Optimal Supply Chain Contracts under Asymmetric Information*

Ying-Ju Chen † Zhengqing Gui ‡ Ernst-Ludwig von Thadden §
Xiaojian Zhao ¶

September 2022

Abstract

A central problem in vertical relationships is to coordinate the mismatch between supply and demand. This paper studies a problem of contracting between a manufacturer and a retailer who privately observes the retail demand materialized after the contracting stage. Under very general assumptions, we show that the optimal contract can be implemented by either a wholesale contract or a buyback contract, depending on the retailer's ex-ante liquidity constraint and ex-post bargaining power. In a buyback contract, the manufacturer requests an upfront payment from the retailer and buys back the unsold inventories with possibly nonlinear buyback price. Since return shipments are inefficient, retail supply and price will be lower than the first-best level. The optimal contracts are robust to several scenarios including multiple retailers.

Keywords: Retail contracts, return policies, buyback contracts, incentive problems, limited liability.

JEL Classification: D82, D86, L42, L60.

---

*von Thadden gratefully acknowledges financial support from the German Research Foundation (DFG) through CRC TR 224, C03.

†School of Business and Management (ISOM), Hong Kong University of Science and Technology. Email: imchen@ust.hk.

‡Risk Management Institute, National University of Singapore, and Economics and Management School, Wuhan University. Email: zgui@nus.edu.sg.

§University of Mannheim and CEPR. Email: vthadden@uni-mannheim.de.

¶Monash University. Email: xjzhao81@gmail.com.

1 Introduction

A central problem in vertical contracting is to minimize the mismatch between supply and demand. Manufacturers, in many situations, have to rely on retailers to sell their products in the market. Due to unavoidable long lead times, production by manufacturers must occur before retail demand is realized, and supply-demand mismatch may arise. When demand is large, the retailer can only sell up to the quantity he has received, and the excess demand is lost. When demand falls short, the unsold inventory may be salvaged by either the retailer or the manufacturer at a discount. The vertical contracting relationship between the manufacturer and the retailer determines the retailer's order quantity, monetary transfers, return policies, and other elements that coordinate the supply chain.

In practice, a number of contracts are used, including buyback, franchise, revenue sharing, wholesale price, fixed transfer contracts, etc. There is large literature investigating these contracts and their consequences on supply chain performance in various contexts (see Cachon (2003) for an extensive discussion). The conceptual question *what is the optimal contract*, however, has received less attention. As Cachon (2003) puts it succinctly: "practice has been used as a motivation for theoretical work, but theoretical work has not found its way into practice".

In this paper, we present a model in which a manufacturer sells its products through a retailer to meet the market demand. Two factors are key to our model. First, production precedes sales, thus contracting parties have to write down all the terms, including the price and the quantity to be produced before any demand uncertainty is resolved. This assumption is widely used in the literature of vertical contracting (e.g., Deneckere et al., 1996, 1997; Montez, 2015) and is the origin of the supply-demand mismatch. Second, the retailer is privately informed about the realized demand, so the manufacturer cannot observe directly the retailer's sales and revenue. As a result, a vertical contract should: (1) specify the appropriate price and quantity as a response to the supply-demand mismatch, and (2) be incentive-compatible so as to induce the retailer to truthfully report the realized demand.

In particular, we study a game with two dates. Contracting happens at date 0, which determines the quantity and an immediate cash transfer. The retail demand is realized at date 1, after which there could be a delayed transfer and return of unsold inventories. The date-1 cash transfer and return shipments are contingent on the retailer's report. Therefore, the contracting parties face the tradeoff between cash and returns. The two channels differ in many aspects. The most immediate one is that cash can happen in both dates, whereas unsold inventories can only be returned after sales. Besides, cash repayments are bounded by the retailer's initial wealth plus his date-1 revenue, which is increasing in the realized demand. This limited liability constraint is considered in the literature studying contracting problems in industrial organization (e.g., Brander and Lewis, 1986) and captures the fundamental feature of small and medium enterprises: They are typically resource-constrained and thus the only collateral that can be pledged is the business value they have created.[1] Return shipments, however, cannot exceed the total amount of leftover inventories, which is negatively related to the demand. Finally, cash transfers are efficient, while we assume that returning unsold inventories leads to deadweight loss.[2] This assumption, together with the retailer's limited liability, makes the return of unsold inventories a screening device. As a result, the manufacturer's objective is to minimize the use of returns with an admissible and incentive-compatible contract.

Without assuming any functional form of contracts, very generally, we find that the optimal contract takes a rather simple form: The retailer has an obligation to make an upfront payment at date 1. If the lower bound of his revenue can cover this payment, the optimal contract is essentially a wholesale contract. Otherwise, the optimal contract is a buyback contract. In the wholesale contract, the retailer's cash transfer is invariant to the realized demand, and there is no return shipment. In the buyback contract, the retailer transfers a fixed amount of cash to the manufacturer when the realized demand is high, and returns part of unsold inventories to the manufacturer when the realized demand is too low for the retailer to pay the fixed amount in full. The rationale for this result is the following: Facing the adverse selection problem, the manufacturer wants to elicit the retailer's private information on the retail demand, so the return of unsold inventory is used as a punishment when the reported demand is low. However, the manufacturer, who has full bargaining power, also aims to minimize unsold inventories returned by the retailer, since it leads to efficiency loss. Therefore, return policies will be offered only when the reported demand is sufficiently low. Notably, the buyback price may not be constant, thus optimal contracts are indeed buyback contracts with nonlinear pricing. The wholesale contract turns out to be a special case.

Proposition 4 states our main results. It shows that the optimal contract shifts from buyback to wholesale as the retailer's bargaining power increases. Therefore, our paper can be viewed as a unified micro-foundation for both wholesale contracts and buyback contracts in the real world. Moreover, the optimal quantity is strictly lower than the first-best level

[1] Limited liability is sometimes used in lieu of risk-aversion. The latter assumption is also in line with our focus on small and medium retailers.

[2] In part of the literature, the salvage value of unsold inventories is assumed to be zero (e.g., Marvel and Peck, 1995; Arya and Mittendorf, 2004). We argue that this is less realistic in situations with many non-perishable goods, such as clothes and electronic devices. Even for perishable goods, returning unsold inventory to the manufacturer may involve certain transportation costs, leading the retailer to be strictly more efficient in keeping unsold inventories.

if the optimal contract is a buyback contract. That is, supply is rationed as a response to information asymmetry. Intuitively, the manufacturer has a bias towards high-demand states to avoid inefficient return shipments, which give her extra incentive to reduce the retail supply. We then discuss two key assumptions of our model: limited liability and incentive-compatibility, and show how our predictions may change if we relax any of them.

We finally extend the benchmark model in various aspects. When the retail price is endogenous and influences the distribution of demand, Proposition 6 shows that the contract optimally makes the price lower than the first-best level. In other words, information asymmetry restricts the manufacturer's market power. When the retailer is allowed to reorder additional quantities after observing the realized demand, Proposition 7 says that the optimal contract is a weighted combination of wholesale and buyback. The weights are determined by the depreciation rate of unfulfilled demand. When the manufacturer contracts with multiple retailers, Proposition 8 characterizes a set of symmetric optimal contracts. The "sum" of these contracts is equivalent to the optimal contract in a single-retailer model in which the retailer has larger bargaining power. It is as if the retailers are merged into one big entity and contract with the manufacturer, after which they split the contract terms equally. Based on this observation, we state in Corollary 1 that when the number of retailers are sufficiently large, optimal contracts will switch from buyback to wholesale contracts without returns, and price and quantity will go back to the first-best level. Put differently, introducing extra retailers in a vertical relationship may push the market supply back to the efficient level, but will be accompanied by an increase in price.

1.1 Related literature

Our paper contributes to the growing literature on vertical relationships with asymmetric information and uncertain demand. The vast majority of this literature focuses on vertical restraints (e.g., Winter, 1993; Deneckere et al., 1996, 1997; Dana and Spier, 2001; Harstad and Mideksa, 2020), i.e., how (and why) the manufacturer controls the retail price, order quantity, or competition between retailers. In these papers, retail contracts are usually given by two-part tariffs, or revenue-sharing schemes, plus some specific restraints, such as the Resale Price Maintenance. However, a fundamental question of whether theses contractual forms are indeed optimal remains a problem. A related stream of literature studies the newsvendor problem(e.g., Pasternack, 1985; Marvel and Peck, 1995; Krishnan and Winter, 2007), in which the manufacturer proposes a contract to induce the retailer choosing the optimal price and inventory. More recently, Montez and Schutz (2021) introduce production-in-advance as inventory to serve demand in the oligopoly games that echoes all-pay contests. However, in this line of models, the retailer's payments to the manufacturer are independent of the realized demand. Thus, the adverse selection as a result of the limited liability problem due to insufficient demand realization has been assumed away. In a related literature (e.g., Rey and Tirole, 1986; Blair and Lewis, 1994), optimal retail contracts are derived under demand uncertainty, but they do not consider the prescribed newsvendor problem. Wang et al. (2020) examine the signaling role of buyback contracts, while they take the buyback form as exogenously given. The paper perhaps closest to the one presented here is by Arya and Mittendorf (2004), in which the manufacturer uses a return allowance to elicit the retailer's private information on demand. In their model the return policy is characterized by the price offered by the manufacturer, so the retailer's choice is all-or-nothing: Full return if the return allowance is higher than the retail price, zero return otherwise. In the present paper, the contract determines the quantity to be returned, so partial return polices are allowed.

Technically, our model is an ex-post screening problem with hidden characteristics. When the type set is a continuum, as in our model, the standard methodology is to use control theory. This approach is pioneered by Guesnerie and Laffont (1984), and further developed by Hellwig (2010) who proposes a unified approach that only requires the compactness of the type set, and allows for mass points. However, the control-theoretic approach cannot be applied in the present paper. In our model, each type of the retailer's set of deviation is bounded by the limited liability and the feasibility constraint, and thus depends on the endogenous contract. Therefore, the retailer's incentive constraint cannot be simplified into a local differential align. This feature is similar to the financial contracting literature by Townsend (1979) and Gale and Hellwig (1985), but in their settings there is no feasibility constraint, which substantially complicates the problem in our retail contracting context. Relatedly, Gui et al. (2019) provide a detailed discussion on how the presence of limited liability affects the analysis of incentive constraint in the financial contracting literature. In particular, overlooking the role of limited liability in specifying the incentive constraint may lead to an over-simplified analysis and sub-optimal contracts. This paper also shows that relaxing the limited liability constraint ex-post may be harmful to firms.

Our paper is part of a more general approach that tries to provide the foundation of observed economics or financial institutions as outcomes of optimal contracting (see, e.g., Nöldeke and Schmidt (1995) in the context of buyer-seller relationships, Aghion and Tirole (1997) in a model of hierarchical authority, or Schmidt (2003) for venture capital arrangements). Our paper follows a similar vein to justify the buyback contract in the retail contract setting without imposing any functional form assumptions on the contract space. Some researchers have also demonstrated, in various contexts, that simple practical contracts seem to perform well even if they are known to be sub-optimal (e.g., Bower, 1993; Rogerson, 2003;

Chu and Sappington, 2007). Unlike these papers, we show that the popular buyback contract in practice may indeed be the optimal contract form.

