_r^1) = h_s, \tag{8.4}$$

and $\tilde{s}_s^1$ satisfies $\partial\Pi(\tilde{s}_r^1, \tilde{s}_s^1)/\partial s_s = 0$. Eq. (8.4) simplifies to

$$F_r(\tilde{s}_r^1) = \frac{h_s + \beta}{h_r + \beta},$$

so it is apparent $\tilde{s}_r^1$ exists and is unique. Since $\Pi(\tilde{s}_r^1, s_s)$ is strictly convex in s_s, $\tilde{s}_s^1$ exists and is unique.

There may also exist an optimal policy with $s_s \le 0$. In that case the candidate policies are $\{\tilde{s}_r^2, \tilde{s}_s^2\}$, where $\tilde{s}_s^2 \le 0$, $\tilde{s}_r^2 + \tilde{s}_s^2 = \bar{s}$ and $\bar{s}$ satisfies

$$\int_0^{\infty} c'(\bar{s} - x)f_s(x)\,dx = 0. \tag{8.5}$$

The above simplifies to

[35] Clark and Scarf (1960) pioneered that approach for the finite horizon problem. Federgruen and Zipkin (1984) make the extension to the infinite horizon case.

[36] This procedure does not prove base-stock policies are optimal, it merely finds the optimal base-stock policies. Furthermore, it relies on the continuity of $\Pi(s_r, s_s)$, so it does not trivially extend to discrete demand.

$$\Pr(D_r + D_s \leq \bar{s}) = \frac{\beta}{h_r + \beta},$$

so $\bar{s}$ exists and is unique.

Given that $\Pi(\tilde{s}_r^1, s_s)$ is strictly convex in s_s, $\Pi(\tilde{s}_r^1, \tilde{s}_s^1) < \Pi(\tilde{s}_r^2, \tilde{s}_s^2)$ whenever $\partial\Pi(\tilde{s}_r^1, 0)/\partial s_s < 0$. Since $\partial\Pi(s_r, 0)/\partial s_s$ is increasing in s_r, from Eq. (8.5), that condition holds when $\tilde{s}_r^1 < \bar{s}$, otherwise it does not. Therefore, $\{\tilde{s}_r^1, \tilde{s}_s^1\}$ is the unique optimal policy when $\tilde{s}_r^1 < \bar{s}$, otherwise any $\{\tilde{s}_r^2, \tilde{s}_s^2\}$ is optimal:

$$\{s_r^o, s_s^o\} = \begin{cases} \{\tilde{s}_r^1, \tilde{s}_s^1\} & \tilde{s}_r^1 < \tilde{s}_r^2 \\ \{\tilde{s}_r^2, \tilde{s}_s^2\} & \tilde{s}_r^1 \geq \tilde{s}_r^2 \end{cases}.$$

Now consider the firms' behavior with the following set of contracts parameterized by $\lambda \in (0, 1]$,

$$t_I = (1 - \lambda)h_r, \tag{8.6}$$

$$t_B^r = \beta_r - \lambda\beta, \tag{8.7}$$

$$t_B^s = \lambda h_s \left(\frac{F_s(s_s^o)}{1 - F_s(s_s^o)} \right). \tag{8.8}$$

Cachon and Zipkin (1999) also propose those contracts, but they do not include the λ parameter.

The retailer's cost function, adjusted for the above contracts is

$$\begin{aligned} c_r(y) &= (h_r - t_I)I_r(y) + (\beta_r - t_B^r)B_r(y) - t_B^s B_s(s_s) \\ &= \lambda c(y) - t_B^s B_s(s_s) \end{aligned}$$

and so

$$\pi_r(s_r, s_s) = \lambda\Pi(s_r, s_s) - t_B^s B_s(s_s). \tag{8.9}$$

Recall $c(s_r, s_s) = c_r(s_r, s_s) + c_s(s_r, s_s)$, so

$$\begin{aligned} \pi_s(s_r, s_s) &= h_s I_s(s_s) + (1 - \lambda)c(s_r, s_s) + t_B^s B_s(s_s) \\ &= (h_s + t_B^s)I_s(s_s) + (1 - \lambda)c(s_r, s_s) + t_B^s(\mu_s - s_s). \end{aligned} \tag{8.10}$$

There are two cases to consider: either $s_s^o > 0$ or $s_s^o = 0$. Take the first case. If $s_s^o > 0$, then Eq. (8.4) implies

$$\frac{\partial \pi_r(s_r^o, s_s^o)}{\partial s_r} = \left(\frac{\lambda}{1-\lambda}\right)\frac{\partial \pi_s(s_r^o, s_s^o)}{\partial s_s} = \lambda \frac{\partial \Pi(s_r^o, s_s^o)}{\partial s_s} = 0.$$

Further, $\pi_r(s_r, s_s)$ is strictly convex in s_r and $\pi_s(s_r^o, s_s)$ is strictly convex in s_s, so $\{s_r^o, s_s^o\}$ is indeed a Nash equilibrium. In fact, it is the unique Nash equilibrium. From the implicit function theorem, $s_r(s_s)$ is decreasing, where, as in the decentralized game without the contract, $s_i(s_j)$ is firm i's best response to firm j's strategy,c

$$\frac{\partial s_r(s_s)}{\partial s_s} = -\frac{\int_{s_s}^{\infty} c''(s_r + s_s - x) f_s(x)\,dx}{f_s(s_s)c''(s_r) + \int_{s_s}^{\infty} c''(s_r + s_s - x) f_s(x)\,dx} \leq 0.$$

Hence, for $s_r = s_r(s_s)$ and $\lambda \leq 1$ the supplier's marginal cost is increasing:

$$\frac{\partial \pi_s(s_r(s_s), s_s)}{\partial s_s} = F_s(s_s)\big(h_s - (1-\lambda)c'(s_r(s_s)) + t_B^s\big) - t_B^s.$$

Thus, there is a unique s_s that satisfies $s_s(s_r(s_s)) = s_s$, i.e., there is a unique Nash equilibrium.

Now suppose $s_s^o \leq 0$. It is straightforward to confirm all of the $\{\tilde{s}_r^2, \tilde{s}_s^2\}$ pairs satisfy the firms' first and second-order conditions. Hence, they are all Nash equilibria. Even though there is not a unique Nash equilibrium, the firms' costs are identical across the equilibria.

These contracts do allow the firms to arbitrarily allocate the retail level costs in the system, but they do not allow the firms to arbitrarily allocate all of the supply chain's costs. This limitation is due to the $\lambda \leq 1$ restriction, i.e., it is not possible with these contracts to allocate to the retailer more than the optimal retail level costs: while the retailer's cost function is well behaved even if $\lambda > 1$, the supplier's is not; with $\lambda > 1$ the supplier has a strong incentive to increase the retail level costs. Of course, fixed payments could be used to achieve those allocations if necessary. But since it is unlikely a retailer would agree to such a burden, this limitation is not too restrictive.

An interesting feature of these contracts is that the t_I and t_B^r transfer payments are identical to the ones used in the single-location model. This is remarkable because the retailer's critical ratio differs across the models: in the single-location model the retailer picks s_r such that

$$F_r(s_r) = \frac{\beta}{\beta + h_r}$$

whereas in the two-location model the retailer picks s_r such that

$$F_r(s_r) = \frac{\beta + h_s}{\beta + h_r}.$$

8.5 Other coordination methods

Alternatives to Cachon and Zipkin's linear scheme have been studied to coordinate similar supply chains under the special case that $\beta_s = 0$, i.e., the supplier incurs no cost associated with retail backorders. That preference structure is most appropriate for an internal supply chain in which each location is operated by a separate manager. For example, instead of a supplier, suppose the second stage in the supply chain is controlled by a warehouse manager. That manager may have no direct interest in the availability of the firm's product at the retail level.

Lee and Whang (1999) base their coordination scheme on the work by Clark and Scarf (1960). (They consider a periodic review model and their firms minimize discounted costs rather than average costs.) Clark and Scarf (1960), which focuses only on system wide performance, demonstrates that base-stock policies are optimal and they can be evaluated from a series of simple single-location optimization problems after the costs in the system are reallocated among the locations. Lee and Whang (1999) take the Clark and Scarf cost reallocation and show it can be used to coordinate decentralized operations. In their arrangement the supplier subsidizes the retailer's holding cost at the rate of h_s and charges the retailer an additional backorder penalty cost per unit at rate h_s. Given those transfers let $g_r(y)$ be the retailer's expected cost at time $t + L_r$ when the retailer's inventory level is y at time t,

$$\begin{aligned} g_r(y) &= (h_r - h_s)I_r(y) + (\beta + h_s)B_r(y) \\ &= (h_r + \beta)I_r(y) + (\beta + h_s)(\mu_r - y) \end{aligned}$$

where $\beta = \beta_r$ since $\beta_s = 0$. $g_r(y)$ is strictly convex and minimized by s_r^o. However, due to shortages at the supplier, the retailer's inventory level, IL_r, may be less than his inventory position, IP_r. To penalize the supplier for those shortages, the supplier transfers to the retailer at time t, $g^p(\mathrm{IL}_r, \mathrm{IP}_r)$,

$$g^p(x, y) = g_r(x) - g_r(y).$$

That transfer may be negative, i.e., the retailer pays the supplier, if the retailer (for irrational reasons) orders $\mathrm{IP}_r > s_r^o$ and the supplier does not fill that request completely, $\mathrm{IL}_r < \mathrm{IP}_r$. In addition, that transfer is not linear in the retailer's inventory position.

The retailer's final cost function is $\pi_r(s_r) = g_r(s_r)$, i.e., from the retailer's perspective the supplier provides perfectly reliable deliveries, since the supplier exactly compensates the retailer for any expected cost consequence of delivering less than the retailer's order. Hence, s_r^o is the retailer's optimal base-stock level.

The supplier's cost function with this arrangement is

$$\begin{aligned} c_s(s_r, s_s) &= h_s[I_s(s_s) + I_r(s_r, s_s) - B_r(s_r, s_s)] + \int_{s_s}^{\infty} g^p(s_r + s_s - x, s_r) f_s(x)\,dx \\ &= h_s[s_s + s_r - \mu_r - \mu_s] - (1 - F_s(s_s) g_r(s_r) + \int_{s_s}^{\infty} g_r(s_r + s_s - x) f_s(x)\,dx. \end{aligned}$$

Differentiate:

$$\begin{aligned} \frac{\partial c_s(s_r, s_s)}{\partial s_s} &= h_s + \int_{s_s}^{\infty} g_r'(s_r + s_s - x) f_s(x)\,dx \\ &= -\beta + (h_s + \beta)\left[F_s(s_s) + \int_{s_s}^{\infty} F_r(s_r + s_s - x) f_s(x)\,dx\right], \\ \frac{\partial^2 c_s(s_r, s_s)}{\partial s_s^2} &= (h_s + \beta)\left[f_s(s_s)(1 - F_r(s_r)) + \int_{s_s}^{\infty} f_r(s_r + s_s - x) f_s(x)\,dx\right] > 0. \end{aligned}$$

So the supplier's cost function is strictly convex and if the retailer's base-stock level is s_r^o, then the supplier's optimal base-stock level is s_s^o, i.e., $\{s_r^o, s_s^o\}$ is a Nash equilibrium. It is also not difficult to show that $\{s_r^o, s_s^o\}$ is the unique Nash equilibrium (assuming $s_s^o > 0$ is optimal).[37] However, this arrangement provides for only one division of the supply chain's profit. Fixed payments could be used to reallocate costs differently.

There are several differences between this transfer payment contract and Cachon and Zipkin's contract. Lee and Whang charge the supplier a nonlinear cost for supplier backorders, whereas Cachon and Zipkin charge a linear one. In the linear contract the retailer must account for supplier shortages, i.e., he does not receive direct compensation for those shortages, whereas with the nonlinear contract the retailer need not be concerned with the supplier's inventory management decision. (In other words, with the linear contract $s_r(s_s)$ is not independent of s_s, whereas it is with the nonlinear contract.) Furthermore, they both have linear transfers associated with I_r and B_r, but those transfers are different: with the nonlinear contract $t_I = h_s$ and $t_B^r = -h_s$, but that pair is not a member of the linear contracts ($t_I = h_s$ requires $\lambda = 1 - h_s/h_r$ whereas $t_B^r = -h_s$ requires $\lambda = (\beta_r + h_s)/\beta$).