The rest of this paper is organized as follows. Section 2 introduces the model setup. Section 3 proves the optimality of buyback contracts. Section 4 discusses several key assumptions of the model. Section 5 extends our benchmark model to different environments. Section 6 concludes. Most of the proofs are relegated to the appendix.

2 Model

A manufacturer (she) contracts with a retailer (he) on the delivery of a homogeneous product. Production has no fixed costs and constant marginal costs $c > 0$. Given any retail price p, retail demand ω is stochastic and characterized by the distribution function $F(\cdot; p)$ over $[0, +\infty)$.[3] In our baseline model we assume that p is exogenous and observable. We therefore drop the reference to p in this section and the next.

Retail demand ω is realized after the quantity q has been produced and delivered to the retailer and can only be observed freely by the retailer. The manufacturer only knows $F(\cdot)$. Therefore, two distortions may arise. First, production must take place prior to the realization of demand, thus there will be a supply-demand mismatch. In Section 5.2 we extent our model to allow for reordering to address this mismatch. Second, the realization of demand is the retailer's private information, so all contractual obligations after sales must be incentive-compatible for the retailer. By applying the Revelation Principle, we focus on direct mechanisms in which the retailer simply reports his demand (or type) $\hat{\omega}$ and the contract is executed correspondingly.

After observing ω, the retailer determines the volume of sales, s. When there is supply shortage, i.e., $q < \omega$, the retailer can only sell up to the quantity q, and the excess demand is lost. When there is insufficient demand, i.e., $q \geq \omega$, the retailer can only sell up to ω. Thus realized sales satisfy

$$s \in [0, \min(\omega, q)]. \tag{FS}$$

s, too, is unobservable to the manufacturer.

Moreover, the retailer is able to salvage unsold inventories at a constant salvage value v_r per unit. If instead the manufacturer possesses unsold inventories, her per unit salvage value

[3]This formulation of stochastic outcomes, the "parameterized-distribution-function approach", was pioneered by Mirrlees (1974) and Holmström (1979). Early traces in the IO/price theory literature include Burns and Walsh (1981). See Section 5 for a brief structural discussion of this model.

is v_m. However, the retailer's salvage value is higher than the manufacturer's, i.e., $v_m < v_r$. Thus, it is more efficient for the retailer to keep unsold inventories.[4] To make the analysis non-trivial, we assume

$$p > c > v_r. \tag{1}$$

Hence, producing to sell at salvage value is not profitable, but normal sales are profitable and are known to be profitable.

A contract $\Gamma = (q, T_0, s, T_1, R)$ specifies: (1) the quantity q delivered to the retailer; (2) the cash transfer from the retailer to the manufacturer before sales (T_0), (3) the level of sales by the retailer (s), (4) after sales payments by the retailer to the manufacturer (T_1); (5) the return shipment of unsold inventory R. The last three components depend on ω and are therefore to be understood as functions.[5] This description captures many different types of retail contracts in practice.

Example 1 (Wholesale price). In a wholesale price contract, the manufacturer charges the retailer a constant wholesale price p_w per unit purchased at date 0, with no state-contingent transfer at date 1. The corresponding transfers and returns are, respectively,

$$T_0 + T_1 = p_w q, \;\; R = 0.$$

Example 2 (buyback). In a buyback contract, the manufacturer charges the wholesale price p_w and buys back unsold units at the price $b < p_w$ per unit. Therefore,

$$T_0 = p_w q, \;\; T_1 = -bR, \;\; R = q - s.$$

Example 3 (Revenue sharing). In a revenue sharing contract, in addition to the wholesale price, the manufacturer also obtains a fraction α of the retailer's revenue. In this case,

$$T_0 = p_w q, \;\; T_1 = \alpha p s, \;\; R = 0.$$

If demand is commonly observable and the retailer faces no limited liability, all these contracts are enforceable. We assume that this is not possible, as the manufacturer does not

[4]In the framework developed here, this assumption can hold even if $v_r \leq v_m$ as long as transportation costs for return shipments are taken into account.

[5]As usual, when there is no risk of confusion, we shall denote the quantity (a number) and the function (mapping ω into these quantities) by the same symbol.

have enough information about the retailer's activity. As discussed in the introduction, this is the case in many applications in practice.

The timing of the game is depicted in Figure 1, where we use "m" to represent the manufacturer and "r" to represent the retailer. At date 0, the manufacturer offers the retailer a take-it-or-leave-it contract Γ. If the retailer accepts the contract, he makes an initial payment T_0 to the manufacturer in exchange for the delivery of q units of the product. At date 1, retail demand ω is realized. The retailer observes ω and sells the quantity s. He then makes a report $\hat{\omega}$ to the manufacturer, pays her T_1, and returns R units, based on $\hat{\omega}$.

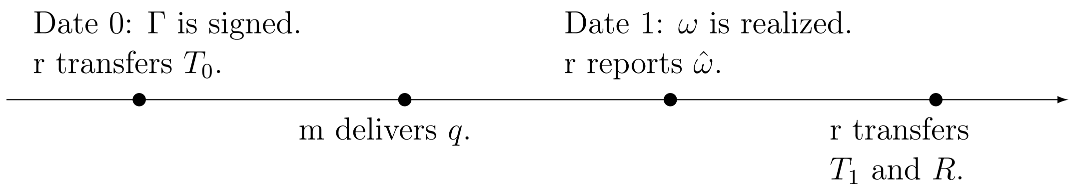


Figure 1: Timeline

For simplicity, we assume that both contracting parties are risk-neutral, and there is no discounting. Let the retailer's initial wealth be $W \geq 0$. Under contract Γ, the retailer's ex-post profit from realized demand ω, reported demand $\hat{\omega}$ and sales s is

$$u_r(\omega, \hat{\omega}, s) = W - T_0 + ps - T_1(\hat{\omega}) + v_r[q - s - R(\hat{\omega})].$$

Here, $W - T_0$ is the retailer's cash position at date 0, ps is the gross revenue from sales, so $ps - T_1(\hat{\omega})$ is the retailer's cash flow at date 1, and $v_r[q - s - R(\hat{\omega})]$ is the salvage value of the retailer's inventory after returns.

Since the manufacturer has no fixed costs and constant marginal costs, her ex-post payoff is

$$u_m(\hat{\omega}) = T_0 - cq + T_1(\hat{\omega}) + v_m R(\hat{\omega}).$$

Note that the problem is one of private values: the manufacturer is exposed to the demand shock ω only through the retailer's ex-post actions (T_1, R).

Ex-post, for each realization of ω, the retailer chooses not only his report strategically optimally, but also his sales level s. The contracting problem becomes interesting because of the liquidity and information restrictions faced by the retailer.

The first such restriction is that the retailer cannot return more than the amount of unsold inventory he has and cannot re-order after strong demand. This implies the following feasibility constraint:

$$0 \leq R \leq q - s. \tag{FR}$$

Second, the retailer cannot pay the manufacturer more than what he has at any time of the game. This implies the following liquidity constraints, at dates 0 and 1, respectively:

$$T_0 \leq W, \tag{L_0}$$

$$T_1 \leq W - T_0 + ps. \tag{L_1}$$

In practice, liquidity constraints (or limited liability) arises for various reasons, such as the retailer's inability to raise additional external finance, his option to quit the relationship ex-post, or legislation banning exploitative contracts. Note that we do not consider the salvage value of unsold inventory on the right-hand side of (L_1), because in practice liquidating leftover inventory typically takes time. In Section 4.1, we discuss a variation of the model where the retailer can use cash generated by salvaging.

Third, sales by the retailer must satisfy the feasibility constraint (FS). Fourth, by the revelation principle, the retailer must have the incentive to report his type ω truthfully. And fifth, he must carry out sales $s(\omega)$ as planned. This leads to the following incentive-compatibility constraint:

$$ps(\omega) - T_1(\omega) + v_r[q - s(\omega) - R(\omega)] \geq p\hat{s} - T_1(\hat{\omega}) + v_r[q - \hat{s} - R(\hat{\omega})] \tag{IC}$$

for all ω, $\hat{\omega}$, and $\hat{s}$ such that

$$0 \leq \hat{s} \leq \min(\omega, q) \tag{IC-FS}$$

$$0 \leq R(\hat{\omega}) \leq q - \hat{s} \tag{IC-FR}$$

$$T_1(\hat{\omega}) \leq W - T_0 + p\hat{s} \tag{IC-L}$$

Note that as the type-ω retailer misreports to be type-$\hat{\omega}$, the transfer and the return shipment change accordingly. Hence, deviations of transfers and returns, $(\hat{T}_1, \hat{R})$, are restricted to lie in the range of the functions T_1, R. However, any deviation of the retailer's ex-post choice of sales, $\hat{s}$, is unobserved and therefore unrestricted, as long as it satisfies the incentive-feasibility constraint (IC-FS). We therefore face a problem of partially verifiable mechanism design, where the disclosure of private information (ω) through observable actions (T_1, R) is obfuscated by some other unobservable action (s).

The incentive constraint (IC)-(IC-L) is remarkable in that it restricts the choice of possible deviations to *feasible lies*. We only require each type of retailer to have no incentive to choose the contract designed for other types when his after-sales wealth, $W - T_0 + p\hat{s}$, and unsold inventory, $q - \hat{s}$, permit this. Hence, not only does the incentive condition (IC) depend on the retailer's type, but also the feasibility of her deviations in (IC-FS)-(IC-L). The overall incentive constraint therefore is weaker than in standard problems where (IC) holds for all $\omega, \hat{\omega}$. Incorporating the qualifications (IC-FS)-(IC-L) makes it difficult to simplify the global incentive constraint to a set of local first-order conditions, and apply the well-established control-theoretic approach to solve for the optimal contract, as in the literature of mechanism design with hidden characteristics.[6] We will discuss how we circumvent this problem in Section 3 and what would happen if we dropped these restrictions to the incentive constraint in Section 4.2.

Finally, the retailer has a monetary outside option, denoted by $W + \underline{u}$. Naturally, $\underline{u} \geq 0$. The manufacturer's outside option is normalized to zero. $\underline{u}$ therefore measures the retailer's relative bargaining power vis-à-vis the manufacturer, who makes the take-it-or-leave-it offer in the contract proposal game. Thus, the contracting parties have the participation constraints

$$\mathrm{E}_\omega\, u_r(\omega, \omega, s(\omega)) \geq W + \underline{u} \tag{PC$_r$}$$

$$\mathrm{E}_\omega\, u_m(\omega) \geq 0. \tag{PC$_m$}$$

Hence, a full statement of the contracting problem between manufacturer and retailer is

$$\begin{aligned} \max_{\Gamma} \quad & \mathrm{E}_\omega\, u_m(\omega) \\ \text{subject to} \quad & \text{(FS), (FR), (L}_0\text{), (L}_1\text{), (IC)-(IC-L), (PC}_r\text{), (PC}_m\text{).} \end{aligned}$$

We call a contract *admissible* if it satisfies the constraints, and *optimal* if it is a solution of this problem. Moreover, if two admissible contracts generate identical payoffs to both contracting parties, we say they are *equivalent*. If an admissible contract Γ generates less expected payoff to the manufacturer than an admissible contract $\hat{\Gamma}$, we say Γ is *dominated* by $\hat{\Gamma}$. And finally, we say that a quantity q can be *implemented* if there is an admissible contract $\Gamma = (q, T_0, s, T_1, R)$.