[37] The retailer's best reply function is independent of s_s. From the implicit function theorem, the supplier's reaction function is decreasing in s_r; hence there is a unique Nash equilibrium.

Chen (1999a) uses an accounting inventory approach to coordinate a serial supply chain.[38] One firm is chosen to compensate all other firms for their actual costs.[39] In this case, assume that firm is the supplier, i.e., the supplier's transfer payment rate to the retailer equals h_r per unit of inventory and β_r per backorder. That leaves the retailer with no actual cost, and so no incentive to choose s_r^o. To provide that incentive, the supplier charges the retailer t_I^a per unit of accounting inventory and t_B^a per accounting backorder, where the retailer's accounting inventory equals the inventory the retailer would have if the supplier always delivered the retailer's full order. Hence, the retailer's expected payment to the supplier per unit time is

$$t_I^a I_r(s_r) + t_B^a B_r(s_r) = (t_I^a + t_B^a) I_r(s_r) + t_B^a(\mu_r - s_r).$$

The retailer's optimal base-stock level, s_r^*, minimizes that payment,

$$F_r(s_r^*) = \frac{t_B^a}{t_I^a + t_B^a}.$$

Now set

$$t_I^a = \lambda(h_r - h_s)$$
$$t_B^a = \lambda(h_s + \beta)$$

for $\lambda > 0$. It follows that

$$\frac{t_B^a}{t_I^a + t_B^a} = \frac{h_s + \beta}{h_r + \beta},$$

and so from Eq. (8.4) the retailer chooses s_r^o. With those contracts the retailer's expected payment is

$$\lambda(h_r + \beta)[I_r(s_r) + F_r(s_r^o)(\mu_r - s_r)].$$

[38] There are several differences between this model and his. In his model there are order-processing delays between stages whereas this model assumes orders are received immediately. In addition, he accommodates more than two levels in the supply chain. Finally, he studies a periodic review model and allows for discrete demand.

[39] Chen also considers the possibility that there exists an owner of the entire supply chain who does not make any operating decisions. Those decisions are made by the managers. In that situation the owner can adopt all of the supply chain's costs.

Hence, the retailer's action can be coordinated and any cost can be assigned to the retailer. The supplier has no control over the transfer payment received (once the terms are set), so the supplier minimizes the costs under her control, which equal the supply chain's costs. Thus, $\{s_r^o, s_s^o\}$ is a Nash equilibrium.

Relative to the two previously discussed approaches for coordination, Chen's accounting inventory is most closely related to Lee and Whang's approach. In fact, in this setting they are essentially equivalent. In both approaches the retailer's cost function is based on the presumption that the retailer's orders are always filled immediately. Second, in Lee and Whang the retailer's effective holding cost per unit is $h_r - h_s$, and in Chen the retailer's holding cost is $\lambda(h_r - h_s)$. Similarly, the backorder penalty costs are $\beta + h_s$ and $\lambda(\beta + h_s)$. When $\lambda = 1$, the approaches are the same. By allowing for $\lambda \neq 1$, the accounting inventory approach allows for any division of profit, whereas the Lee and Whang approach does not have that flexibility (merely because it lacks that parameter). Finally, in both cases the supplier bears all remaining costs in the supply chain, and so the supplier's cost function is equivalent with either scheme.

The equivalence between Chen's accounting inventory and Lee and Whang's nonlinear contract is surprising because Chen does not appear to charge the supplier for her backorders whereas Lee and Whang do charge a nonlinear penalty function. However, accounting inventory does charge a nonlinear penalty because the supplier compensates the retailer's actual costs. Thus, accounting inventory and the Lee and Whang approach are two different ways to describe the same transfer payments. This is probably a general result, but additional research is needed to confirm that conjecture after all differences between the models are reconciled.

Porteus (2000) offers responsibility tokens to coordinate the supply chain, which is also closely related to the Lee and Whang approach.[40] As with Lee and Whang, in Porteus the transfer rate from the supplier to the retailer is $h_s I_r - h_s B_r$. However, Porteus does not include an explicit charge associated with B_s. Instead, in lieu of actual inventory, the supplier issues a responsibility token for each unit the retailer orders that the supplier is unable to fill. From the retailer's perspective that token is equivalent to inventory: the retailer incurs a holding cost of $h_r - h_s$ per token per unit time and incurs no backorder penalty cost if the token is used to 'fill' demand. If a token is used to fill demand then the supplier is charged the full backorder cost until the supplier provides a unit of actual inventory to fill that demand. Similarly, the supplier receives the retailer's holding cost on each token since the supply chain does not actually incur a holding cost on these imaginary tokens. Hence, with this system the retailer receives a perfectly reliable supply from the supplier and the supplier bears the consequence of her backorders, just as with Lee and Whang. However, in Lee and Whang the supplier pays

[40] He considers a periodic review, finite horizon model with multiple firms in a serial supply chain.

the expected cost consequence of a backorder whereas with responsibility tokens the supplier pays the actual cost consequence. The same holds with accounting inventory. When all players are risk neutral this distinction does not matter.

Watson (2002) considers coordination of a serial supply chain with AR(1) demand. Given this demand process the optimal order at each stage of the supply chain is not equal to demand, as it is with independent demand. Hence, schemes that use local penalties, e.g., Chen's (1999a) accounting inventory or the linear transfer payments of Cachon and Zipkin (1999), require each stage to forecast the ordering process of the subsequent stage, which is nontrivial with this demand process, especially for the highest stages. Watson proposes an alternative which is computationally friendlier. In his scheme the stage is given incentives to manage echelon inventory, and so each stage needs only to observe the demand process.

8.6 Discussion

In the two-location base-stock model decentralized operations always leads to suboptimal performance, but the extent of the performance deterioration (i.e., the competition penalty) depends on the supply chain parameters. If the firms' backorder costs are similar then the competition penalty is often reasonably small. The competition penalty can be small even if the supplier does not care about customer service because her operations role in the supply chain may not be too important (e.g., if h_s is large or if L_s is small).

To coordinate the retailer's action the firms can agree to a pair of linear transfer payments that function like the buyback contract in the single-period newsvendor model. To coordinate the supplier's action Cachon and Zipkin (1999) propose a linear transfer payment based on the supplier's backorders. With these contracts the optimal policy is the unique Nash equilibrium. Further, they allow the firms to arbitrarily divide the retail level costs. Lee and Whang (1999), Chen (1999a) and Porteus (2000) propose alternative coordination schemes for the special case in which the supplier does not care about retail level backorders.

There are a number of worthwhile extensions to the two-location base-stock model. Caldentey and Wein (2000) study a model that is the two-location model considered here, with the exception that the supplier chooses a production rate rather than an inventory policy. They demonstrate that many of the qualitative insights from Cachon and Zipkin (1999) continue to hold. Cachon (1999) obtains the same finding for a two-echelon serial supply chain with lost sales. Duenyas and Tsai (2001) also study a two-echelon serial supply chain with lost sales but they include several twists. First, they consider dynamic policies rather than just base-stock policies. Second, upstream inventory is needed by the downstream firm to satisfy its demand, but that inventory can also be used to satisfy demand in some outside market.

Given a choice between the two demands, the supply chain prefers to use the inventory to meet the needs of the downstream firm. But since demand is stochastic, it may be optimal to satisfy some outside demand. They demonstrate that while fixed wholesale-price contracts do not coordinate the supply chain, they are nevertheless quite effective, i.e., the competition penalty is small for most parameter settings. Finally, Parker and Kapuscinsky (2001) tackle the considerably harder problem of coordinating a serial supply chain with capacity constraints.

A natural next step after serial supply chains is to consider supply chains with multiple retailers. However, the analytical complexity of even centralized operations in those models poses quite a challenge. Cachon (2001) obtains results for a two-echelon model with multiple retailers and discrete stochastic demand using the theory of supermodular games (see Topkis, 1998). He demonstrates that multiple Nash equilibria may exist, and the optimal policy may even be a Nash equilibrium. Hence, decentralized operations do not necessarily lead to suboptimal performance.[41] Andersson and Marklund (2000) and Axsater (2001) consider a similar model but approach coordination differently.

Chen, Federgruen and Zheng (2001) (CFZ) study a model with one supplier, multiple noncompeting retailers and deterministic demand. [Bernstein and Federgruen (1999) study a closely related model with competing retailers.] While the centralized solution is intractable, for the case with fixed prices, and therefore fixed demand rates, Roundy (1985) provides a class of policies that is nearly optimal. Within that class CFZ find the centralized solution. They show that a single-order quantity-discount policy cannot coordinate the action of heterogenous retailers and they propose a set of transfer payments that does coordinate the supply chain. The coordination issues in this model are quite different than in the previously discussed multiechelon models with stochastic demand. For example, in the two-location model the retailer's action has no impact on the supplier's holding cost, whereas in the CFZ model the retailer's action impacts the supplier's holding and order processing costs. Furthermore, in the CFZ model, unlike with stochastic demand, all customer demands are met without backordering and the supplier never delays a shipment.

9 Coordination with internal markets

In each of the previous models the firms agree to a contract that explicitly stipulates transfer payments, e.g., the buyback rate is b or the revenue share is ϕ. However, there are situations that call for more flexibility, i.e., the transfer payment rates are contingent on the realization of some

[41] The true optimal policy is unknown for that supply chain, so performance is measured relative to the optimal policy within the class of reorder point policies.

random event that occurs after the contract is signed. This section studies a model in which such contingency is provided via an internal market.

9.1 Model and analysis

Suppose there is one supplier, one production manager and two retailers. The production manager is the supplier's employee and the retailers are independent firms. The following sequence of events occurs: the production manager chooses a production input level e, which yields an output of $Q = Ye$ finished units, where $Y \in [0, 1]$ is a random variable; the production manager incurs cost $c(e)$, where $c(e)$ is strictly convex and increasing; retailer i observes α_i, the realization of the random variable $A_i > 0$; each retailer submits an order to the supplier; the supplier allocates q_i units to retailer i, being sure that q_i does not exceed retailer i's order and $q_1 + q_2 \leq Q$; and finally retailer i earns revenue $q_i p_i(q_i)$, where $p_i(q_i) = \alpha_i q_i^{-1/\eta}$ and $\eta > 1$ is the constant demand elasticity. Let θ be the realization of Y, $A = \{A_1, A_2\}$ and $\alpha = \{\alpha_1, \alpha_2\}$. This model is closely related, albeit with different notation, to one developed by Kouvelis and Lariviere (2000), which in turn is a variant of the model developed by Porteus and Whang (1991).[42] See Agrawal and Tsay (2001) and Erkoc and Wu (2002b) for related models.

Before considering transfer payments between the firms, consider the supply chain optimal actions. Given that each retailer's revenue is strictly increasing in his allocation, it is always optimal to allocate the entire output to the retailers. Thus, let γ be the fraction of Q that is allocated to retailer one. Let $\pi(\gamma, \alpha, Q)$ be total retailer revenue if retailer one is allocated γQ units and retailer two is allocated $(1-\gamma)Q$ units:

$$\pi(\gamma, \alpha, Q) = (\alpha_1 \gamma^{(\eta-1)/\eta} + \alpha_2 (1-\gamma)^{(\eta-1)/\eta}) Q^{(\eta-1)/\eta}.$$

The optimal allocation of production to the two retailers depends on the demand realizations but not on the production output. Further, revenue is concave in γ, so let $\gamma^{o}(\alpha)$ be the optimal share to allocate to retailer one:

$$\gamma^{o}(\alpha) = \frac{\alpha_1^{\eta}}{\alpha_1^{\eta} + \alpha_2^{\eta}}. \tag{9.1}$$

Conditional on an optimal allocation, the retailers' total revenue is

$$\pi(\alpha, Q) = \pi(\gamma^{o}(\alpha), \alpha, Q) = (\alpha_1^{\eta} + \alpha_2^{\eta})^{1/\eta} Q^{(\eta-1)/\eta}.$$

[42] Kouvelis and Lariviere (2000) do not label their players 'supplier' and 'retailers'. More importantly, they have two agents responsible for production: the output of production is $\theta q_1 q_2$, where q_i is the action taken by the ith agent. Porteus and Whang (1991) have a single production agent and N demand agents (which are the retailers in this model). Their production agent faces an additive output shock rather than a multiplicative one. Their demand agents face a newsvendor problem with effort-dependent demand.

Total expected supply chain profit, $\Pi(e, A, Y)$, is thus

$$\Pi(e, A, Y) = E[\pi(A, Ye)] - c(e).$$

Profit is concave in e, so the unique optimal production effort level, e^{o}, satisfies

$$\Pi'(e^{o}) = \left(\frac{\eta - 1}{\eta}\right)(e^{o})^{-1/\eta} E[(A_1^{\eta} + A_2^{\eta})^{1/\eta} Y^{(\eta-1)/\eta}] - c'(e^{o}) = 0. \tag{9.2}$$

Now consider decentralized operations. To achieve channel coordination it must be that the retailers purchase all of the supplier's output and that output must be allocated to the retailers properly. It is now apparent that a fixed per unit wholesale price cannot achieve those tasks: for any fixed wholesale price there is some realization of Y such that the retailers do not purchase all of the supplier's output. In addition, with a fixed wholesale price it is possible the retailers desire more than the supplier's output, in which case the supplier must implement some allocation rule. The possibility of rationing could cause the retailers to submit strategic orders, which in turn could lead to an inefficient allocation of output (see Cachon & Lariviere, 1999). To alleviate those problems the supplier could make the transfer price contingent on the realization of A and Q. To be specific, suppose the supplier charges the retailers w per unit. Assuming retailer i receives q_i units, retailer i's profit is

$$\pi_i(q_i, w) = \alpha_i q_i^{(\eta-1)/\eta} - w q_i.$$

Retailer i's optimal quantity, q_i^*, satisfies

$$\frac{\partial \pi_i(q_i^*, w)}{\partial q_i} = 0 = \left(\frac{\eta - 1}{\eta}\right)\alpha_i (q_i^*)^{-1/\eta} - w.$$

It follows that $q_i^* = \gamma^{o}(\alpha) Q$ when $w = w(\alpha, Q)$,

$$w(\alpha, Q) = \left(\frac{\eta - 1}{\eta}\right)(\alpha_1^{\eta} + \alpha_2^{\eta})^{1/\eta} Q^{-1/\eta}.$$

Hence, when the supplier charges $w(\alpha, Q)$ the retailers order exactly Q units in total and the allocation of inventory between them maximizes supply chain revenue. Note that $w(\alpha, Q)$ is precisely the marginal value of additional production,

$$\frac{\partial \pi(\alpha, Q)}{\partial Q} = w(\alpha, Q).$$

The supplier could offer the retailers a contract that identifies $w(\alpha, Q)$ as the wholesale price contingent on the realizations of A and Y, but that surely would be an unruly contract (in length and complexity). Furthermore, implementation of that contract requires the supplier to actually observe the realizations of A, which may not occur. Fortunately, there is a simpler alternative. The supplier merely commits at the start of the game to hold a market for output after the retailers observe their demand realizations. The unique market-clearing price is $w(\alpha, Q)$, and so the market optimizes the supply chain's profit without the supplier observing A.

It remains to determine the supplier's compensation scheme for her production manager. To complicate matters, assume the supplier is unable to observe the production manager's effort, i.e., e is noncontractible. But the supplier is able to observe the firm's final output, Ye. So suppose the supplier pays the production manager

$$\left(\frac{\eta-1}{\eta}\right)(e^{\mathrm{o}})^{-1/\eta}E[(A_1^\eta + A_2^\eta)^{1/\eta}Y^{(\eta-1)/\eta}]/E[Y] = E[Qw(A,Q)\,|\,e^{\mathrm{o}}]/E[Q\,|\,e^{\mathrm{o}}] \tag{9.3}$$

per unit of realized output. Hence, the supplier pays the production manager the expected shadow price of capacity and sells capacity to the retailers for the realized shadow price of capacity. Given that scheme, the production manager's expected utility is

$$u(e) = \left(\frac{\eta-1}{\eta}\right)(e^{\mathrm{o}})^{-1/\eta}E[(A_1^\eta + A_2^\eta)^{1/\eta}Y^{(\eta-1)/\eta}]e - c(e)$$

and the marginal utility is

$$u'(e) = \left(\frac{\eta-1}{\eta}\right)(e^{\mathrm{o}})^{-1/\eta}E[(A_1^\eta + A_2^\eta)^{1/\eta}Y^{(\eta-1)/\eta}] - c'(e).$$

A comparison with Eq. (9.2) reveals that the production manager's optimal effort is e^{o}.

The supplier earns zero profit in expectation from the internal market: $E[Qw(A,Q)\,|\,e^{\mathrm{o}}]$ is the expected revenue from the retailers, and, from Eq. (9.3), it is also the expected payout to the production manager. To earn a positive profit the supplier can charge the production manager and/or the retailers fixed fees. In fact, in more general settings, Kouvelis and Lariviere (2000) show that supplier breaks even or loses money with the internal market approach to supply chain coordination.[43] Hence, this is a viable strategy for the supplier only if it is coupled with fixed fees.

[43] For example, suppose there were two production managers and output equaled Ye_1e_2, where e_i is the effort level of production manager i. In that case Kouvelis and Lariviere (2000) show that the supplier/market maker loses money on the market. See Holmstrom (1982) for additional discussion on coordination and budget balancing.

9.2 Discussion

In this model one agent in the supply chain produces a resource (production output) that has uncertain value to the other part of the supply chain (i.e., the retailers). Coordination therefore requires the proper incentive to produce the resource as well as the proper incentive to consume the resource. Production can be coordinated with a single price (the expected value of output) but its consumption requires a state-dependent price, which can be provided via a market mechanism. Interestingly, the expected revenue from selling the resource via the market may be less than the expected cost to purchase the resource. Hence, there must be a market maker that stands between the producers and consumers of the resource and the market maker must be ready to lose money in this market. So the market maker is willing to participate only if there exists some other instrument to extract rents from the participants (e.g., fixed fees). In other words, the market maker uses the market to align incentives, but does not directly profit from the market.

The Donohue (2000) model is similar to this one in the sense that the value of first-period production is uncertain: it depends on the realization of the demand signal at the start of period 2. A market was not necessary in that model because it is optimal for the retailer to purchase the supplier's entire period 1 output. Now suppose there is a holding cost for inventory and the retailer's holding cost is higher than the supplier's. In that case it may be optimal for the supply chain to produce more in period 1 than the retailer orders, i.e., the excess production is held in storage at the supplier until needed in period 2. It remains optimal to move all period 1 production to the retailer in period 2: inventory at the supplier cannot satisfy demand. But only $w_2 = 0$ can ensure the retailer orders all of the supplier's inventory. Unfortunately, with that price any period 2 order is optimal for the retailer. Hence, the buyback contract with fixed wholesale prices in each period is no longer a practical coordination scheme for the supply chain. See Barnes-Schuster et al. (1998) for a more formal treatment of this argument.

10 Asymmetric information

In all of the models considered so far the firms are blessed with full information, i.e., all firms possess the same information when making their decisions. Hence, any coordination failure is due exclusively to incentive conflicts and not due to an inability by one firm to evaluate the optimal policy. However, in practice full information is rare. Given the complexity and geographic breadth of most modern supply chains it is not surprising that at least one firm lacks some important piece of information that another firm possesses. For instance, the manufacturer of a product may have a more accurate forecast of demand than the manufacturer's supplier of a critical component. In that case optimal supply chain performance requires more than the coordination of actions. It also requires the sharing of information so that each firm in the supply chain is able to determine the precise set of optimal actions.

Sometimes sharing information is not difficult. For example, suppose the relevant information is the demand distribution of a product with stationary stochastic demand. Hence, the demand forecast can be shared by sharing past sales data. (That might be technically challenging, but the credibility of the forecast is not in doubt.) In that case the interesting research questions include how to use that information to improve supply chain performance and by how much performance improves (see Cachon & Fisher, 2000; Chen, 1998; Gavirneni et al., 1999; Lee, So & Tang, 2000).

Unfortunately, there is also the possibility of opportunistic behavior with information sharing. For example, a manufacturer may tell her supplier that demand will be quite high to get the supplier to build a substantial amount of capacity. This is particularly problematic when the demand forecast is constructed from diverse and unverifiable pools of information. Consider the demand forecast for a new product. The manufacturer's sales manager may incorporate consumer panel data into her forecast, which could be shared with a supplier, but her forecast may also include her subjective opinion based on a myriad of information gathered from her years of experience in the industry. If the sales manager knows her market well, those guesses and hunches may be quite informative, yet there really is no obvious way for her to convey that information to the supplier other than to say in her opinion expected demand is 'x'. In other words, the sales manager has important information that the supplier cannot easily verify, so the supplier may not be sure 'x' is indeed the sales manager's best forecast. Furthermore, it is even difficult for the supplier to verify the forecast *ex post*: if demand turns out to be less than 'x' the supplier cannot be sure the sale manager gave a biased forecast since a low-demand realization is possible even if 'x' is the correct forecast.

This section considers a supply chain contracting model with asymmetric demand forecasts that is based on Cachon and Lariviere (2001).[44] As before, the main issues are which contracts, if any, achieve supply chain coordination and how are rents distributed with those contracts. In this model coordination requires (1) the supplier takes the correct action and (2) an accurate demand forecast is shared.

In addition to information sharing, this model highlights the issue of contract compliance, as was first discussed in Section 2. With forced compliance (i.e., all firms must choose the actions specified in the contract)

[44] See Riordan (1984) for a similar model: he has asymmetric information regarding demand, asymmetric information regarding the supplier's cost and the capacity and production decisions are joined, i.e., production always equals capacity.

supply chain coordination and accurate forecast sharing are possible. However, with voluntary compliance (i.e., each firm chooses optimal actions even if they deviate from those in the contract) information sharing is possible but only if optimal supply chain performance is sacrificed.

10.1 The capacity procurement game

In the capacity procurement game a manufacturer, M, develops a new product with uncertain demand. There is a single potential supplier, S, for a critical component, i.e., even M is not able to make this component. [See Milner and Pinker (2001) for a capacity contracting model without asymmetric information in which the downstream firm is able to provide some of its own capacity but still may depend on the upstream firm's capacity when demand is high.] Let D_θ be demand, where $\theta \in \{h, l\}$, Let $F(x \mid \theta)$ be the distribution function of demand, where $F(x \mid \theta) = 0$ for all $x < 0$, $F(x \mid \theta) > 0$ for all $x \geq 0$, and $F(x \mid \theta)$ is increasing and differentiable. Furthermore, D_h stochastically dominates D_l, i.e., $F(x \mid h) < F(x \mid l)$ for all $x \geq 0$.

With full information both firms observe the θ parameter. With asymmetric information the θ parameter is observed only by the manufacturer. In that case the supplier's prior beliefs are that $\Pr(\theta = h) = \rho$ and $\Pr(\theta = l) = 1 - \rho$. The manufacturer also knows ρ, i.e., the prior is common knowledge.

The interactions between M and S are divided into two stages. In stage 1, M gives S a demand forecast and offers S a contract which includes an initial order, q_i. Assuming S accepts the contract, S then constructs k units of capacity at a cost $c_k > 0$ per unit. In stage 2, M observes D_θ and places her final order with S, q_{f}, where the contract specifies the set of feasible final orders. Then S produces $\min\{D_\theta, k\}$ units at a cost of $c_p > 0$ per unit and delivers those units to M. Finally, M pays S based on the agreed contract and M earns $r > c_p + c_k$ per unit of demand satisfied. The salvage value of unused units of capacity is normalized to zero. The qualitative behavior of the model is unchanged if M only observes an imperfect signal of demand in stage 2.

Like the newsvendor model studied in Section 2, this model has only one demand period. But in the newsvendor model production occurs before the demand realization is observed, whereas in this model production, constrained by the initial capacity choice, occurs after the demand realization is observed. This model is also different than the two-stage newsvendor model considered in Section 6. In that model some production can be deferred until after demand information is learned, but that production is more expensive than early production. In this model the cost of production is the same no matter in which stage it occurs. Hence, unlike in the Section 6 model, it is never optimal to produce before the demand information is learned.