[6] See the large literature starting from Guesnerie and Laffont (1984).

3 Analysis

3.1 First best

The contracting problem described above features two main frictions: one physical in the sense that quantities must be determined before the realization of demand, the other informational in the sense that demand is private information of the retailer. We take the former friction as given and immutable, and in this section investigate the benchmark of symmetric full information.

In this case, it is not efficient to return merchandise to the manufacturer, as this is value-reducing compared to salvaging by the retailer and provides no informational benefit. Furthermore, sales are equal to maximum feasible demand, $s = \min(\omega, q)$. Social surplus from producing quantity q therefore is

$$S(q) = \int_0^{+\infty} p \min(\omega, q) + v_r(q - \min(\omega, q)) dF(\omega) - cq.$$

Denote by $Q(q)$ the expected feasible demand given q and price p,

$$Q(q) = \int_0^{+\infty} \min(\omega, q) dF(\omega) = q - \int_0^q F(\omega) d\omega, \tag{2}$$

where the last equality follows by partial integration. Then

$$S(q) = (p - v_r)Q(q) - (c - v_r)q. \tag{3}$$

This reformulation of total surplus has a natural interpretation. Since the product can always be salvaged with a per unit value v_r, $p - v_r$ and $c - v_r$ are the "real" price and marginal cost of the retailer, respectively. Therefore, $S(q)$ is similar to the standard profit function of a monopolist facing a demand function $Q(q)$.

We have

$$Q'(q) = 1 - F(q),$$
$$Q''(q) = -f(q).$$

Hence, $Q'(0) > 0$, and it is optimal to produce a positive quantity. By (1), the first-best quantity therefore is uniquely pinned down by the first-order condition.

Proposition 1. *The first-best quantity q^{FB} is unique and satisfies*

$$F(q^{FB}) = \frac{p-c}{p-v_r}. \tag{4}$$

The first-best surplus is

$$S(q^{FB}) = (p-c)q^{FB} - (p-v_r)\int_0^{q^{FB}} F(\omega)d\omega > 0. \tag{5}$$

Proof. (4) follows directly from the first-order condition, (5) from (2) and (3). □

3.2 Second best

For the second-best analysis it is useful to distinguish two parts of a contract. The first part consists of q and T_0, deliveries and transfers that happen at date 0. The second part consists of T_1, R and s; they are functions of ω and $\hat{\omega}$ and are key to the incentive-compatibility constraint. We will sometimes refer to (q, T_0) as the *date-0 component*, and the triple (s, T_1, R) as the *date-1 component*. It is important to realize that the choice of s and $\hat{\omega}$ at date 1 must be optimal for each ω given the schedules T_1 and R, since the retailer has private information and no commitment power.

3.2.1 Implementation by wholesale contract

It has been widely recognized by the literature that wholesale contracts have the feature of making non state-contingent transfers. In our framework, the date-0 component cannot depend on the state, while the date-1 component is non-state-contingent only if

$$T_1(\omega) = T_1 \text{ for all } \omega, \text{ and } T_1 \leq W - T_0,$$
$$R(\omega) = 0 \text{ for all } \omega.$$

In this case, the contract has no costly incentive, so social surplus is split between contracting parties without efficiency loss. More generally, the following characterization of wholesale contracts is useful.

Lemma 1. *The quantity q can be implemented by a wholesale contract with full surplus extraction if and only if*

$$S(q) + cq \leq W + \underline{u}. \tag{6}$$

Proof. Under a wholesale contract that implements q, the retailer gets

$$u_r = W - T_0 + S(q) + cq - T_1.$$

Under the constraint $T_1 \leq W - T_0$, the total payment $T_0 + T_1$ necessary to achieve a binding (PC_r) is feasible if and only if (6) holds. □

Lemma 1 characterizes the situations in which the retailer's initial liquidity constraint (L_0) does not bind. This occurs if she either has sufficient funds or sufficiently high bargaining power.

Note that the left hand side of (6) is strictly monotone in q. Let $\underline{q} = \underline{q}(W, \underline{u})$ be the greatest q for which (6) holds. Hence, if $q^{FB} \leq \underline{q}$ the first-best can be implemented by a wholesale contract. If $q^{FB} > \underline{q}$, the first-best cannot be implemented by a wholesale contract with full surplus extraction. By Lemma 1 the condition $q^{FB} \leq \underline{q}$ is therefore necessary and sufficient for the first-best to be second-best optimal. By (5) and (6), this condition is equivalent to

$$pq^{FB} - (p - v_r) \int_0^{q^{FB}} F(\omega) d\omega \leq W + \underline{u}. \tag{7}$$

If (7) does not hold, the manufacturer has two options. First, she can implement the quantity $\underline{q}$ by an optimal wholesale contract with full surplus extraction, as given in Lemma 1. And second, she can offer a contract that implements a quantity $q > \underline{q}$ by requiring some ex-post state-contingent repayment T_1 supported by costly incentives.[7]

3.2.2 Sales

Next, we characterize the second-best contract implementing $q > \underline{q}$ through costly incentives, taking $q^{FB} > \underline{q}$ as given. For this we should start from the retailer's sales decision. Ex-post we have

$$u_r(\omega, \hat{\omega}, s) = W - T_0 + ps - T_1(\hat{\omega}) + v_r[q - s - R(\hat{\omega})] \tag{8}$$

$$= W - T_0 + (p - v_r)s + v_r[q - R(\hat{\omega})] - T_1(\hat{\omega}). \tag{9}$$

As (9) shows, this is strictly increasing in s for each ω, and ex-post payments T_1 would not be affected by increasing s. However, as (8) shows, if the contract is incentive-compatible, increasing s is only possible if $s < q - R(\omega)$, which may not hold due to the overall contract

[7] Clearly, the option of supplying q^{FB} and leaving some rents over and above $W + \underline{u}$ to the retailer is dominated by the first alternative.

structure. However, as the following lemma shows, such a structure would not be optimal ex ante.

Lemma 2. *If Γ is optimal, then $s(\omega) = \min(\omega, q)$ for almost all ω.*

Proof. For any admissible Γ, denote by A the set of states ω in which sales are not maximal, i.e.,

$$A = \{\omega : s(\omega) < \min(\omega, q)\}.$$

Suppose that A has positive measure. Ignoring some measure-theoretic fine points, consider an alternative contract $\hat{\Gamma}$ that on A increases $s(\omega)$ to $s'(\omega) = s(\omega) + \varepsilon(\omega) < \min(\omega, q)$, reduces $R(\omega)$ to $R'(\omega) = R(\omega) - \varepsilon(\omega) > 0$,[8] and increases $T_1(\omega)$ to $T_1'(\omega)$ in a way that keeps the retailer's utility unchanged:

$$\begin{aligned} T_1' &= (p - v_r)(s' - s) + v_r(R - R') + T_1 \\ &= p\varepsilon + T_1 \end{aligned}$$

Since the retailer's utility is unchanged, the new contract is incentive-compatible. By construction it satisfies (FR) and (L_1). Since the extra revenue from $\hat{\Gamma}$ goes entirely to the manufacturer, she is strictly better off:

$$T_1' + v_m R' = T_1 + p\varepsilon + v_m R' > T_1$$

for all $\omega \in A$. Therefore, the retailer sells more and repays more to the manufacturer at any $\omega \in A$ in $\hat{\Gamma}$, which implies that Γ is not optimal. □

Intuitively, the only reason that the retailer may want to undersell in some state ω is that the contract specifies a large return shipment of unsold inventory in order to relax the incentive constraint for the report of $\hat{\omega}$. However, this yields lower profits on the equilibrium path. Therefore, the manufacturer can be made better off by reducing the return shipment without violating the incentive-constraint, which is possible because if the retailer in any state has the ability and incentive to misreport ω in $\hat{\Gamma}$, he would have done so already in the original contract Γ by selling less.

As any optimal contract will have $s(\omega) = \min(\omega, q)$ for almost all ω, we can simplify the retailer's ex-post objective function to

$$p \min(\omega, q) - T_1(\hat{\omega}) - v_r R(\hat{\omega}). \tag{10}$$

[8] Note that by construction of A, $R(\omega) \geq q - s(\omega) > \max(0, q - \omega) \geq 0$.

3.2.3 Date-1 component

We then turn to the optimal date-1 component (T_1, R) for any given $q > \underline{q}$, using that necessarily $s = \min(\omega, q)$ by Lemma 2. To simplify notations, let V be a function indicating the retailer's total payout ex-post, i.e.,

$$V(\omega) = T_1(\omega) + v_r R(\omega). \tag{11}$$

It is instructive to note that any demand realization larger than q is useless to the retailer since q is determined ex-ante. Therefore, the total payout for $\omega > q$ should be identical to that of q. We highlight this result in Lemma 3.

Lemma 3. *If Γ satisfies (IC)-(IC-L) for any $\omega \leq q$, then it satisfies (IC)-(IC-L) for all ω if and only if $V(\omega) = V(q)$ for any $\omega > q$.*

Lemma 3 allows us to restrict attention to the contract structure on $[0, q]$. On $[0, q]$, the main difficulty comes from (IC-FR) and (IC-L). If we ignore both of them, the retailer is free to report any other types. (IC) itself implies that $V(\omega)$ is constant for all ω. If we only ignore (IC-FR), the right-hand side of (IC-L) increases with ω, so the retailer can always understate his type. (IC)-(IC-L) implies that $V(\omega)$ is decreasing in ω. However, when we consider (IC-FR) and (IC-L) together, such monotonicity of the retailer's total payout disappears.

Figure 2 is an example of how (FR) and (L_1) may invalidate the monotonicity of $V(\omega)$. In the example, $R(\omega) > q - \omega_2$ for any $\omega \in (\omega_1, \omega_2)$, while $T_1(\omega) > W - T_0 + p\omega_2$ for any $\omega \in (\omega_2, \omega_3)$. The retailer with type $\omega \in (\omega_1, \omega_2)$ cannot exaggerate his type above ω_2 due to the liquidity constraint for cash. The retailer with type $\omega \in (\omega_2, \omega_3)$ cannot understate his type below ω_2 due to the feasibility constraint for returns. One can hardly infer the shape of contract from the incentive constraint, since (IC)-(IC-L) only imposes restriction on (ω_1, ω_2) and (ω_2, ω_3) separately.[9]

As a result, we propose a new constructive method. On the whole interval (ω_1, ω_3), we reduce the return shipments and increase the cash repayments while keeping the expected total payout $\mathrm{E}_\omega[V(\omega)]$ unchanged. The resulting contract is depicted in Figure 3.