10.2 Full information

To establish a benchmark, in this section assume both firms observe θ and begin the analysis with the supply chain optimal solution. The supply chain makes two decisions: how much capacity to construct, k, and how much to produce. The latter is simple, produce $\min\{k, D_\theta\}$ after observing demand. Hence, the only substantive decision is how much capacity to build. Let $S_\theta(x)$ be expected sales with x units of capacity,

$$S_\theta(x) = x - E[(x - D_\theta)^+] \\ = x - \int_0^x F_\theta(x)\,\mathrm{d}x.$$

Let $\Omega_\theta(k)$ be the supply chain's expected profit with k units of capacity,

$$\Omega_\theta(k) = (r - c_p)S_\theta(k) - c_k k.$$

Given that $\Omega_\theta(k)$ is concave, the optimal capacity, k_θ^{o}, satisfies the newsvendor critical ratio:

$$\overline{F}_\theta(k_\theta^{\mathrm{o}}) = \frac{c_k}{r - c_p},$$

where $\overline{F}_\theta(x) = 1 - F_\theta(x)$.[45] Let $\Omega_\theta^{\mathrm{o}} = \Omega_\theta(k_\theta^{\mathrm{o}})$. Thus, supply chain coordination is achieved if the supplier builds k_θ^{o} units of capacity and defers all production until after receiving the manufacturer's final order.

Now turn to the game between M and S. There are many different types of contracts the manufacturer could offer the supplier. Consider an options contract: M purchases q_i options for w_{o} per option at stage 1 and then pays w_{e} to exercise each option at stage 2. Hence, the total expected transfer payment is

$$w_{\mathrm{o}} q_i + w_{\mathrm{e}} S_\theta(q_i).$$

That contract could also be described as a buyback contract: M pays $w = w_{\mathrm{o}} + w_{\mathrm{e}}$ at stage 1 for an order quantity of q_i and S pays $b = w_{\mathrm{e}}$ per unit in stage 2 that M 'returns', i.e., does not take actual delivery. Alternatively, that contract could be described as a wholesale-price contract ($w_{\mathrm{o}} + w_{\mathrm{e}}$ is the wholesale price) with a termination penalty charged for each unit M cancels in stage 2 (where w_{o} is the termination penalty). Erkoc and Wu (2002a) study

[45] Given that $F_\theta(0) > 0$, it is possible that $k_\theta^{\mathrm{o}} = 0$, but that case is not too interesting, so assume $k_\theta^{\mathrm{o}} > 0$. Boundary conditions are also ignored in the remainder of the analysis.

reservation contracts in a capacity procurement game with convex capacity costs: with a reservation contract M reserves a particular amount of capacity before observing demand and then pays a fee to S for each unit of reserved capacity that is not utilized once demand is observed. That contract is not considered in this section.

On the assumption the supplier builds enough capacity to cover the manufacturer's options, $k=q_i$, the manufacturer's expected profit is

$$\Pi_\theta(q_i) = (r - w_e)S_\theta(q_i) - w_o q_i.$$

If contract parameters are chosen so that $(r-w_e)=\lambda(r-c_p)$ and $w_o=\lambda c_k$, where $\lambda \in [0, 1]$, then

$$\Pi_\theta(q_i) = \lambda\Omega_\theta(q_i).$$

Hence, $q_i = k_\theta^o$ is the manufacturer's optimal order, the supply chain is coordinated and the supply chain's profit can be arbitrarily allocated between the firms. Indeed, the supplier's profit is $(1-\lambda)\Omega_\theta(q_i)$, so k_θ^o also maximizes the supplier's profit, apparently confirming the initial $k=q_i$ assumption.

But there is an important caveat to the above analysis. Can the manufacturer be sure the supplier indeed builds $k=q_i$? Suppose the manufacturer is unable to verify the supplier actually builds $k=q_i$. Given the supplier's capacity depends on a number of factors that are hard for the manufacturer to verify (labor practices, production schedules, component yields, etc.), it is not surprising if there were situations in which the manufacturer would be unable to prove the supplier built less than q_i capacity. With that in mind, consider the following profit function for a supplier (assuming $k<q_i$) who believes demand is τ,

$$\begin{aligned}\pi(k, q_i, \tau) &= (w_e - c_p)s_\tau(k) + w_o q_i - c_k k \\ &= (1-\lambda)(r - c_p)s_\tau(k) - c_k(k - \lambda q_i).\end{aligned}$$

It follows that

$$\frac{\partial\pi(k_\theta^o, k_\theta^o, \theta)}{\partial k} < 0,$$

i.e., k_θ^o does not maximize the supplier's profit if $q_i = k_\theta^o$. The source of the problem is that the above cost function assumes the manufacturer pays w_o per option no matter what capacity is constructed. Therefore, only the w_e parameter of the contract impacts the supplier's decision on the margin, i.e., the supplier sets his capacity as if the supplier is offered just a wholesale-price contract.

Cachon and Lariviere (2001) define forced compliance to be the case when the supplier must choose $k=q_i$ and voluntary compliance to be the case when the supplier can choose $k<q_i$ even though the manufacturer pays $w_o q_i$ for q_i options. Both situations represent extreme ends of a spectrum: with forced compliance the supplier acts as if any deviation from $k<q_i$ is infinitely costly, whereas with voluntary compliance the supplier acts as if there are no consequences. Reality is somewhere in the middle. (This discussion is analogous to the one with retail effort, which in reality is neither fully contractible nor fully uncontractible.) Nevertheless, voluntary compliance is worth study because the supplier is likely to build $k<q_i$ even if there is some penalty for doing so. For example, suppose the supplier must refund the manufacturer for any option the manufacturer purchased that the supplier is unable to exercise, i.e., if $D_\theta > k$, the supplier refunds the manufacturer $(\min\{q_i, D_\theta\} - k)^+ w_o$. Even with such a penalty, the supplier has a $1-F(k)$ chance of pocketing the fee for $(q_i - k)$ options without having built the capacity to cover those options: if $D_\theta < k$, then the manufacturer would not know the supplier was unable to cover all options because the supplier covers the options the manufacturer exercises. That incentive is enough to cause the supplier to choose $k<q_i$. Erkoc and Wu (2002a) propose an alternative approach in the context of their reservation contract: they study a game in which the supplier pays a penalty for each unit of capacity that is reserved but is not delivered. They find sufficient penalties such that compliance is achieved.

To summarize, with forced compliance the manufacturer can use a number of contracts to coordinate the supply chain and divide its profit. However, coordination with those contracts is not assured with anything less than forced compliance. As a result, voluntary compliance is a more conservative assumption (albeit, possibly too conservative).[46]

The remainder of this section studies the manufacturer's contract offer under voluntary compliance. As suggested above, with voluntary compliance the manufacturer's initial order has no impact on the supplier's capacity decision. It follows that transfer payments based on the initial order also have no impact on the supplier's capacity decision. To influence the supplier's capacity decision the manufacturer is relegated to a contract based on his final order, q_f. An obvious candidate is the wholesale-price contract.

With a wholesale-price contract the supplier's profit is

$$\pi_\theta(k) = (w - c_p)S_\theta(k) - c_k k.$$

[46] Intermediate compliance regimes are challenging to study because the penalty for noncompliance is not well behaved in k: in general $1-F(k)$ is neither concave nor convex in k. See Krasa and Villamil (2000) for a model in which the compliance regime is an endogenous variable. As already mentioned, see also Erkoc and Wu (2002a) for an approach to achieve compliance.

Note that the supplier evaluates his expected profit on the assumption that θ is certainly the demand parameter. The supplier's profit is strictly concave in k, so there exists a unique wholesale price for any k such that k is optimal for the supplier. Let $w_\theta(k)$ be that wholesale price:

$$w_\theta(k) = \frac{c_k}{\overline{F}_\theta(k)} + c_p.$$

The manufacturer's profit function can now be expressed as

$$\Pi_\theta(k) = (r - w_\theta(k))S_\theta(k),$$

i.e., the manufacturer can choose a desired capacity by offering the wholesale-price contract $w_\theta(k)$. Differentiate the manufacturer's profit function:

$$\begin{aligned}
\Pi'_\theta(k) &= (r - w_\theta(k))S'_\theta(k) - w'_\theta(k)S_\theta(k),\\
\Pi''_\theta(k) &= (r - w_\theta(k))S''_\theta(k) - w''_\theta(k)S_\theta(k) - 2w'_\theta(k)S'_\theta(k)\\
&= -\left(r - c_p + \frac{c_k}{\overline{F}_\theta(k)}\right)f_\theta(k) - w''_\theta(k)S_\theta(k).
\end{aligned}$$

If $w''_\theta(k) > 0$, then $\Pi_\theta(k)$ is strictly concave. So for convenience assume $w''_\theta(k) > 0$, which holds for any demand distribution with an increasing failure rate, i.e., the hazard rate, $f_\theta(x)/(1 - F_\theta(x))$, is increasing. The normal, exponential and the uniform meet that condition, as well as the gamma and Weibull with some parameter restrictions (see Barlow & Proschan, 1965). Therefore, there is a unique optimal capacity, k^*_θ, and a unique wholesale price that induces that capacity, $w^*_\theta = w(k^*_\theta)$. It follows from $\Pi'_\theta(k^*_\theta) = 0$ that the supply chain is not coordinated, $k^*_\theta < k^o_\theta$:

$$\begin{aligned}
\overline{F}_\theta(k^*_\theta) &= \frac{c_k}{r - c_p}\left(1 + \frac{f_\theta(k^*_\theta)}{\overline{F}_\theta(k^*_\theta)^2}\, S_\theta(k^*_\theta)\right)\\
&= \overline{F}_\theta(k^o_\theta)\left(1 + \frac{f_\theta(k^*_\theta)}{\overline{F}_\theta(k^*_\theta)^2}\, S_\theta(k^*_\theta)\right).
\end{aligned}$$

10.3 Forecast sharing

Now suppose the supplier does not observe θ. The supplier has a prior belief regarding θ, but to do better the supplier might ask the manufacturer for her forecast of demand, since the supplier knows the manufacturer knows θ. If the manufacturer announces demand is low, the supplier should believe the forecast: a high-demand manufacturer is unlikely to bias her forecast down. But a high-demand forecast is suspect, since there is the real possibility a low-demand manufacturer would offer a high-demand forecast to get the supplier to build more capacity: the manufacturer always wants more capacity built if the manufacturer is not paying for it. Thus, a sophisticated supplier ignores the manufacturer's verbal forecast, and instead infers the manufacturer's demand from the contract the manufacturer offers. With that understanding, the manufacturer shares her demand forecast by offering the right contract.

To continue with that reasoning, in this model forecast sharing takes place via a signaling equilibrium. With a signaling equilibrium the supplier assigns a belief to each possible contract the manufacturer could offer: either the contract is offered only by a high-demand manufacturer, offered only by a low-demand manufacturer or it could be offered by either type. There are different kinds of signaling equilibria. With a separating equilibrium the supplier divides the contracts into the former two sets: either the contract is a high-type contract or a low-type contract, but no contract is offered by both types. With a pooling equilibrium the contracts are divided into the latter two sets: either the contract is offered by a low-type or it is offered by both types. In a pooling equilibrium there is no means to share a high-demand forecast (since no contract is designated for only the high-type manufacturer) and the low-demand forecast is shared only out of equilibrium (since each type prefers the contract designated for both types, which conveys no information regarding the manufacturer's demand). Thus, forecast sharing only occurs with a separating equilibrium: a high-type M offers the best contract among those designated for high-types and a low-type M offers the best contract among those designated for low-types.

A separating equilibrium is rational if a high-type M indeed prefers to offer a high-type contract rather than to offer a low-type contract. The high-type recognizes that if she offers a low-type contract the supplier builds capacity based on the assumption demand is indeed low. Similarly, a separating equilibrium must also have that a low-type M prefers to offer a low-type contract rather than to offer a high-type contact. In other words, the low-type manufacturer must not prefer to mimic a high-type manufacturer by choosing a high-type contract. That condition is more onerous because if the low-type M offers a high-type contract then the supplier builds capacity under the assumption demand is high. See Cachon and Lariviere (2001) for a more formal description of these conditions.

First consider separating equilibria with forced compliance. Recall, the manufacturer can coordinate the supply chain with an options contract (and the parameters of those contracts do not depend on the demand distribution). The type θ manufacturer's profit with one of those contracts is $\lambda\Omega_\theta(q_i)$: with forced compliance the supplier must build $k = q_i$ if the supplier accepts the contract, so the supplier's belief regarding demand has no impact on the capacity choice given the supplier accepts the contract. The supplier's profit is

$(1-\lambda)\Omega_\theta(q_i)$. Let $\hat{\pi}$ be the supplier's minimum acceptable profit, i.e., the supplier rejects the contract if his expected profit is less than $\hat{\pi}$.[47] Clearly both types of manufacturers want a large share of the supply chain profit and for a fixed share, both manufacturers want to maximize the supply chain's profit. However, a manufacturer could prefer a larger share of suboptimal profit over a smaller share of optimal profit. Hence, the supplier must be diligent about biased forecasts.

Suppose the low-type manufacturer offers an options contract with

$$\lambda_{\rm l} = 1 - \frac{\hat{\pi}}{\Omega_{\rm l}^{\rm o}}$$

and the high-type manufacturer offers $\min\{\lambda_{\rm h}, \hat{\lambda}_{\rm h}\}$, where

$$\lambda_{\rm h} = 1 - \frac{\hat{\pi}}{\Omega_{\rm h}^{\rm o}}$$

and

$$\hat{\lambda}_{\rm h} = \frac{\Omega_{\rm l}^{\rm o} - \hat{\pi}}{\Omega_{\rm l}(k_{\rm h}^{\rm o})}.$$

Since $\min\{\lambda_{\rm h}, \hat{\lambda}_{\rm h}\} > \lambda_{\rm l}$, the high-type manufacturer has no interest in offering the low-type's contract: with that contract the manufacturer earns a smaller share ($\lambda_{\rm l}$ vs. $\min\{\lambda_{\rm h}, \hat{\lambda}_{\rm h}\}$) of a lower profit ($\Omega_{\rm h}(k_{\rm l}^{\rm o})$ vs. $\Omega_{\rm h}^{\rm o}$). The low-type manufacturer also has no interest in offering the high-type's contract. By construction, the low-type manufacturer is indifferent between earning her low-type profit, $\Omega_{\rm l}^{\rm o} - \hat{\pi}$, and earning $\hat{\lambda}_{\rm h}$ percent of the high-type contract profit, $\Omega_{\rm l}(k_{\rm h}^{\rm o})$. As long as the high-type captures no more than $\hat{\lambda}_{\rm h}$ percent of the supply chain's profit, the low-type has no interest in pretending to be a high-type. Thus, the above contracts are a separating equilibrium.

With low demand the supplier earns $\hat{\pi}$. With high-demand the supplier earns the same amount if $\lambda_{\rm h} < \hat{\lambda}_{\rm h}$, otherwise the supplier earns more. The manufacturer would prefer to take more from the supplier when $\lambda_{\rm h} > \hat{\lambda}_{\rm h}$, but then the supplier cannot trust the manufacturer's forecast (because a low-type would gladly accept some suboptimal supply chain performance in exchange for a large fraction of the profit). Even though the manufacturer may be unable to drive the supplier's profit down to $\hat{\pi}$, the supply chain is coordinated in all situations. Hence, with forced compliance

[47] It is assumed $\hat{\pi}$ is independent of θ. But it is certainly plausible that a supplier's outside opportunity is correlated with the manufacturer's demand: if the manufacturer has high demand, then other manufacturers may have high demand, leading to a higher than average opportunity cost to the supplier. Additional research is needed to explore this issue.

information is exchanged via the parameters of the contract and not via the form of the contract.

Now consider voluntary compliance. Here the manufacturer is relegated to offering a wholesale-price contract. In a separating equilibrium the low-type manufacturer offers w_{l}^{*} and the high-type manufacturer would like to offer w_{h}^{*}. If $w_{\mathrm{h}}^{*} > w_{\mathrm{l}}^{*}$ then it is possible the low-type manufacturer will not mimic the high-type: mimicking gets the low-type manufacturer more capacity, but she must pay a higher price. If $w_{\mathrm{h}}^{*} \leq w_{\mathrm{l}}^{*}$ then mimicking certainly occurs: the low-type manufacturer gets more capacity and pays no more per unit. In that case the high-type manufacturer needs to supplement the wholesale-price contract with some additional transfer payment that only a high-type manufacturer would be willing to pay.

One suggestion is for the high-type M to also offer the supplier a fixed fee, A. The low-type manufacturer earns $\Pi_{\mathrm{l}}(w_{\mathrm{l}}^{*}, l)$ when she reveals herself to be a low-type, where $\Pi_{\theta}(w, \tau)$ is a type θ manufacturer's profit when the supplier believes demand is type $\tau \in \{l, h\}$. She earns $\Pi_{\mathrm{l}}(w_{\mathrm{h}}^{*}, h) - A$ when she mimics the high-type (i.e., offers the high-type contract), so she does not wish to mimic when

$$\Pi_{\mathrm{l}}(w_{\mathrm{l}}^{*}, l) \geq \Pi_{\mathrm{l}}(w_{\mathrm{h}}^{*}, h) - A.$$

The high-type prefers to offer that fixed fee rather than to offer the low-type's contract when

$$\Pi_{\mathrm{h}}(w_{\mathrm{h}}^{*}, h) - A > \Pi_{\mathrm{h}}(w_{\mathrm{l}}^{*}, l).$$

There exists a fixed fee that satisfies both conditions when

$$\Pi_{\mathrm{h}}(w_{\mathrm{h}}^{*}, h) - \Pi_{\mathrm{h}}(w_{\mathrm{l}}^{*}, l) > \Pi_{\mathrm{l}}(w_{\mathrm{h}}^{*}, h) - \Pi_{\mathrm{l}}(w_{\mathrm{l}}^{*}, l).$$

The above states that the high-type manufacturer has more to gain from the supplier believing demand is high than the low-type manufacturer.

While the fixed fee works, there may be a cheaper approach for the high-type manufacturer to signal. An ideal signal is not costly to a high-type manufacturer but very costly to a low-type manufacturer. Clearly the fixed fee is not ideal because it is equally costly to each type. A better signal is for the high-type manufacturer to offer a higher wholesale price. For the high-type manufacturer, a higher wholesale price is not costly at all initially, while it is costly for the low-type:

$$\frac{\partial \Pi_{\mathrm{h}}(w_{\mathrm{h}}^{*}, h)}{\partial w} = 0,$$

whereas, if $w_h^* \geq w_l^*$, then

$$\frac{\partial \Pi_l(w_h^*, h)}{\partial w} < 0.$$

Another option for the high-type manufacturer is to offer a firm commitment: in stage 1 the manufacturer commits to purchase at least m units in stage 2, i.e., $q_f \geq m$, with any remaining units purchased for a wholesale price w. That contract could also be called a capacity reservation contract: at stage 1 the manufacturer reserves m units of the supplier's capacity that she promises to utilize fully at stage 2. The firm commitment is not costly to the manufacturer when $D_\theta \geq m$, but is costly when $D_\theta < m$. Since $D_l < m$ is more likely than $D_h < m$, a firm commitment is more costly to the low-demand manufacturer than the high-demand manufacturer. Hence, it too is a cheaper signal than the fixed payment. Interestingly, a firm commitment is not desirable with full information because it may lead to an ex-post-inefficient action: if $D_\theta < m$ then the production of $m - D_\theta$ units is wasted. However, with asymmetric information these contracts can enhance supply chain performance by allowing for the credible communication of essential information. See Cachon and Lariviere (2001) for a more detailed analysis of these different types of signaling instruments.

10.4 Discussion

This section considers a model in which one member of the supply chain has a better forecast of demand than the other. Since supply chain coordination requires that the amount of capacity the supplier builds depend on the demand forecast, supply chain coordination is achieved only if the demand forecast is shared accurately. With forced compliance the manufacturer can use an options contract to coordinate the supplier's action and to share information. However, sharing information is more costly with voluntary compliance. Nevertheless, some techniques for credibly sharing forecasts are cheaper than others. In particular, firm commitments, which are not optimal with full information, are effective for a manufacturer that needs to convince a supplier that her high-demand forecast is genuine.

Given that forecast sharing is costly even with the best signaling instrument, the high-demand manufacturer may wish to consider options other than signaling. One option is for the manufacturer to pay the supplier for units and take delivery of them in stage 1, i.e., before the demand realization is observed. In that case the supplier's profit does not depend on the manufacturer's demand distribution, so there is no need to share information. However, this option completely disregards the benefit of deferring production until after the demand realization is observed. A second option is for the manufacturer to choose a contract associated with a pooling equilibrium. In that case the supplier evaluates the contract as if he is dealing with a representative manufacturer. Hence, the terms are not as good as the high-demand manufacturer could get with full information, but this option could be attractive if there is a substantial cost to signal her high demand. (For example, a person in excellent health might opt for a standard health group life insurance plan merely to avoid the hassle of medical exams to demonstrate her excellent health.) Integration of all of these options awaits further research.

There are several other papers on supply chain contracting with asymmetric information. Cohen, Ho, Ren and Terwiesch (2001) study the forecasting process in the semiconductor equipment supply chain. In this setting the supplier has a long lead time to complete a piece of equipment and the manufacturer's desired completion time is uncertain. The manufacturer has an internal forecast for the desired delivery date and can provide a forecast of the delivery date to the supplier. However, since the manufacturer does not want the supplier to be late with a delivery, the manufacturer is biased to forecast a delivery date that is sooner than really needed. (The manufacturer is powerful in this supply chain and so can refuse to take delivery of completed equipment until the equipment is actually needed.) The supplier is well aware of this bias, but the research question is whether the supplier acts as if the forecast is biased. For example, if the manufacturer announces that the equipment is needed in the third quarter of a year does the supplier act as if the manufacturer really needs it in the first quarter of the following year. Using data from the industry they find that the supplier indeed acts as if the forecasts are biased. Terwiesch, Ren, Ho and Cohen (2002) extend this result to demonstrate empirically that suppliers given poor delivery lead times to manufacturers that are notorious for biased forecasting. Terwiesch and Loch (2002) also study signaling, but in the context of a product designer with an unknown ability to create valuable designs to their clients.

Porteus and Whang (1999) study a model that closely resembles the model in this section except they have the supplier offering the contract rather than the manufacturer. Hence, they study screening (the party without the information designs the contract to learn information) rather than signaling (the party with the information designs the contract to communicate information). As a result, the supplier offers a menu of contracts, one designed for each type of manufacturer.

Ha (1996) also studies a screening model. He has a supplier offering a contract to a manufacturer with stochastic demand. [Corbett and Tang (1999) study a similar model with deterministic demand.] The manufacturer knows his cost, but that cost is uncertain to the supplier. After the supplier offers the contract the manufacturer orders q units and sets the retail price. Supply chain coordination is possible with full information; however, the coordinating parameters depend on the manufacturer's cost. (His coordinating contract prohibits the manufacturer from setting a price below a specified threshold, i.e., a resale price maintenance provision.) If the supplier does not know the manufacturer's cost, Ha suggests the supplier offer a menu of contracts: for each order quantity the supplier stipulates a transfer payment and a minimum resale price. Each contract is targeted to a particular manufacturer, i.e., manufacturers with different costs choose different contracts and each manufacturer chooses the contract designed for his cost. That latter property is due to the revelation principle (Myerson, 1979).[48] Unfortunately, supply chain coordination is no longer possible. See Lim (2001) for a screening model that separates producers that vary in the quality of their output.

Corbett and de Groote (2000) study a model with one buyer, one supplier, deterministic demand and fixed ordering costs for each level in the supply chain. As in Porteus and Whang (1999) and Ha (1996), the contract designer has a prior belief regarding the other firm's cost and that firm knows his cost precisely. They propose a quantity-discount schedule, which is like a contract menu: there is a unique per unit price for each quantity the buyer may choose. As in Ha (1996) supply chain coordination is not achieved if there is asymmetric information. Corbett (2001) studies coordination with one supplier, one buyer, stochastic demand, fixed ordering costs and asymmetric information with respect to either the fixed ordering cost or the backorder penalty cost. He finds that consignment stock influences incentives, sometimes in a beneficial way, sometimes in a destructive way. (Consignment assigns ownership of inventory at the downstream firm to the upstream firm.)

Brown (1999) is related to the capacity procurement model. He has one supplier, one manufacturer and a single demand period with stochastic demand. But there are some important differences. First, he considers only forced compliance. Second, in his model the manufacturer announces her demand forecast to the supplier, and then immediately places her final order, i.e., there is no intermediate step in which the supplier builds capacity between the initial and final order. Brown requires that the manufacturer's order be consistent with the announced forecast, i.e., assuming the forecast is true the manufacturer's order is optimal. This constraint is reasonable because the supplier is able to immediately verify with certainty any inconsistency between the forecast and the order quantity. In the capacity procurement model the supplier is never able to verify for sure whether a biased forecast was provided, so that constraint would be problematic. Furthermore, it is necessary for the manufacturer to provide both an order quantity and a forecast, because there is a continuum of forecasts (where a forecast includes a mean and a standard

[48] Kreps (1990) describes the revelation principle as "... obvious once you understand it but somewhat cumbersome to explain." (p. 691). See his book for a good entry into mechanism design. In a nutshell, the revelation principle states that if there exists an optimal mechanism and that mechanism does not completely reveal the player's types, then the outcome of that mechanism can be replicated with a mechanism that does reveal the player's types. Thus, the search for optimal mechanisms can be restricted to truth-inducing mechanisms.

deviation) that yield the same optimal order quantity. In other words, he allows for a continuum of manufacturer types.