Parts (a) and (b) of Figure 3 are constructed in the same manner. The retailer is supposed to repay a constant amount of cash $W - T_0 + t$ to the manufacturer with minimum return shipments $q - \omega_3$. If his cash falls short, he has to increase return shipments as much as he could to fill in the gap. Therefore, if $t \leq p\omega_1$, the contract will look like part (a) of Figure

[9] Note that this argument does no depend on whether (FR) and (L_1) bind on (ω_1, ω_3). It also has nothing to do with the monotonicy of T_1 and R. What really matters is that T_1 and R are separated by ω_2.

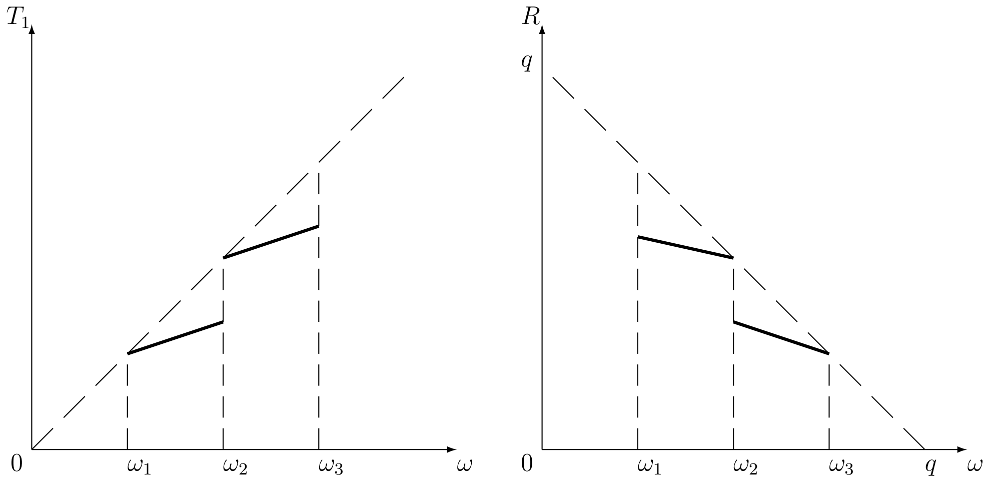


Figure 2: An example of how (IC)-(IC-L) imposes only local restrictions.

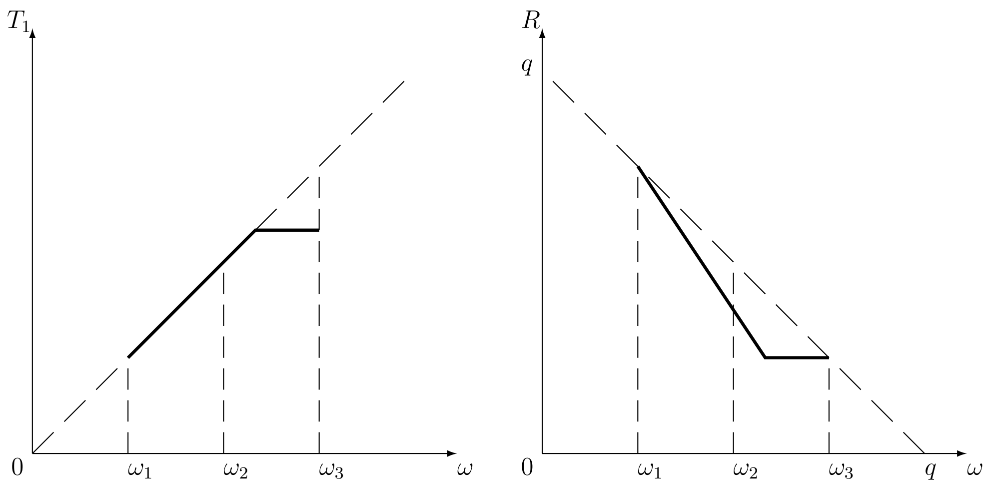
3. If $t > p\omega_1$, there are some sufficiently bad states in which the retailer cannot fulfill the obligation to repay t even if he transfers everything to the manufacturer. The contract will then looks like part (b) of Figure 3.

The improved contract has several features. First, it entails larger expected cash transfers to the manufacturer and smaller expected return shipments than the initial contract. Since returning unsold inventories is inefficient, this contract outperforms the contract in Figure 2. Second, both contracting parties' expected payoffs on (ω_1, ω_3) can be fully determined by three parameters: ω_1, ω_3, and a constant t. Hence, the problem of finding optimal contracts can be simplified into finding a sequence of parameters. Finally, the improved contract has a natural economic interpretation. The retailer is obligated to repay t in cash and return $q - \omega_3$ unsold inventories to the manufacturer when $\omega \in (\omega_1, \omega_3)$ is realized. If the cash repayment exceeds the retailer's total wealth, i.e., (L_1) binds, the manufacturer buys back some extra unsold inventories to fill in the gap until (FR) also binds.

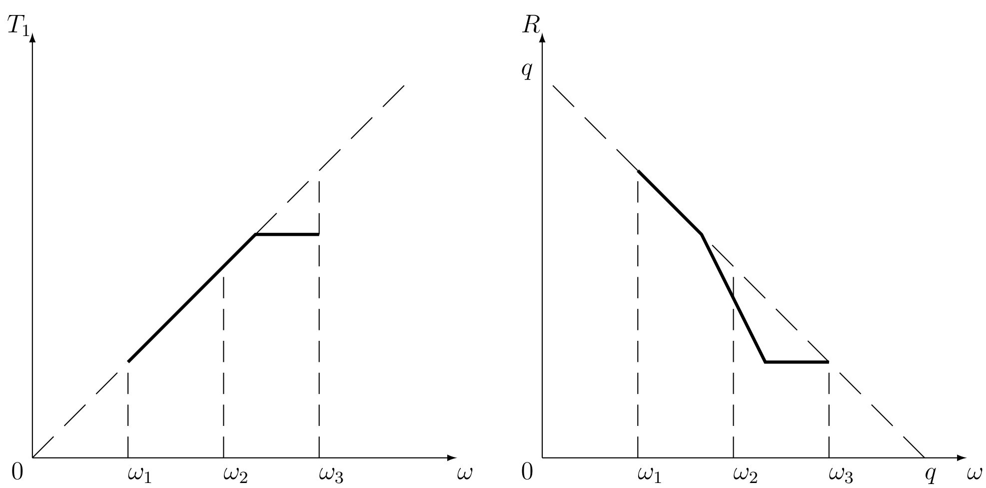
Because of these nice properties, we say Γ is a *local buyback contract* if one can divide $[0, q]$ into finitely many intervals like (ω_1, ω_3), and Γ has the features demonstrated in Figure 3 on each of these intervals. A formal definition is stated as below.

Definition 1. Γ *is a* ***local buyback contract*** *if there exists a sequence of pairs* $\{(\omega_i, t_i) : i = 0, 1, \ldots, n\}$*, such that:*

(a) $0 \leq \omega_0 < \omega_1 < \cdots < \omega_n \leq q$*;*



(a) $t \leq p\omega_1$



(b) $t > p\omega_1$

Figure 3: A local improvement of the contract in Figure 2

(b) for any $i \geq 1$ and $\omega \in (\omega_{i-1}, \omega_i)$,

$$T_1(\omega) = W - T_0 + \min(t_i, p\omega),$$
$$R(\omega) = \min\left(q - \omega_i + \frac{t_i - p\omega}{v_r}, q - \omega\right).$$

As we have discussed, in a local buyback contract, both contracting parties' payoffs conditional on (ω_{i-1}, ω_i) can be fully determined by the two endpoints as well as t_i. The retailer loses all his cash if $p\omega \leq t_i$, or equivalently,

$$\omega \leq \overline{\omega} = \frac{t_i}{p},$$

and loses all his unsold inventories if $q - \omega \leq q - \omega_i + (t_i - p\omega)/v_r$, or equivalently,

$$\omega \leq \underline{\omega} = \frac{t_i - v_r \omega_i}{p - v_r}.$$

Therefore, when $\underline{\omega} > \omega_i$, the retailer receives nothing on $(\omega_{i-1}, \underline{\omega})$, and pays out a constant amount $t + v_r(q - \omega_i)$ on $(\underline{\omega}, \omega_i)$:

$$\mathrm{E}_\omega[u_r | \omega_{i-1} < \omega < \omega_i] = \mathrm{E}_\omega[u_r | \underline{\omega} < \omega < \omega_i] = S(q | \underline{\omega} < \omega < \omega_i) + cq - t - v_r(q - \omega_i).$$

When $\underline{\omega} \leq \omega_i$ the retailer's total payout is always $t + v_r(q - \omega_i)$ on (ω_{i-1}, ω_i):

$$\mathrm{E}_\omega[u_r | \omega_{i-1} < \omega < \omega_i] = S(q | \omega_{i-1} < \omega < \omega_i) + cq - t - v_r(q - \omega_i).$$

Hence, by choosing t_i appropriately, one can transform an admissible contract into a local buyback contract without changing the retailer's expected payoff. This also suggests that the expected total payout $\mathrm{E}_\omega[V(\omega)]$ is the same after the transformation. Note that V is nondecreasing on every (ω_{i-1}, ω_i) in a local buyback contract. If we further know that V is nonincreasing in the original contract, then $T_1(\omega_i) \leq W - T_0 + t_i$, implying that the local buyback contract requests more cash repayment than the original contract. Consequently, switching from the original contract into a local buyback contract also benefits the manufacturer, i.e., moving from Figure 2 to Figure 3 is really an improvement.

How to divide $[0, q]$ into disjoint intervals so that V is nonincreasing on each of them? Our observation is that when $R(\omega) < q - \omega$, V is nonincreasing on $(\omega, q - R(\omega))$; this can be proved immediately by applying $s(\omega) = \hat{s} = \omega$ to (IC)-(IC-L). Let

$$\Omega = \bigcup_{\omega \in [0,q]} (\omega, q - R(\omega)),$$

which represents the union of all the open intervals $(\omega, q - R(\omega))$ generated from ω. By definition, Ω is an open set, so it is the union of countably many disjoint open intervals (Royden and Fitzpatrick, 2010). If Ω can be partitioned into finitely many intervals, then V must be nonincreasing on any of these intervals. Between these intervals are states with a binding (FR). Therefore, performing the transformation from Figure 2 to Figure 3 on every interval will finally give us a local buyback contract.

If the decomposition of Ω consists of infinitely many intervals, and they may not even be well-ordered, then we approximate the original contract with a local buyback contract. In particular, for any $\varepsilon > 0$, we restrict attention to intervals with measure larger than ε. Summarizing the two cases gives us Lemma 4. The proof of Lemma 4, due to the possible complex structure of Ω, is involving and relegated to Appendix.

Lemma 4. *If Γ implements $q > \underline{q}$, then it is either dominated or approximated by a local buyback contract.*

Proof. See Appendix A.1. □

The key implication of Lemma 4 is that, if we are able to find an optimal local buyback contract, then it is indeed optimal. Our next step is to find the optimal local buyback contract and see whether it is admissible. Note that, given a local buyback contract, both contracting parties' expected payoffs are entirely pinned down by $\{(\omega_i, t_i) : i = 0, 1, \dots, n\}$. Thus the standard technique for constrained optimization problems can be applied.