Brown studies two related contracts, an options contract (which he refers to as a buyback contract) and an options-futures contract. The latter is a buyback contract with a maximum threshold that the manufacturer can return: the manufacturer orders z units, pays c_f for the first y units, pays $w > c_f$ for the remaining $z-y$ units, and can return up to $z-y$ units for credit $b < w$. Hence, the options-futures contract contains firm commitments.

Both types of contracts coordinate the supply chain with full information and arbitrarily allocate profit. With asymmetric information Brown assumes the supplier accepts the contract only if it yields a minimum expect profit. However, with the buyback contract this fixed profit benchmark creates an incentive for the manufacturer to announce a biased forecast. To explain, let $\hat{\pi}$ be that benchmark, let $\{\mu(\sigma), \sigma\}$ be the mean–standard deviation pairs such that the manufacturer's order quantity is optimal and let $\Omega(\mu(\sigma), \sigma)$ be the supply chain's profit with that order quantity. The manufacturer must offer a contract in which the supplier's share of supply chain profit, $1-\lambda$, is at least $\hat{\pi}/\Omega^o(\mu(\sigma), \sigma)$. In the newsvendor setting $\Omega(\mu(\sigma), \sigma)$ is decreasing in σ, which implies $1-\lambda$ is increasing in σ, and the manufacturer's share, λ, is decreasing in σ. Hence, the manufacturer's optimal forecast has $\sigma = 0$ even if in reality $\sigma > 0$. If the supplier accepts that contract then his expected profit is in fact less than $\hat{\pi}$. Brown shows this incentive is eliminated if the manufacturer offers a futures-options contract: i.e., there exists a one for one relationship between the set of options-futures contracts and the set of manufacturer types such that a manufacturer always prefers the contract designated for his type. Hence, as in capacity procurement model, firm commitments are a useful tool for conveying information. They are particularly attractive in Brown's model because they do not result in suboptimal performance.

11 Conclusion

Over the last decade the legitimacy of supply chain contracting research has been established and many research veins have been tapped. Several key conclusions have emerged. First, coordination failure is common; incentive conflicts plausibly arise in a wide range of operational situations. As a result, suboptimal supply chain performance is not necessarily due to incompetent managers or naive operating policies. Rather, poor supply chain performance may be due to conflicting incentives and these incentive conflicts can be managed. Second, in many situations there are multiple kinds of contracts that achieve coordination and arbitrarily divide profit. Hence, the contract selection process in practice must depend on criteria or objectives that have not been fully explored, i.e., there is a need for additional research that investigates the subtle, but possibly quite important, differences among the set of coordinating contracts. Third, the consequence of coordination failure is context specific: there are situations in which supply chain performance is nearly optimal with naive and simple contracts, but there are also situations in which decentralized operations without proper incentive management leads to substantially deteriorated performance. It is quite useful to have theory that can help to contrast those cases. Fourth, this body of work emphasizes that managing incentive conflicts can lead to Pareto improvements, which is often referred to as a 'win-win' situation in practice. This insight should help to break the 'zero sum game' mentality which is understandably so prevalent among supply chain managers and is a strong impediment to significant supply chain progress.[49] While vigilance is always prudent, a wise supply chain manager recognizes that not every offer is a wolf in sheep's clothing.

Unfortunately, theory has almost exclusively followed practice in this domain, i.e., practice has been used as a motivation for theoretical work, but theoretical work has not found its way into practice. This need not be so. As already mentioned, one of the surprises of this research is that coordination can be achieved with many different contractual forms. An understanding of the subtle differences among these contracts may allow a researcher to identify a particularly suitable contract form for an industry, even if that contract form has no precedence in the industry. Just as there has been documented improvements with innovations like delayed differentiation (Lee, 1996) and accurate response (Fisher & Raman, 1996), it should be possible to generate equally valuable improvements via innovations in incentive design.

As a first step toward wider implementation, this research needs to develop an empirical–theoretical feedback loop. As this chapter illustrates, the literature contains a considerable amount of theory, but an embarrassingly paltry amount of empiricism. Thus, we have little guidance on how the theory should now proceed. For example, we have identified a number of contracts that coordinate a supplier selling to a newsvendor but can we explain why certain types have been adopted in certain industries and not others? Can we explain why these contracts have not completely eliminated the Pareto inferior wholesale-price contract? A standard argument is that the wholesale-price contract is cheaper to administer, but we lack any evidence regarding the magnitude of the administrative cost of the more complex contracts. And even if the coordinating contracts are adopted, such as buybacks or revenue sharing, are coordinating parameters chosen in practice? For example, the set of revenue-sharing contracts is much larger than the set of coordinating revenue-sharing contracts. If we observe that firms choose noncoordinating contracts, then we need an explanation. Irrational or incompetent behavior on the part of managers is a convenient explanation, but it is not satisfying to build a theory on irrational behavior. A theory is

[49] In a zero sum game one player's payoff is decreasing in the other player's payoff, so one player can be made better off only by making the other player worse off.

interesting only if it can be refuted and irrational behavior cannot be refuted. A better approach is to challenge the assumptions and analysis of the theory. With some empiricism we should be able to identify which parts of the theory are sound and which deserve more scrutiny.

The franchise literature could provide a useful guide to researchers in supply chain contracting. An excellent starting point is Lafontaine and Slade (2001). They review and compare the extensive theoretical and empirical results on franchising. Some of the predictions from theory are indeed supported by numerous empirical studies, while others are lacking. It is clear that the give-and-take between theory and data has been enormously successful for that body of work.

On a hopeful note, some preliminary activity in the empirical domain has fortunately begun. Mortimer (2000) provides an analysis of revenue-sharing contracts in the video cassette rental industry. Cohen, Ho, Ren and Terwiesch (2002) carefully evaluate forecast sharing in the semiconductor equipment industry. Their findings are consistent with the premise in Cachon and Lariviere (2001): if forecasts are not credible, then they will be ignored and supply chain performance suffers. Follow on work of theirs demonstrates that providing poor forecasts leads to lower future credibility and lower received service from a supplier. Novak and Eppinger (2001) empirically evaluate the interaction between product complexity and the make or buy decision, and find support for the property rights theory of vertical integration. Finally, Randall, Netessine and Rudi (2002) study whether e-retailers choose to drop ship or hold their own inventory. The appropriate strategy for a retailer depends on the characteristics of its product and industry, as predicted by the theoretical work in Netessine and Rudi (2000a), and they indeed find that e-retailers that chose the appropriate strategy were less likely to bankrupt.

Even though our most rewarding efforts now lie with collecting data, it is still worthwhile to comment on areas of the theory that need on additional investigation. Current models are too dependent on single-shot contracting. Most supply chain interactions occur over long periods of time with many opportunities to renegotiate or to interact with spot markets. For some steps in this direction, see Kranton and Minehart (2001) for work on buyer–supplier networks and long-run relationships; Plambeck and Taylor (2002) for a model with renegotiation of quantity-flexibility contracts; and Wu, Kleindorfer and Zhang (2002) and Lee and Whang (2002) for the impact of spot markets on capacity contracting and inventory procurement, respectively.

More research is needed on how multiple suppliers compete for the affection of multiple retailers, i.e., additional emphasis is needed on many-to-one or many-to-many supply chain structures. In the context of auction theory, Jin and Wu (2002) and Chen (2001) study procurement in the many-to-one structure and Bernstein and Véricourt (2002) offer some initial work on a many-to-many supply chain. Forecasting and other types of information sharing require much more attention. Lariviere (2002) provides recent work in this area. Finally, more work is needed on how scarce capacity is allocated in a supply chain and how scarce capacity influences behavior in the supply chain. Recent work in this area is provided by Zhao, Deshpande and Ryan (2002) and Deshpande and Schwarz (2002).

To summarize, opportunities abound.

Acknowledgements

I would like to thank the many people who carefully read and commented on the first two drafts of this manuscript: Ravi Anupindi, Fangruo Chen, Charles Corbett, James Dana, Ananth Iyer, Ton de Kok, Yigal Gerchak, Mark Ferguson, Paul Kleindorfer, Howard Kunreuther, Marty Lariviere, Serguei Netessine, Ediel Pinker, Nils Rudi, Leroy Schwarz, Sridhar Seshadri, Greg Shaffer, Yossi Sheffi, Terry Taylor, Christian Terwiesch, Andy Tsay and Kevin Weng. I am, of course, responsible for all remaining errors.

References

Agrawal, N., A.A. Tsay (2001). Intrafirm incentives and supply chain performance, in: J.S. Song, D.D. Yao (eds.), *Supply Chain Structures: Coordination, Information and Optimization (Volume 42 of the International Series in Operations Research and Management Science)*, Norwell, MA, Kluwer Academic Publishers, pp. 44–72.

Agrawal, V., S. Seshadri (2000). Risk intermediation in supply chains. *IIE Transactions* 32, 819–831.

Ailawadi, K. (2001). The retail power-performance conundrum: what have we learned. *Journal of Retailing* 77, 299–318.

Anand, K., H. Mendelson (1997). Information and organization for horizontal multimarket coordination. *Management Science* 43, 1609–1627.

Anand, K., R. Anupindi, Y. Bassok (2001). Strategic inventories in procurement contracts. University of Pennsylvania Working Paper, Philadelphia, PA.

Andersson, J., J. Marklund (2000). Decentralized inventory control in a two-level distribution system. *European Journal of Operational Research* 127(3), 483–506.

Anupindi, R., R. Akella (1993). An inventory model with commitments. University of Michigan Working Paper.

Anupindi, R., Y. Bassok (1999). Centralization of stocks: retailers vs. manufacturer. *Management Science* 45(2), 178–191.

Anupindi, R., Y. Bassok, E. Zemel (2001). A general framework for the study of decentralized distribution systems. *Manufacturing and Service Operations Management* 3(4), 349–368.

Atkins, D., X. Zhao (2002). Supply chain structure under price and service competition. University of British Columbia Working Paper.

Aviv, Y. (2001). The effect of collaborative forecasting on supply chain performance. *Management Science* 47(10), 1326–1343.

Aviv, Y., A. Federgruen (1998). The operational benefits of information sharing and vendor managed inventory (VMI) programs. Washington University Working Paper, St. Louis, MO.

Axsater, S. (2001). A framework for decentralized multi-echelon inventory control. *IIE Transactions* 33, 91–97.

Baiman, S., P. Fischer, M. Rajan (2000). Information, contracting and quality costs. *Management Science* 46(6), 776–789.

Baiman, S., P. Fischer, M. Rajan (2001). Performance measurement and design in supply chains. *Management Science* 47(1), 173–188.

Barlow, R.E., F. Proschan (1965). *Mathematical Theory of Reliability*, John Wiley & Sons.

Barnes-Schuster, D., Y. Bassok, R. Anupindi (2002). Coordination and flexibility in supply contracts with options. *Manufacturing and Service Operations Management* 4, 171–207.

Bassok, Y., R. Anupindi (1997a). Analysis of supply contracts with total minimum commitment. *IIE Transactions* 29(5).

Bassok, Y., R. Anupindi (1997b). Analysis of supply contracts with commitments and flexibility. University of Southern California Working Paper.

Bernstein, F., G. DeCroix (2002). Decentralized pricing and capacity decisions in a multi-tier system with modular assembly. Duke University Working Paper.

Bernstein, F., A. Federgruen (1999). Pricing and replenishment strategies in a distribution system with competing retailers. Forthcoming *Operations Research*.

Bernstein, F., A. Federgruen (2000). Decentralized supply chains with competing retailers under demand uncertainty. Forthcoming *Management Science*.

Bernstein, F., F. Véricourt (2002). Allocation of supply contracts with service guarantees. Duke University Working Paper.

Bernstein, F., F. Chen, A. Federgruen (2002). Vendor managed inventories and supply chain coordination: the case with one supplier and competing retailers. Duke University Working Paper.

Bloom, P., V. Perry (2001). Retailer power and supplier welfare: the case of Wal-Mart. *Journal of Retailing* 77, 379–396.

Boyaci, T., G. Gallego (1997). Coordination issues in simple supply chains. Columbia University Working Paper, New York, NY.

Bresnahan, T., P. Reiss (1985). Dealer and manufacturer margins. *Rand Journal of Economics* 16(2), 253–268.

Brown, A. (1999). A coordinating supply contract under asymmetric demand information: guaranteeing honest information sharing. Vanderbilt University Working Paper, Nashville, TN.

Brown, A., H. Lee (1998). The win-win nature of options based capacity reservation arrangements. Vanderbilt University Working Paper, Nashville, TN.

Butz, D. (1997). Vertical price controls with uncertain demand. *Journal of Law and Economics* 40, 433–459.

Cachon, G. (1998). Competitive supply chain inventory management, in: S. Tayur, R. Ganeshan, M. Magazine (eds.), *Quantitative Models for Supply Chain Management*, Boston, Kluwer.

Cachon, G. (1999). Competitive and cooperative inventory management in a two-echelon supply chain with lost sales. University of Pennsylvania Working Paper, Philadelphia, PA.

Cachon, G. (2001). Stock wars: inventory competition in a two echelon supply chain. *Operations Research* 49(5), 658–674.

Cachon, G. (2002). The allocation of inventory risk and advanced purchase discounts in a supply chain. University of Pennsylvania Working Paper.

Cachon, G., M. Fisher (2000). Supply chain inventory management and the value of shared information. *Management Science* 46(8), 1032–1048.

Cachon, G., M. Lariviere (1997). Capacity allocation with past sales: when to turn-and-earn. *Management Science* 45(5), 685–703.

Cachon, G., M. Lariviere (1999). Capacity choice and allocation: strategic behavior and supply chain performance. *Management Science* 45(8), 1091–1108.

Cachon, G., M. Lariviere (2000). Supply chain coordination with revenue sharing: strengths and limitations. Forthcoming. *Management Science*.

Cachon, G., M. Lariviere (2001). Contracting to assure supply: how to share demand forecasts in a supply chain. *Management Science* 47(5), 629–646.

Cachon, G., P. Zipkin (1999). Competitive and cooperative inventory policies in a two-stage supply chain. *Management Science* 45(7), 936–953.

Caldentey, R., L. Wein (1999). Analysis of a production-inventory system. Massachusetts Institute of Technology Working Paper, Cambridge, MA.
Chen, F. (1998). Echelon reorder points, installation reorder points, and the value of centralized demand information. *Management Science* 44(12), S221–S234.
Chen, F. (1999). Decentralized supply chains subject to information delays. *Management Science* 45(8), 1076–1090.
Chen, Y. (1999). Oligopoly price discrimination and resale price maintenance. *Rand Journal of Economics* 30(3), 441–455.
Chen, F. (2001). Auctioning supply contracts. Columbia University Working Paper.
Chen, F., A. Federgruen (2000). Mean-variance analysis of basic inventory models. Columbia University Working Paper.
Chen, F., Y.S. Zheng (1994). Lower bounds for multi-echelon stochastic inventory system. *Management Science* 40(11), 1426–1443.
Chen, F., A. Federgruen, Y. Zheng (2001). Coordination mechanisms for decentralized distribution systems with one supplier and multiple retailers. *Management Science* 47(5), 693–708.
Cho, R., Y. Gerchak (2001). Efficiency of independent downstream firm could counteract coordination difficulties. University of Waterloo Working Paper.
Chod, J., N. Rudi (2002). Resource flexibility with responsive pricing.
Choi, K.S., J.G. Dai, J.S. Song (2002). On measuring supplier performance under vendor-managed inventory programs. University of California at Irvine Working Paper.
Chu, W., P. Desai (1995). Channel coordination mechanisms for customer satisfaction. *Marketing Science* 14(4), 343–359.
Clark, A.J., H.E. Scarf (1960). Optimal policies for a multi-echelon inventory problem. *Management Science* 6, 475–490.
Cohen, M., T. Ho, J. Ren, C. Terwiesch (2001). Measuring inputed costs in the semiconductor equipment supply chain. University of Pensylvania Working Paper.
Corbett, C. (2001). Stochastic inventory systems in a supply chain with asymmetric information: cycle stocks, safety stocks, and consignment stock. *Operations Research* 49(4), 487–500.
Corbett, C., G. DeCroix (2001). Shared savings contracts in supply chains. *Management Science* 47(7), 881–893.
Corbett, C., X. de Groote (2000). A supplier's optimal quantity discount policy under asymmetric information. *Management Science* 46(3), 444–450.
Corbett, C., U. Karmarkar (2001). Competition and structure in serial supply chains. *Management Science* 47(7), 966–978.
Corbett, C., C. Tang (1999). Designing supply contracts: contract type and information asymmetry, in: S. Tayur, R. Ganeshan, M. Magazine (eds.), *Quantitative Models for Supply Chain Management*, Boston, Kluwer Academic Publishers.
Dana, J., K. Spier (2001). Revenue sharing and vertical control in the video rental industry. *The Journal of Industrial Economics* 49(3), 223–245.
Debo, L. (1999). Repeatedly selling to an impatient newsvendor when demand fluctuates: a supergame theoretic framework for co-operation in a supply chain. Carnegie Mellon University Working Paper.
Demougin, D.M. (1989). A renegotiation-proof mechanism for a principal-agent model with moral hazard and adverse selection. *Rand Journal of Economics* 20(2), 256–267.
Deneckere, R., H. Marvel, J. Peck (1996). Demand uncertainty, inventories, and resale price maintenance. *Quarterly Journal of Economics* 111, 885–913.
Deneckere, R., H. Marvel, J. Peck (1997). Demand uncertainty and price maintenance: markdowns as destructive competition. *American Economic Review* 87(4), 619–641.
Desai, P., K. Srinivasan (1995). Demand signalling under unobservable effort in franchising. *Management Science* 41(10), 1608–1623.
Desiraju, R., S. Moorthy (1997). Managing a distribution channel under asymmetric information with performance requirements. *Management Science* 43, 1628–1644.
Deshpande, V., L. Schwarz (2002). Optimal capacity allocation in decentralized supply chains. Purdue University Working Paper.

Dolan, R., J.B. Frey (1987). Quantity discounts: managerial issues and research opportunities/commentary/reply. *Marketing Science* 6(1), 1–24.

Dong, L., N. Rudi (2001). Supply chain interaction under transshipments. Washington University Working Paper.

Donohue, K. (2000). Efficient supply contracts for fashion goods with forecast updating and two production modes. *Management Science* 46(11), 1397–1411.

Duenyas, I., C.-Y. Tsai (2001). Centralized and decentralized control of a two-stage tandem manufacturing system with demand for intermediate and end products. University of Michigan Working Paper, Ann Arbor, MI.

Duenyas, I., W. Hopp, Y. Bassok (1997). Production quotas as bounds on interplant JIT contracts. *Management Science* 43(10), 1372–1386.

Eeckhoudt, L., C. Gollier, H. Schlesinger (1995). The risk-averse (and prudent) newsboy. *Management Science* 41(5), 786–794.

Emmons, H., S. Gilbert (1998). Returns policies in pricing and inventory decisions for catalogue goods. *Management Science* 44(2), 276–283.

Eppen, G., A. Iyer (1997). Backup agreements in fashion buying – the value of upstream flexibility. *Management Science* 43(11), 1469–1484.

Erhun, F., P. Keskinocak, S. Tayur (2000). Spot markets for capacity and supply chain coordination. Carnegie Mellon Working Paper, Pittsburgh, PA.

Erkoc, M., S.D. Wu (2002a). Managing high-tech capacity expansion via reservation contracts. Lehigh University Working Paper.

Erkoc, M., S.D. Wu (2002b). Due-date coordination in an internal market via risk sharing. Lehigh University Working Paper.

Ertogral, K., D. Wu (2001). A bargaining game for supply chain contracting. Lehigh University Working Paper.

Federgruen, A., P. Zipkin (1984). Computational issues in an infinite-horizon, multiechelon inventory model. *Operations Research* 32(4), 818–836.

Ferguson, M., G. DeCroix, P. Zipkin (2002). When to commit in a multi-echelon supply chain with partial information updating. Duke University Working Paper.

Fisher, M., A. Raman (1996). Reducing the cost of demand uncertainty through accurate response to early sales. *Operations Research* 44, 87–99.

Friedman, J. (1986). *Game Theory with Applications to Economics*, Oxford, Oxford University Press.

Gallini, N., N. Lutz (1992). Dual distribution and royalty fees in franchising. *The Journal of Law, Economics, and Organization* 8, 471–501.

Gans, N. (2002). Customer loyalty and supplier quality competition. *Management Science* 48(2), 207–221.

Gavirneni, S., R. Kapuscinski, S. Tayur (1999). Value of information in capacitated supply chains. *Management Science* 45(1), 16–24.

Gerchak, Y., Y. Wang (1999). Coordination in decentralized assembly systems with random demand. University of Waterloo Working Paper, Waterloo, Ontario, Canada.

Gerchak, Y., R. Cho, S. Ray (2001). Coordination and dynamic shelf-space management of video movie rentals. University of Waterloo Working Paper, Waterloo, Ontario.

Gilbert, S., V. Cvsa (2000). Strategic supply chain contracting to stimulate downstream process innovation. University of Texas at Austin Working Paper.

Gurnani, H., Y. Gerchak (1998). Coordination in decentralized assembly systems with uncertain component yield. University of Waterloo Working Paper, Waterloo, Ontario, Canada.

Ha, A. (2001). Supplier-buyer contracting: asymmetric cost information and the cut-off level policy for buyer participation. *Naval Research Logistics* 48(1), 41–64.

Ha, A., L. Li, S.-M. Ng (2000). Price and delivery logistics competition in a supply chain. Yale School of Management Working Paper, New Haven, CT.

Holden, S. (1999). Renegotiation and the efficiency of investments. *Rand Journal of Economics* 30(1), 106–119.

Holmstrom, B. (1979). Moral hazard and observability. *Bell Journal of Economics* 10(1), 74–91.

Holmstrom, B. (1982). Moral hazard in teams. *Bell Journal of Economics* 13(2), 324–340.

Ippolito, P. (1991). Resale price maintenance: economic evidence from litigation. *Journal of Law and Economics* 34, 263–294.

Iyer, A., M. Bergen (1997). Quick response in manufacturer-retailer channels. *Management Science* 43(4), 559–570.

Jeuland, A., S. Shugan (1983). Managing channel profits. *Marketing Science* 2, 239–272.

Jin, M., S.D. Wu (2002). Procurement auctions with supplier coalitions: validity requirements and mechanism design. Lehigh University Working Paper.

Karjalainen, R. (1992). The newsboy game. University of Pennsylvania Working Paper, Philadelphia, PA.

Katz, M. (1989). Vertical contractual relations, in: R. Schmalensee, R. Willig (eds.), *Handbook of Industrial Organization*, Vol. 1. Boston, North-Holland.

Kolay, S., G. Shaffer (2002). All-unit discounts in retail contracts. University of Rochester Working Paper.

Kouvelis, P., G. Gutierrez (1997). The newsvendor problem in a global market: optimal centralized and decentralized control policies for a two-market stochastic inventory system. *Management Science* 43(5), 571–585.

Kouvelis, P., M. Lariviere (2000). Decentralizing cross-functional decisions: coordination through internal markets. *Management Science* 46(8), 1049–1058.

Kranton, R., D. Minehart (2001). A theory of buyer-seller networks. *American Economic Review* 91(3), 485–508.

Krasa, S., A. Villamil (2000). Optimal contracts when enforcement is a decision variable. *Econometrica* 68(1), 119–134.

Kreps, D.M. (1990). *A Course in Microeconomic Theory*, Princton, Princeton University Press.

Krishnan, H., R. Kapuscinski, D. Butz (2001). Coordinating contracts for decentralized supply chains with retailer promotional effort. University of Michigan Working Paper.

Lafountaine, F. (1992). Agency theory and franchising: some empirical results. *Rand Journal of Economics* 23(2), 263–283.

Lafontaine, F., M. Slade (2001). Incentive contracting and the franchise decision, in: K. Chatterjee, W. Samuelson (eds.), *Game Theory and Business Applications*, Boston, Kluwer Academic Publishing.