Let L be the Lagrangian of the manufacturer's optimization problem, and λ be the Lagrangian multiplier of (PC_r). If Γ maximizes $\text{E}_\omega[u_m(\omega)]$ subject to (PC_r) and (PC_m), $\{(\omega_i, t_i) : i = 0, 1, \dots, n\}$ should maximize

$$L = \text{E}_\omega[u_m(\omega)] + \lambda\{\text{E}_\omega[u_r(\omega, \omega, \min(\omega, q))] - (W + \underline{u})\},$$

subject to (PC_r) and (PC_m), as well as several boundary constraints:

$$0 = \omega_0 \leq \omega_1 \leq \cdots \leq \omega_n = q, \tag{12}$$

$$t_i \leq p\omega_i + v_r(q - \omega_i) \text{ for any } i = 1, 2, \dots, n, \tag{13}$$

and the complementary slackness constraint:

$$\lambda \geq 0, \quad \lambda\{\mathrm{E}_\omega[u_r(\omega, \omega, \min(\omega, q))] - (W + \underline{u})\} = 0. \tag{14}$$

We observe from first-order necessary conditions that $\lambda > 0$, which means (PC_r) binds in the optimal contract. We can also show $\partial L/\partial\omega_i < 0$ and $\partial L/\partial t_i > 0$ for all $i \geq 1$, which suggest that $t_i \leq p\omega_i + v_r(q - \omega_i)$ binds. Hence, the optimal date-1 component is a local buyback contract with $n = 1$ and a binding (PC_r). We call it a *buyback contract* and make a formal definition below.

Definition 2. *Γ is a* ***buyback contract*** *if there exists $t > 0$ such that for any ω:*

$$T_1(\omega) = \begin{cases} W - T_0 + p\omega & \omega < \overline{\omega}, \\ W - T_0 + t & \omega \geq \overline{\omega}; \end{cases}$$

$$R(\omega) = \begin{cases} q - \omega & \omega < \underline{\omega}, \\ \frac{1}{v_r}(t - p\omega) & \underline{\omega} \leq \omega < \overline{\omega}, \\ 0 & \omega \geq \overline{\omega}; \end{cases}$$

where $\overline{\omega} = t/p$, $\underline{\omega} = 0$ if $t \leq v_r q$, and $\underline{\omega} = \frac{t - v_r q}{p - v_r}$ if $t > v_r q$.

By Definition 2, in a buyback contract the retailer is obligated to repay t to the manufacturer by cash. The $W - T_0$ part is the cash inherited from date 0, and t shall be paid by the retailer's date-1 revenue $p\omega$. If $p\omega < t$, the retailer repays all his cash. The definition of $R(\omega)$ indicates that the retailer has minimum return shipments $q - b$. Moreover, if his cash repayment $T_1(\omega)$ is not enough to cover t, some extra amount must be returned to make sure that the retail feels indifferent, unless $R(\omega)$ reaches its upper bound $q - \omega$. Lemma 5 confirms our discussion on the optimality of buyback contracts.

Lemma 5. *If Γ implements $q > \underline{q}$, then it is dominated by a buyback contract with t determined by a binding (PC_r).*

Proof. See Appendix A.2. □

As a final remark, we show that the buyback contract characterized in Lemma 5 is indeed admissible.

Lemma 6. *If Γ is a buyback contract satisfying (L_0), (PC_r) and (PC_m), then it is admissible.*

Proof. It is straightforward to verify (FS), (FR), and (L_1) from Defintion 2. For (IC)-(IC-L), note that upward misreporting is infeasible in a buyback contract due to the cutoff structure of T_1. In terms of downward misreporting, the retailer at state ω has no incentive to misreport any $\hat{\omega} > \underline{\omega}$, because $V(\omega) = V(\hat{\omega})$. He also has no incentive to misreport any $\hat{\omega} \leq \underline{\omega}$, because that requires selling less to fulfill the return obligation with full cash repayments, which gives the retailer zero utility. Therefore, Γ also satisfies (IC)-(IC-L). □

Lemmas 5 and 6 jointly confirm that a buyback contract with a binding (PC_r) is both feasible and optimal to implement $q > \underline{q}$. We summarize them in Proposition 2.

Proposition 2. *The optimal contract implementing $q > \underline{q}$ is a buyback contract.*

By Definition 2, a buyback contract may exhibit two different structures, depending on the relationship between t and $v_r q$.

Case 1: If $t \leq v_r q$, then $\underline{\omega} = 0$, the ex-post cash transfer T_1 and return R each has at most one kink at $\overline{\omega}$. One can understand the contract as an upfront payment t with a constant buyback price v_r, because $V(\omega) = t$ for all ω. We would refer to this case as a *linear buyback price.*

The manufacturer's expected utility is

$$\begin{aligned}\mathrm{E}_\omega\, u_m(\omega) &= W - cq + \int_0^{\overline{\omega}} [p\omega + \frac{v_m}{v_r}(t - p\omega)]dF(\omega) + \int_{\overline{\omega}}^{+\infty} t dF(\omega) \\ &= W - cq + t - \left(1 - \frac{v_m}{v_r}\right)[t - pQ(\overline{\omega})],\end{aligned}$$

where t is determined by a binding (PC_r),

$$\mathrm{E}_\omega\, u_r(\omega, \omega, \min(\omega, q)) = S(q) + cq - t = W + \underline{u}. \tag{15}$$

Hence,

$$\mathrm{E}_\omega\, u_m(\omega) = S(q) - \underline{u} - \left(1 - \frac{v_m}{v_r}\right)[t - pQ(\overline{\omega})]. \tag{16}$$

Here, $(1 - v_m/v_r)[t - pQ(\overline{\omega})]$ is the efficiency loss. If the retailer is obligated to make an upfront payment t at date-1, he has to order at least $\overline{\omega} = t/p$ units, which generates an expected revenue $pQ(\overline{\omega})$. Thus, $t - pQ(\overline{\omega})$ is the expected gap that has to be fulfilled by returns. Since one unit of unsold inventory generates a value v_r to the retailer and v_m to the manufacturer, $1 - v_m/v_r$ is the efficiency loss induced by one unit of shortfall in revenue.

Case 2: If $t > v_r q$, then $\underline{\omega} > 0$, T_1 and R each has two kinks at $\underline{\omega}$ and $\overline{\omega}$. The buyback price in this case is never constant when $\omega < \underline{\omega}$, because in this range:

$$\frac{t - T_1(\omega)}{R(\omega)} = \frac{t - p\omega}{q - \omega}.$$

Therefore, we say in this case the buyback contract has a *nonlinear buyback price.*

The manufacturer's expected utility is

$$\begin{aligned} \mathrm{E}_\omega\, u_m(\omega) = W - cq &+ \int_0^{\underline{\omega}} [p\omega + v_m(q-\omega)] dF(\omega) \\ &+ \int_{\underline{\omega}}^{\overline{\omega}} [p\omega + \frac{v_m}{v_r}(t - p\omega)] dF(\omega) + \int_{\overline{\omega}}^{+\infty} t dF(\omega), \end{aligned}$$

where t is determined by a binding (PC_r),

$$\mathrm{E}_\omega\, u_r(\omega, \omega, \min(\omega, q)) = [S(q|\omega > \underline{\omega}) + cq - t][1 - F(\underline{\omega})] = W + \underline{u}. \tag{17}$$

Hence,

$$\mathrm{E}_\omega\, u_m(\omega) = S(q) - \underline{u} - \left(1 - \frac{v_m}{v_r}\right) [v_r q + (p - v_r) Q(\underline{\omega}) - pQ(\overline{\omega})]. \tag{18}$$

Comparing (16) and (18), we see that the efficiency loss is now related to $\underline{\omega}$. That is because when $\omega \le \underline{\omega}$, the retailer returns all the unsold inventories but still cannot fulfill the gap between t and his revenue $p\omega$ if the buyback price is v_r. Therefore, the manufacturer has to increase the per unit buyback price. Put differently, the retailer returns less than that is expected in a linear-pricing buyback contract.

The cutoff between Case 1 and Case 2 can be derived from taking $t \to v_r q$ on the right-hand side of (15) (or equivalently the right-hand side of (17)). That is, let $\overline{q}$ be the solution for

$$S(q) + (c - v_r)q = W + \underline{u}. \tag{19}$$

One can also verify that the left-hand side of (19) is strictly increasing in q. Then, $t \le v_r q$ if and only if $q \le \overline{q}$, and $t > v_r q$ if and only if $q > \overline{q}$. It is immediate from (6) and (19) that $\underline{q} < \overline{q}$.

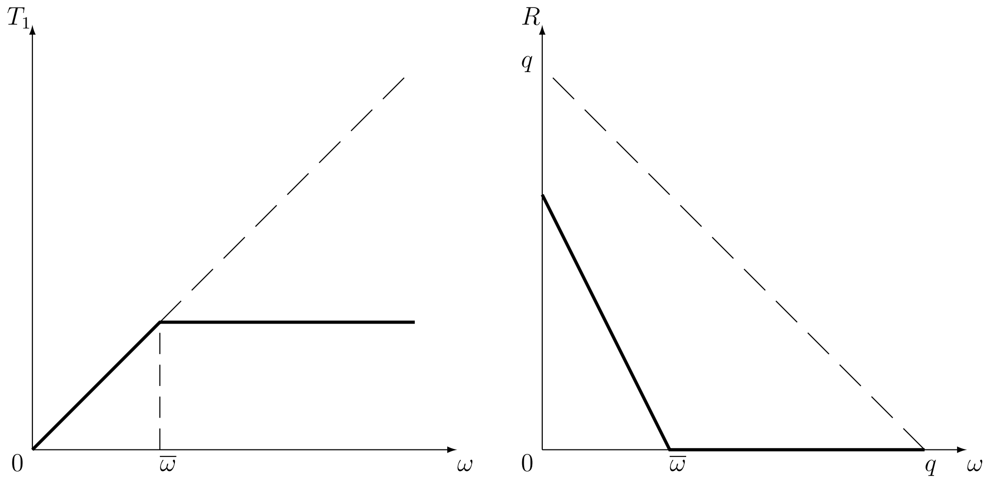
The preceding discussion is summarized in Proposition (3).

Proposition 3. *The optimal contract implementing $q > \underline{q}$*

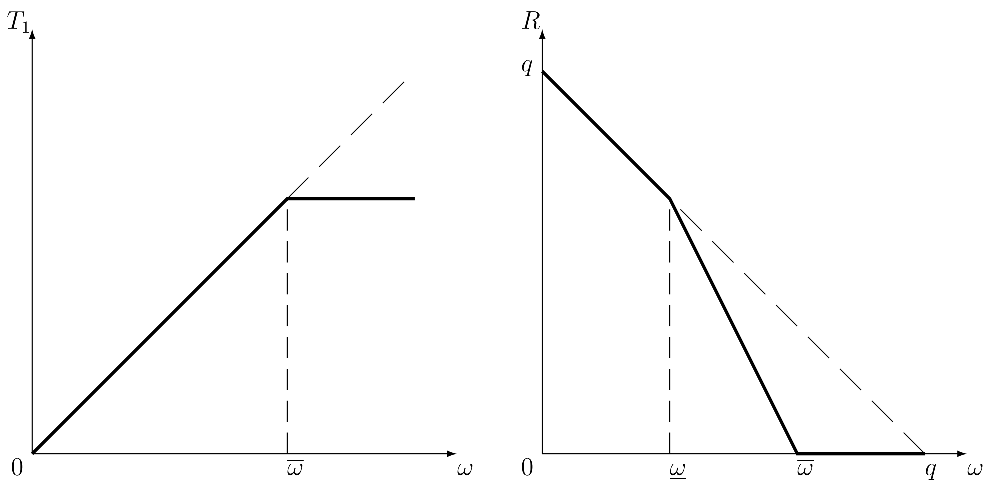
(a) has a linear buyback price when $\underline{q} < q \le \overline{q}$;

(b) has a nonlinear buyback price when $q > \overline{q}$.

Figure 4 graphically illustrates the two cases of Proposition 3.



(a) Linear buyback price ($q \leq \overline{q}$)



(b) Nonlinear buyback price ($q > \overline{q}$)

Figure 4: Buyback contracts on $[0, +\infty)$

3.2.4 Date-0 component

The date-0 component consists of q and T_0. It is straightforward to see that the choice of T_0 is irrelevant, since the change of T_0 is always shifted to the change of $W - T_0$ at date 1.

Therefore, our final step is to investigate whether the manufacturer benefits from implementing a quantity higher than $\underline{q}$. Note that if the manufacturer offers a wholesale contract implementing $\underline{q}$, her payoff is simply $W - c\underline{q}$. If she offers a buyback contract that implements $q \in (\underline{q}, \overline{q}]$, her payoff is determined by (16). The change of $\mathrm{E}_\omega\, u_m(\omega)$ with respect to q is captured by

$$\frac{\mathrm{d}\,\mathrm{E}_\omega\, u_m(\omega)}{\mathrm{d}q} = S'(q) - \left(1 - \frac{v_m}{v_r}\right)\left[\frac{\mathrm{d}t}{\mathrm{d}q} - pQ'(\overline{\omega})\frac{\mathrm{d}\overline{\omega}}{\mathrm{d}q}\right] \tag{20}$$

$$= \left[1 - \left(1 - \frac{v_m}{v_r}\right)F(\overline{\omega})\right][p - (p - v_r)F(q)] - c. \tag{21}$$

In the second equality, we use the facts:

$$\frac{\mathrm{d}\overline{\omega}}{\mathrm{d}q} = \frac{1}{p}\left(\frac{\mathrm{d}t}{\mathrm{d}q}\right),$$
$$\frac{\mathrm{d}t}{\mathrm{d}q} = S'(q) + c.$$

When $q \to \underline{q}$ from the right, we have $\overline{\omega} \to 0$, $t \to 0$, and, more importantly,

$$\lim_{q \to \underline{q}+} \mathrm{E}_\omega\, u_m(\omega) = S(\underline{q}) - \underline{u} = W - c\underline{q},$$
$$\lim_{q \to \underline{q}+} \frac{\mathrm{d}\,\mathrm{E}_\omega\, u_m(\omega)}{\mathrm{d}q} = p - (p - v_r)F(\underline{q}) - c > 0.$$

The last inequality follows from (4) and $q^{FB} > \underline{q}$. In other words, there must be some $q > \underline{q}$ that gives the manufacturer a strictly higher payoff than the wholesale contract.

We denote by q^* the optimal quantity that the manufacturer wants to implement using a buyback contract. Then, q^* is either a root of (21), or a root of the first-order condition of the manufacturer's payoff when $q > \overline{q}$. That is,

$$\frac{\mathrm{d}\,\mathrm{E}_\omega\, u_m(\omega)}{\mathrm{d}q} = S'(q) - (1 - \frac{v_m}{v_r})\left[v_r + (p - v_r)Q'(\underline{\omega})\frac{\mathrm{d}\underline{\omega}}{\mathrm{d}q} - pQ'(\overline{\omega})\frac{\mathrm{d}\overline{\omega}}{\mathrm{d}q}\right] \tag{22}$$

$$= \left[1 - \left(1 - \frac{v_m}{v_r}\right)F(\overline{\omega}) - \frac{v_m}{v_r}F(\underline{\omega})\right]\left[p - (p - v_r)\frac{F(q) - F(\underline{\omega})}{1 - F(\underline{\omega})}\right] + v_m F(\underline{\omega}) - c. \tag{23}$$

In the second equality, we use the facts:

$$\frac{\mathrm{d}\underline{\omega}}{\mathrm{d}q} = \frac{1}{p - v_r}\left(\frac{\mathrm{d}t}{\mathrm{d}q} - v_r\right),$$
$$\frac{\mathrm{d}t}{\mathrm{d}q} = p - (p - v_r)\frac{F(q) - F(\underline{\omega})}{1 - F(\underline{\omega})}.$$

From either (21) or (23), we have

$$\frac{\mathrm{d}\,\mathrm{E}_{\omega}\,u_m(\omega)}{\mathrm{d}q} = 0 \implies p - (p - v_r)F(q^*) > c,$$

which implies $q^* < q^{FB}$ from (4).

In a nutshell, when (6) holds, optimal contracts are wholesale contracts with first-best quantity. When (6) fails, optimal contracts are buyback contracts. There must be inefficient return of unsold inventories at date 1, which incentivizes the manufacturer to reduce the probability of oversupply. Consequently, the second-best quantity is smaller than the first-best. The date-0 cash repayments T_0 does not enter the manufacturer's objective function, thus it is irrelevant to the optimality of contracts. Proposition 4 formally states our results.

Proposition 4. *(a) When $q^{FB} \leq \underline{q}$, the optimal contract is a wholesale contract implementing q^{FB}.*

(b) When $q^{FB} > \underline{q}$, the optimal contract is a buyback contract implementing $q^* \in (\underline{q}, q^{FB})$; moreover, if $q^* > \overline{q}$, the buyback price is nonlinear.

One way to understand Proposition 4 is to interpret $\underline{u}$ as the retailer's bargaining power in the vertical relationship. When the retailer has to obtain a large fraction of the gains from trade, the optimal contract is a wholesale contract with a fixed date-1 cash transfer and zero return shipment. As the retailer's bargaining power decreases, the optimal contract shifts from wholesale to buyback and the retailer's date-1 obligation increases. When the retailer's reservation value is sufficiently low, the optimal contract becomes a buyback contract with nonlinear pricing. The manufacturer has to increase the buyback price in low-demand states so as to extract more revenue from the retailer in high-demand states.

The relationship between bargaining power and contract structures can also be observed in practice. Large retailers such as Walmart or Target are less likely to delay payments to suppliers as they face weaker financial constraints, while small groceries or bookstores may specify buyback terms in their contracts with producers. Thus, our analysis provides a foundation of retail contracts. In the supply chain contracting literature, a pre-dominant paradigm is to compare amongst various contracts observed in practice (e.g., Cachon, 2003; Chen, 2003). While these comparisons generate useful managerial implications, a potential caveat is that the contracts considered may be sub-optimal. By taking a different approach, our analysis speaks directly to the question of contract optimality. Remarkably, even though salvaging unsold inventories at the retailer is more efficient, the manufacturer buys back some of them in order to alleviate the ex-post adverse selection problem.

4 Discussion

4.1 Limited liability

The first key friction of our model is that the retailer is subject to limited liability in both dates. If limited liability is absent at date 0, the manufacturer will simply charge a fixed cash transfer from the retailer and implement the first-best price and quantity. This can be viewed as an extreme case of part (a) of Proposition 4 where W is sufficiently large. In this case, the optimal contract is a wholesale contract with first-best quantities, so relaxing (L_0) is weakly beneficial to the manufacturer.

However, if limited liability is relaxed at date 1, the effect may not be straightforward. For instance, one can argue that salvaging unsold inventories generates cash flow instead of nontransferable utility to the retailer at date 1. Then, v_r may be interpreted as a fire sale price lower than the retailer's marginal cost. In this case, there is no return of unsold inventories, so the retailer always chooses $s = \min(\omega, q)$. (L_1) becomes

$$T_1(\omega) \leq W - T_0 + p \min(\omega, q) + v_r(q - \min(\omega, q)), \qquad (\mathrm{LL}_1')$$

and (IC)-(IC-L) becomes

$$T_1(\omega) \leq T_1(\hat{\omega}), \text{ for any } \omega, \hat{\omega} \text{ such that } T_1(\hat{\omega}) \leq ps + v_r(q - \min(\omega, q)). \qquad (\mathrm{SIC}')$$

Clearly, (SIC′) suggests that $T_1(\omega)$ is constant for all ω, and by (LL_1'),

$$T_1(\omega) \leq W - T_0 + v_r q. \qquad (24)$$

In other words, now the manufacturer cannot "punish" the retailer by requesting inefficient returns, so the date-1 cash repayment does not depend on the retailer's report $\hat{\omega}$. T_1 is therefore bounded above by the retailer's cash flow at the lowest state. (24) further implies that the manufacturer's expected utility satisfies

$$\mathrm{E}_\omega\, u_m(\omega) = -cq + T_0 + \mathrm{E}_\omega\, T_1(\omega) \leq W - cq + v_r q. \qquad (25)$$

In the benchmark model, by Proposition 4 and Definition 1, the optimal buyback contract can be understood as the retailer tries to make an upfront payment $W - T_0 + t$ at date 1. Therefore, the manufacturer's expected utility satisfies

$$\mathrm{E}_\omega\, u_m(\omega) \leq -cq + T_0 + W - T_0 + t = W - cq + t. \qquad (26)$$

Comparing (25) and (26), one may guess that the manufacturer is worse off in this revised model when the optimal contract has $t > v_r q$. This conjecture is confirmed in Proposition 5.

Proposition 5. *There exists a cutoff $\hat{u} < (p - v_r)Q(q^{FB})$ such that when $W + \underline{u} < \hat{u}$, the manufacturer is worse off when the retailer can salvage cash from unsold inventories. As a result, the optimal contract is less efficient compared with the benchmark model.*

4.2 Incentive-compatibility

Another important feature of our model is that the retailer's incentive-compatibility constraint incorporates his limited liability and feasibility constraints.[10] Such model specification actually implies that (L_1) and (FR) must hold off the equilibrium path. Intuitively, when the retailer makes his report $\hat{\omega}$ to the manufacturer, he has already finished selling the products and collected $p\min(\omega, q)$ units of cash with $q - \min(\omega, q)$ units of unsold inventory. If the contract indicates a cash transfer higher than $p\min(\omega, q)$ or a return shipment larger than $q - \min(\omega, q)$ at some state $\hat{\omega}$, the retailer is unable to report $\hat{\omega}$ even if he find it profitable to do so. Incorporating (L_1) and (FR) into (IC) has two major effects on the model.