Lal, R. (1990). Improving channel coordination through franchising. *Marketing Science* 9, 299–318.

Lariviere, M. (1998). Supply chain contracting and co-ordination with stochastic demand, in: S. Tayur, R. Ganeshan, M. Magazine (eds.), *Quantitative Models for Supply Chain Management*, Boston, Kluwer.

Lariviere, M. (2002). Inducing forecast revelation through restricted returns. Northwestern University Working Paper.

Lariviere, M., V. Padmanabhan (1997). Slotting allowances and new product introductions. *Marketing Science* 16, 112–128.

Lariviere, M., E. Porteus (2001). Selling to the newsvendor: an analysis of price-only contracts. *Manufacturing and Service Operations Management* 3(4), 293–305.

Lee, H. (1987). A multi-echelon inventory model for repairable items with emergency lateral transshipments. *Management Science* 33(10), 1302–1316.

Lee, H. (1996). Effective inventory and service management through product and process redesign. *Operations Research* 44(1), 151–159.

Lee, H., S. Whang (1999). Decentralized multi-echelon supply chains: incentives and information. *Management Science* 45(5), 633–640.

Lee, H., S. Whang (2002). The impact of the secondary market on the supply chain. *Management Science* 48, 719–731.

Lee, H., V. Padmanabhan, T. Taylor, S. Whang (2000). Price protection in the personal computer industry. *Management Science* 46(4), 467–482.

Lee, H., K.C. So, C. Tang (2000). The value of information sharing in a two-level supply chain. *Management Science* 46, 626–643.

Lim, W. (2001). Producer-supplier contracts with incomplete information. *Management Science* 47(5), 709–715.

Lippman, S., K. McCardle (1997). The competitive newsboy. *Operations Research* 45, 54–65.

Lu, X., J.S. Song, A. Regan (2002). Rebate, returns and price protection policies in supply chain coordination. University of California at Irvine Working Paper.

Mahajan, S., G. van Ryzin (1999). Retail inventories and consumer choice, in: S. Tayur, R. Ganeshan, M. Magazine (eds.), *Quantitative Models for Supply Chain Management*, Boston, Kluwer.

Mahajan, S., G. van Ryzin (2001). Inventory competition under dynamic consumer choice. *Operations Research* 49(5), 646–657.

Majumder, P., H. Groenevelt (2001). Competition in remanufacturing. *Production and Operations Management* 10(2), 125–141.

Marvel, H., J. Peck (1995). Demand uncertainty and returns policies. *International Economic Review* 36(3), 691–714.

Marx, L., G. Shaffer (2001a). Bargaining power in sequential contracting. University of Rochester Working Paper.

Marx, L., G. Shaffer (2001b). Rent shifting and efficiency in sequential contracting. University of Rochester Working Paper.

Marx, L., G. Shaffer (2002). Base contracts in games with ex-post observability. University of Rochester Working Paper.

Messinger, P., C. Narasimhan (1995). Has power shifted in the grocery channel. *Marketing Science* 14(2), 189–223.

Milner, J., E. Pinker (2001). Contingent labor contracting under demand and supply uncertainty. *Management Science* 47(8), 1046–1062.

Moinzadeh, K., S. Nahmias (2000). Adjustment strategies for a fixed delivery contract. *Operations Research* 48(3), 408–423.

Moorthy, K.S. (1987). Managing channel profits: comment. *Marketing Science* 6, 375–379.

Mortimer, J.H. (2000). The effects of revenue-sharing contracts on welfare in vertically separated markets: evidence from the video rental industry. University of California at Los Angeles Working Paper, Los Angeles, CA.

Moses, M., S. Seshadri (2000). Policy mechanisms for supply chain coordination. *IIE Transactions* 32, 245–262.

Myerson, R. (1979). Incentive compatibility and the bargaining problem. *Econometrica* 47, 399–404.

Nahmias, S. (1993). *Production and Operations Analysis*, Boston, Irwin.

Narayanan, V., A. Raman (1997). Contracting for inventory in a distribution channel with stochastic demand and substitute products. Harvard University Working Paper.

Netessine, S., N. Rudi (2000). Supply chain structures on the internet: marketing-operations coordination. University of Pennsylvania Working Paper, Philadelphia, PA.

Netessine, S., N. Rudi (2000b). Centralized and competitive inventory models with demand substitution. Forthcoming, *Operations Research*.

Netessine, S., R. Shumsky (2001). Revenue management games. University of Pennsylvania Working Paper, Philadelphia, PA.

Novak, S., S. Eppinger (2001). Sourcing by design: product complexity and the supply chain. *Management Science* 47(1), 189–204.

O'Brien, D., G. Shaffer (1992). Vertical control with bilateral contracts. *Rand Journal of Economics* 23, 299–308.

Padmanabhan, V., I.P.L. Png (1995). Returns policies: make money by making good. *Sloan Management Review* Fall, 65–72.

Padmanabhan, V., I.P.L. Png (1997). Manufacturer's returns policy and retail competition. *Marketing Science* 16(1), 81–94.

Parker, R., R. Kapuscinsky (2001). Managing a non-cooperative supply chain with limited capacity. University of Michigan Working Paper.

Parlar, M. (1988). Game theoretic analysis of the substitutable product inventory problem with random demands. *Naval Research Logistics Quarterly* 35, 397–409.

Pasternack, B. (1985). Optimal pricing and returns policies for perishable commodities. *Marketing Science* 4(2), 166–176.
Pasternack, B. (1999). Using revenue sharing to achieve channel coordination for a newsboy type inventory model. CSU Fullerton Working Paper, Fullerton.
Petruzzi, N., M. Dada (1999). Pricing and the newsvendor problem: a review with extensions. *Operations Research* 47, 183–194.
Plambeck, E., T. Taylor (2002). Sell the plant? The impact of contract manufacturing on innovation, capacity and profitability. Stanford University Working Paper.
Plambeck, E., S. Zenios (2000). Performance-based incentives in a dynamic principal-agent model. *Manufacturing and Service Operations Management* 2, 240–263.
Porteus, E. (2000). Responsibility tokens in supply chain management. *Manufacturing and Service Operations Management* 2(2), 203–219.
Porteus, E., S. Whang (1991). On manufacturing/marketing incentives. *Management Science* 37(9), 1166–1181.
Porteus, E., S. Whang (1999). Supply chain contracting: non-recurring engineering charge, minimum order quantity, and boilerplate contracts. Stanford University Working Paper.
Randall, R., S. Netessine, N. Rudi (2002). Inventory structure and internet retailing: an empirical examination of the role of inventory ownership. University of Utah Working Paper.
Reyniers, D., C. Tapiero (1995). The delivery and control of quality in supplier-producer contracts. *Management Science* 41(10), 1581–1589.
Riordan, M. (1984). Uncertainty, asymmetric information and bilateral contracts. *Review of Economic Studies* 51, 83–93.
Rochet, J., L. Stole (2002). Nonlinear pricing with random participation. *Review of Economic Studies* 69, 277–311.
Roundy, R. (1985). 98%-effective integer ratio lot-sizing for one-warehouse multi-retailer systems. *Management Science* 31, 1416–1430.
Rubinstein, A. (1989). The electronic mail game: strategic behavior under 'almost common knowledge'. *American Economic Review* 79(3), 385–391.
Rudi, N., S. Kapur, D. Pyke (2001). A two-location inventory model with transhipment and local decision making. *Management Science* 47(12), 1668–1680.
Scheller-Wolf, A., S. Tayur (1997). Reducing international risk through quantity contracts. Carnegie Mellon Working Paper, Pittsburgh, PA.
Schweitzer, M., C. Cachon (2000). Decision bias in the newsvendor problem: experimental evidence. *Management Science* 46(3), 404–420.
Shaffer, G. (1991). Slotting allowances and resale price maintenance: a comparison of facilitating practices. *Rand Journal of Economics* 22, 120–135.
Silver, E., D. Pyke, R. Peterson (1998). *Inventory Management and Production Planning and Scheduling*, New York, John Wiley and Sons.
Spengler, J. (1950). Vertical integration and antitrust policy. *Journal of Political Economy*, 347–352.
Stidham, S. (1992). Pricing and capacity decisions for a service facility: stability and multiple local optima. *Management Science* 38(8), 1121–1139.
Tagaras, G., M.A. Cohen (1992). Pooling in two-location inventory systems with non-negligible replenishment lead times. *Management Science* 38(8), 1067–1083.
Taylor, T. (2001). Channel coordination under price protection, midlife returns and end-of-life returns in dynamic markets. *Management Science* 47(9), 1220–1234.
Taylor, T. (2002). Coordination under channel rebates with sales effort effect. *Management Science* 48(8), 992–1007.
Taylor, T. (2002b). Sale timing in a supply chain: when to sell to the retailer. Columbia University Working Paper.
Tayur, S., Ganeshan, R., Magazine, M. (eds.) (1998). *Quantitative Models for Supply Chain Management*, Boston, Kluwer.
Terwiesch, C., C. Loch (2002). Collaborative prototyping and the pricing of custom designed products. University of Pennsylvania Working Paper.

Terwiesch, C., J. Ren, T. Ho, M. Cohen (2002). An empirical analysis of forecast sharing in the semiconductor equipment supply chain. University of Pennsylvania Working Paper.

Tirole, J. (1986). Procurement and renegotiation. *Journal of Political Economy* 94(2), 235–259.

Tomlin, B. (2000). Capacity investments in supply chain: sharing-the-gain rather than sharing-the-pain. University of North Carolina Working Paper.

Topkis, D. (1998). *Supermodularity and Complementarity*, Princeton, Princeton University Press.

Tsay, A. (1999). Quantity-flexibility contract and supplier-customer incentives. *Management Science* 45(10), 1339–1358.

Tsay, A. (2001). Managing retail channel overstock: markdown money and return policies. *Journal of Retailing* 77, 457–492.

Tsay, A. (2002). Risk sensitivity in distribution channel partnership: implications for manufacturer return policies. *Journal of Retailing* 78, 147–160.

Tsay, A., N. Agrawal (1999). Channel conflict and coordination: an investigation of supply chain design. Santa Clara University Working Paper, Santa Clara, CA.

Tsay, A., N. Agrawal (2000). Channel dynamics under price and service competition. *Manufacturing and Service Operations Management*. 2(4), 372–391.

Tsay, A., W.S. Lovejoy (1999). Quantity-flexibility contracts and supply chain performance. *Manufacturing and Service Operations Management* 1(2), 89–111.

Tsay, A., S. Nahmias, N. Agrawal (1998). Modeling supply chain contracts: a review, in: S. Tayur, R. Ganeshan, M. Magazine (eds.), *Quantitative Models for Supply Chain Management*, Boston, Kluwer.

van Mieghem, J. (1999). Coordinating investment, production and subcontracting. *Management Science* 45(7), 954–971.

van Mieghem, J., M. Dada (1999). Price vs production postponement: capacity and competition. *Management Science* 45(12), 1631–1649.

van Ryzin, G., S. Mahajan (1999). Supply chain coordination under horizontal competition. Columbia University Working Paper, New York, NY.

Wang, Y., Y. Gerchak (2001). Supply chain coordination when demand is shelf-space dependent. *Manufacturing and Service Operations Management* 3(1), 82–87.

Wang, Y., L. Jiang, Z.J. Shen (2001). Consignment sales, price-protection decisions and channel performances. Case Western Reserve University Working Paper.

Watson, Noel (2002). Execution in supply chain management: dynamics, mis-steps and mitigation strategies. University of Pennsylvania Dissertation.

Webster, S., S.K. Weng (2000). A risk-free perishable item returns policy. *Manufacturing and Service Operations Management* 2(1), 100–106.

Whang, S. (1995). Coordination in operations: a taxonomy. *Journal of Operations Management* 12, 413–422.

Wilson, R. (1993). *Nonlinear Pricing*, Oxford, Oxford University Press.

Wu, D.J., P. Kleindorfer, J. Zhang (2002). Optimal bidding and contracting strategies for capital-intensive goods. *European Journal of Operational Research* 137, 657–676.

Yüksel, O., H. Lee (2002). Sharing inventory risks for customized components. Stanford University Working Paper.

Zhao, H., V. Deshpande, J. Ryan (2002). Inventory sharing and rationing in decentralized dealer networks. Purdue University Working Paper.