First, our (IC) is weaker than what is standard in the literature of adverse selection. To see this, note that if we allow the retailer to misreport any other types as in the classical screening model, then the incentive-compatibility constraint should be stated as

$$u_r(\omega, \omega, s) \geq u_r(\omega, \hat{\omega}, \hat{s}) \text{ for any } \omega, \hat{\omega}, s \text{ and } \hat{s} \text{ such that } s, \hat{s} \in [0, \min(\omega, q)]. \tag{IC^U}$$

Here the superscript U stands for "unconstrained incentive-compatiblity". Clearly, (IC^U) imposes no restriction on the retailer's set of possible deviation, which suggests that a type-ω retailer can mimic any other types. If a contract specifies $u_r(\omega, \omega, s) < u_r(\omega, \hat{\omega}, \hat{s})$ for some $\omega, \hat{\omega}, s$ and $\hat{s}$, then it should violate (IC^U). However, such contract may still satisfy our (IC) as long as $T(\hat{\omega}) > W - T_0 + p\min(\omega, q)$ or $R(\hat{\omega}) > q - \min(\omega, q)$.

In fact, since in our model both contracting parties are risk-neutral, (IC^U) immediately implies that $V(\omega)$ is constant for all ω, or (IC^U) binds for all ω. However, according to Proposition 4, $V(\omega)$ is increasing when the optimal contract is a buyback with $t > v_r q$. That is, if one use (IC^U) instead of (IC), the resulting optimal contracts will be sub-optimal when

[10] This model specification is in line with the financial contracting literature (e.g., Townsend, 1979; Gale and Hellwig, 1985), where the debtor naturally faces a liquidity constraint. However, as pointed out by (Gui et al., 2019), a number of papers overlooked the privately informed party's limited liability when specifying their incentive constraints, which may lead to sub-optimal contracts in some simple examples.

$\underline{u} < u_0$.

Second, our (IC) gives rise to a novel constructive proof technique. When the type set is a continuum, the standard approach for contracting problems with incentive constraints is optimal control (e.g., Hellwig, 2010). The basic idea is to replace the global incentive constraint with local first-order conditions. When the agent is able to mimic any other type irrespective of his own type or the contract, his indirect utility function is absolute continuous, which serves as the state variable in the control problem.

Nevertheless, the control-theoretic approach cannot be applied in the present paper when (IC) incorporates (FR) and (L_1). Since the retailer's set of possible deviation is type- and contract-dependent, it is possible that the retailer is only able to mimic a subset of types, or even cannot mimic other type. For instance, if the incentive constraint is specified as (IC^U), then one can use local incentive constraints to replace (IC^U); u_r is thus absolutely continuous.[11] For the (IC) presented in Section 2, if (L_1) binds at ω, the retailer with any $\hat{\omega} < \omega$ cannot misreport ω because he cannot afford the cash payment specified in the contract when ω is reported. Hence, the retailer's indirect utility function may have a jump at ω. The possible discontinuities in contracts prevent us from using control theory. Consequently, we apply a step-by-step constructive method to characterize the optimal contract, as shown in the discussion before Lemma 5. This proof technique has been applied in the early literature of Costly State Verification (e.g., Gale and Hellwig, 1985), but is much more involving in this paper.

5 Extensions

5.1 Price-dependent demand

In our benchmark model we have assumed that the retail price p was exogenous. This is probably a good assumption if the manufacturer is sufficiently remote and unacquainted with the retailer's local market, and if that market is sufficiently competitive. Alternatively, the retail price could be contractible and therefore endogenous to the contracting problem. In this section, we therefore relax our restriction and allow for a price-dependent stochastic demand function $F(\cdot; p)$. This parameterized-distribution-function approach provides greater flexibility than standard state-space models and encompasses different specific state-space formulations. As an example, consider the basic demand model $Q = Q(p, \theta)$ where $Q \geq 0$

[11] Actually, assuming that (IC) holds for any ω and $\hat{\omega}$ that lies in an open ball of ω is sufficient for the substitution to be valid. See Hellwig (2000) and Hellwig (2001) for examples. In these papers, the limited liability constraint never binds due to the agent's risk-aversion, so the control-theoretic approach can still be applied. However, in our model even this weaker condition cannot be ensured.

is market demand and $\theta \in \Theta$ a random variable with probability measure μ. For any p and $0 \leq \omega_1 < \omega_2$, we have $0 \leq \mu(\{\theta; Q(p;\theta) \leq \omega_1\} \leq \mu(\{\theta; Q(p;\theta) \leq \omega_2\} \leq 1$. Hence, $F(\omega;p) = \mu(\{\theta; Q(p;\theta) \leq \omega\}$ is a well-defined family of c.d.f.s. Other examples can easily be constructed. That pricing is determined before the demand realization is consistent with the long-standing literature on price-setting newsvendor problems (Petruzzi and Dada (1999)).

We assume that $F(\cdot;\cdot)$ is atomless and differentiable in both ω and p, and let

$$F_p(\omega;p) = \frac{\partial F(\omega;p)}{\partial p},\ f(\omega;p) = \frac{\partial F(\omega;p)}{\partial \omega} > 0.$$

By definition, $F_p(\cdot;p)$ is the marginal effect of price on the distribution of demand, $f(\cdot;p)$ is the density function of F given any price p. In line with traditional models, such as the state-space model sketched above, we assume that $F(\cdot;p)$ satisfies First-Order Stochastic Dominance, i.e., for any p and ω, $F_p(\omega;p) > 0$. This assumption ensures that retail demand is more likely to be realized at a higher level when the price is higher.

The expected feasible demand and social surplus are now given by

$$Q(q;p) = q - \int_0^q F(\omega;p)d\omega,$$
$$S(q;p) = (p - v_r)Q(q;p) - (c - v_r)q,$$

respectively. To avoid excessive technical details, we assume that the social surplus is concave in p and q, and a higher price leads to higher marginal social benefits of quantity. That is, for any p and q,

$$S_{pp}(q;p) < 0,\ S_{qq}(q;p) < 0,\ \text{and } S_{pq}(q;p) > 0. \tag{27}$$

Here, subscripts p and q are used to denote partial derivatives. Assume that p is observable and contractible, then the definition of retail contract should also be extended to $\Gamma = (p, q, T_0, s, T_1, R)$.

When information is symmetric, the first-best price is determined by the first-order conditions of (3):

$$F(q;p) = \frac{p-c}{p-v_r}, \tag{28}$$
$$Q(q;p) = -(p - v_r)Q_p(q;p). \tag{29}$$

(27) ensures that these conditions are sufficient. Let (p^{FB}, q^{FB}) be the solution for (28) and (29). Then, the first-best contract generates a gross surplus $S(q^{FB};p^{FB})$. With slightly

abuse of notations, let $\underline{q}(p)$ be the solution of

$$S(q^{FB}; p) + cq^{FB} = W + \underline{u}.$$

Then, by an argument similar to Lemma 1, q^{FB} can be implemented by a wholesale contract if and only if $q^{FB} \leq \underline{q}(p^{FB})$.

Moreover,

$$\begin{aligned} S_p(q;p) &= Q(q;p) - (p - v_r)\int_0^q F_p(\omega;p)d\omega, \\ Q_p(q;p) &= -\int_0^q F_p(\omega;p)d\omega. \end{aligned}$$

Since $F_p(\omega; p) > 0$, both $Q(q; p)$ and $S_p(q; p)$ decrease with p. Therefore, $S(q; p)$ is concave in p, which implies the existence of $\underline{p}$ and $\overline{p}$ such that

$$S(q^{FB}; p^{FB}) + cq^{FB} \geq W + \underline{u}$$

if and only if $p^{FB} \in [\underline{p}, \overline{p}]$.

When there is asymmetric information, similar to our analysis in Section 3, p belongs to the date-0 component. Therefore, it does not affect the buyback structure of the optimal contract. We can solve for p and q together from first-order conditions. In this case, (16) becomes

$$\mathrm{E}_\omega\, u_m(\omega) = S(q;p) - (1 - \frac{v_m}{v_r})[t - pQ(\overline{\omega};p)]. \tag{30}$$

(18) becomes

$$\mathrm{E}_\omega\, u_m(\omega) = S(q;p) - (1 - \frac{v_m}{v_r})[v_r q + (p - v_r)Q(\underline{\omega};p) - pQ(\overline{\omega};p)]. \tag{31}$$

The optimal price and quantity is characterized in Proposition 6.

Proposition 6. *If Γ is optimal under endogenous and contractible retail price, then*

(a) when $q^{FB} \leq \underline{q}(p^{FB})$, Γ is a wholesale contract implementing q^{FB} at a price p^{FB};

(b) when $q^{FB} > \underline{q}(p^{FB})$, Γ is a wholesale contract implementing $q^* < q^{FB}$ at a price $p^* < p^{FB}$.

Proof. See Appendix A.4. □

Intuitively, efficiency loss in a buyback contract comes from return shipments, so the manufacturer is more reluctant to "excess supply" rather than "excess demand". Consequently, she will deliver less products ex-ante and request a lower retail price to reduce the probability of oversupply. This logic also applies to Proposition 4.

5.2 Reordering

In our benchmark model, production precedes sales, so the order quantity should be determined before demand is observed. This assumption fits into many production-in-advance industries (see, e.g., Montez and Schutz, 2021), but in some situations the retailer is able to reorder additional products when he observes a high demand. In this section, we revise our benchmark model so the retailer can reorder extra units contingent on his private information, and study whether it improves contract efficiency.

Consider the environment described in Section 2. Now assume that the game has three periods. At date 0, an initial contract $\Gamma_0 = (p, q_0, T_0, T_1, R)$ is signed, and q_0 is delivered to the retailer. At date 1, the retailer observes the retail demand ω and makes a report $\hat{\omega}$. T_1 and R are transferred accordingly. Then, the manufacturer offers a follow-up contract $\Gamma_1 = (q_1, T_2)$ to the retailer, which specifies an additional quantity q_1 to be delivered and a cash transfer T_2. Due to the lag between production and sales, the extra quantity q_1 is sold at date 2, and T_2 is made after the retailer collects his revenue at date 2. Naturally, Γ_1 is contingent on the retailer's report $\hat{\omega}$. A timeline for this revised model is shown in Figure 5.

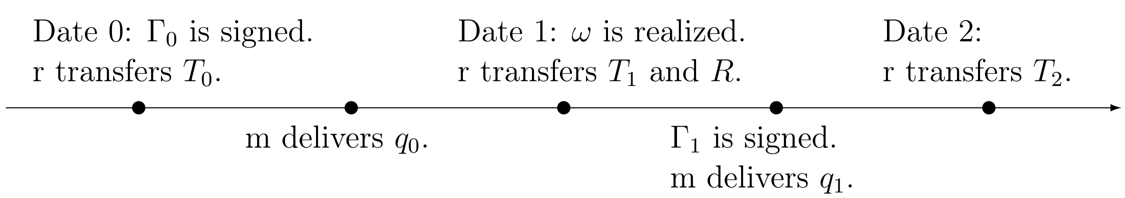


Figure 5: Timeline of the model with reordering.

One central question of this model is to determine the retail demand at date 2. In fact, some literature argues that the retail demand will decay or flow to other products if not being supplied in the first place (Netessine and Rudi, 2003). Therefore, we assume that only a fraction $\beta \in (0, 1)$ of the excess demand at date 1 can be preserved at date 2. That is, the residual demand at date-1 is $\max\{0, \beta(\omega - q_0)\}$. If $\beta = 0$, our model goes back to the benchmark described in Section 2 (which is referred to as the lost-sales model).

If $\beta = 1$, the excess demand at date 1 can be fully captured by the retailer at date 2 as the (fully) backlogging model. The optimal contract should have nothing delivered at date

0 and all demand is fulfilled at date 2. Our model thus boils down to a standard screening problem in which the retailer has private information before signing Γ_1. In this environment, the manufacturer simply offers $T_2(q_1) = pq_1 - \underline{u}$ at date 1. The retailer reports ω truthfully and obtains $q_1 = \omega$. The two contracting parties act as if they are vertically integrated into a monopoly, and choose the monopoly price p^m according to

$$p^m \in \arg\max \ p \int_0^{+\infty} \omega dF(\omega; p).$$

The case becomes interesting if $\beta \in (0, 1)$. Since the manufacturer cannot commit on Γ_1 at date 0, by sequential rationality, she will optimally choose $T_2(q_1) = pq_1$ at date 1. The retailer thus enjoys no surplus from reordering. As a result, the date-0 contract Γ_0 should still take the form of a buyback contract. The date-0 order quantity q_0 will be smaller than q^* characterized in Proposition 4, as the excess demand can be partially recovered now. Consequently, the retail price p will be higher than p^*. These results are formally stated in Proposition 7.

Proposition 7. *If $\beta \in (0, 1)$, and Γ_0 and Γ_1 are optimal in the model of reordering, then:*

(a) Γ_0 is a buyback contract with $p < p^*$ and $q_0 < q^*$. Moreover,

(a.1) when $\beta \to 0$, $p \to p^*$ and $q_0 \to q^*$;

(a.2) when $\beta \to 1$, $p \to p^m$ and $q_0 \to 0$.

(b) Γ_1 satisfies

$$q_1 = \max\{0, \beta(\omega - q_0)\},$$
$$T_2 = pq_1.$$

Proof. See Appendix A.5. □

Intuitively, β is an index of the "position" that this revised model stands between two extreme cases. Although Proposition 7 cannot guarantee the monotonicity of p and q_0 over β, one may still conclude that when β is sufficiently large, our model becomes "close" to a screening model, thus the manufacturer has a stronger incentive to produce less at date 0. Similarly, when β is sufficiently small, our model becomes "close" to the benchmark where information is symmetric at the time of contracting. The manufacturer in fear of losing demand at date 2 will produce more at date 0.

Proposition 7 enables us to make testable predictions based on different real-life interpretations of β. One may translate β as a parameter indicating "how fast will unfulfilled demand vanishes". Then Proposition 7 tells us that retailers selling durable products may enjoy stronger market power than retailers selling newspapers or fast foods, because consumers easily turn to other sellers if they cannot get a newspaper immediately. As a result, newsvendors must order a large quantity ex-ante and cannot rely too much on reordering. An alternative interpretation of β may be "the difficulty of observing realized demand". Machine factories often start to produce after receiving orders from customers, while supermarkets and bakeries can hardly tell the actual demand of the day unless the last customer leaves their stores. Therefore, supermarkets may have to maintain a large storage in advance.

5.3 Multiple retailers

It is common in practice that a manufacturer sells her products through different retailers. The manufacturer may want to maintain a relatively high retail price for her products, but retailers usually compete with each other and attract customers by cutting down retail prices. As a result, the manufacturer sometimes fixes the retail price through contracts. This mechanism is the so called Resale Price Maintenance (RPM) that has been well studied in the literature (e.g., Marvel and McCafferty, 1984; Shaffer, 1991; Deneckere et al., 1996; Jullien and Rey, 2007; Asker and Bar-Isaac, 2014) and intensively discussed in legal practice.[12] However, there is still fierce debate about whether RPM is anti-competitive and should be prohibited by policymakers. In this section, we extend our benchmark model to allow for multiple retailers, and see whether downstream competition changes the manufacturer's incentive to control retail price and quantity.

Consider an environment that is identical to the benchmark model in Section 2 with the only exception that now there are n symmetric retailers, indexed by superscript $j \in \{1, 2, \ldots, n\}$. At date 0, the manufacturer offers a contract to each retailer. The contract for retailer j specifies the date-0 price p^j, quantity q^j, cash transfer T_0^j, the date-1 cash repayment T_1^j and the return shipment R^j. The last two components are contingent on retailer j's report $\hat{\omega}^j$.[13] Retailers then decide whether to accept their corresponding contracts simultaneously. At date 1, the retail demand ω is realized and retailers make their reports. In the spirit of Kreps and Scheinkman (1983), we assume that demand is allocated according to efficient rationing, and when some retailers post the same price, their allocated demand should be equal. Moreover, the distribution of demand $F(\omega; p)$ is determined by the highest price in the market, i.e., $\max\{p^1, p^2, \ldots, p^n\}$. We say that a sequence of contracts $\Gamma^1, \Gamma^2, \ldots, \Gamma^n$ are optimal if they maximize the manufacturer's profits subject to all the constraints listed in

[12] See, e.g., Leegin Creative Leather Products, Inc. v. PSKS, Inc., dba Kay's Kloset...Kay's Shoes, 551 U.S. 877 (2007). https://www.supremecourt.gov/opinions/06pdf/06-480.pdf.

[13] For simplicity, we assume that retailer j's contract cannot depend on the other retailer's report.

in Section 2.

Optimal contracts are then characterized by Proposition 8.

Proposition 8. *If* $\Gamma^1, \Gamma^2, \ldots, \Gamma^n$ *are optimal, then they are identical. Moreover, let*

$$p^* = p^1,\ q^* = nq^1,\ T_0^* = nT_0^1,\ T_1^*(\omega) = nT_1^1(\omega),\ R^*(\omega) = nR^1(\omega).$$

Then $\Gamma^* = (p^*, q^*, T_0^*, T_1^*, R^*)$ *is optimal when there is only one retailer with initial wealth* nW *and reservation utility* $n\underline{u}$*.*

Proof. See Appendix A.6 □

According to Proposition 8, optimal contracts with multiple retailers are closely related to the optimal buyback contract in the single-retailer model. It is as if that retailers are merged together before contracting with the manufacturer. Therefore, by Proposition 4, the structure of optimal contracts as well as the equilibrium price and quantity depends on $n\underline{u}$. In particular, by part (a) of Proposition 4, the retail price and quantity are efficient when the single retailer's reservation utility is sufficiently high, which translates into sufficiently many retailers in the present model. We formally state this result in Corollary 1.

Corollary 1. *When* n *is sufficiently large, the manufacturer proposes the first-best price* p^{FB} *and equally distributes the first-best quantity* q^{FB} *to all retailers. In this case, her profits decreases with* n*.*

Corollary 1 describes the effect of competition under RPM. Since the manufacturer fully controls the retail price through contracts, she equally distributes her products among retailers. As the number of retailers increases, the manufacturer has to produce more to make sure that each retailer receives at least $\underline{u}$. The total supply thus increases to the first-best level q^{FB}, accompanied by an increase in the retail price to the first-best level p^{FB}. After this point, the price and quantity never change, so the manufacturer's profit decreases as competition becomes more intensive.

6 Concluding Remarks

In this paper, we show that the optimal retail contract takes the form of either a wholesale or a buyback contract when the retailer privately observes the realized demand, thereby providing a unified microeconomic foundation for retail contracts. Moreover, as the retailer's bargaining power increases, the optimal contract shifts from wholesale to buyback, and the buyback price becomes nonlinear. The optimal price and quantity are shown to be lower than the first-best level, implying that the manufacturer's market power is reduced by downstream information asymmetry, and supplies are rationed.

Our paper can be regarded as part of the foundation of economic and social institutions with a complete contracting approach. While we take a first step in this direction in the area of retail contracting theory, linking optimal retail contractual forms in response to a variety of economic context to empirical studies on retail markets, especially how vertical relationships, demand fluctuation and inventory management affect the market structure of the retail sector (e.g., Hortaçsu and Syverson, 2015) leaves us a promising research agenda of combining theory and practice in the future.

References

Aghion, P. and J. Tirole (1997): "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105, 1–29.

Arya, A. and B. Mittendorf (2004): "Using Return Policies to Elicit Retailer Information," *RAND Journal of Economics*, 35, 617–630.

Asker, J. and H. Bar-Isaac (2014): "Raising Retailers' Profits: On Vertical Practices and the Exclusion of Rivals," *American Economic Review*, 104, 672–686.

Blair, B. F. and T. R. Lewis (1994): "Optimal Retail Contracts with Asymmetric Information and Moral Hazard," *RAND Journal of Economics*, 25, 284–296.

Bower, A. G. (1993): "Procurement Policy and Contracting Efficiency," *International Economic Review*, 34, 873–901.

Brander, J. A. and T. R. Lewis (1986): "Oligopoly and Financial Structure: The Limited Liability Effect," *American Economic Review*, 76, 956–970.

Burns, M. E. and C. Walsh (1981): "Market Provision of Price-excludable Public Goods: A General Analysis," *Journal of Political Economy*, 89, 166–91.

Cachon, G. P. (2003): "Supply Chain Coordination with Contracts," in *Handbooks in Operations Research and Management Science*, ed. by A. G. de Kok and S. C. Graves, Elsevier, vol. 11, 227–339.

Chen, F. (2003): "Information Sharing and Supply Chain Coordination," in *Handbooks in Operations Research and Management Science*, ed. by A. G. de Kok and S. C. Graves, Elsevier, vol. 11, 341–421.

Chu, L. Y. and D. E. Sappington (2007): "Simple Cost-Sharing Contracts," *American Economic Review*, 97, 419–428.

Dana, J. D. and K. E. Spier (2001): "Revenue Sharing and Vertical Control in the Video Rental Industry," *Journal of Industrial Economics*, 49, 223–245.

Deneckere, R., H. P. Marvel, and J. Peck (1996): "Demand Uncertainty, Inventories, and Resale Price Maintenance," *Quarterly Journal of Economics*, 111, 885–913.

——— (1997): "Demand Uncertainty and Price Maintenance: Markdowns as Destructive Competition," *American Economic Review*, 87, 619–641.

Gale, D. and M. F. Hellwig (1985): "Incentive-Compatible Debt Contracts: The One-Period Problem," *Review of Economic Studies*, 52, 647–663.

Guesnerie, R. and J.-J. Laffont (1984): "A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm," *Journal of Public Economics*, 25, 329–369.

Gui, Z., E.-L. von Thadden, and X. Zhao (2019): "Incentive-Compatibility, Limited Liability and Costly Liquidation in Financial Contracting," *Games and Economic Behavior*, 118, 412–433.

Harstad, B. and T. Mideksa (2020): "Coalitions under Vertical Contracts: A Theory for Mergers, Collusion, and Conservation," Working paper.

Hellwig, M. F. (2000): "Financial Intermediation with Risk Aversion," *Review of Economic Studies*, 67, 719–742.

——— (2001): "Risk Aversion and Incentive Compatibility with Ex Post Information Asymmetry," *Economic Theory*, 18, 415–438.

——— (2010): "Incentive Problems with Unidimensional Hidden Characteristics: A Unified Approach," *Econometrica*, 78, 1201–1237.

Holmström, B. (1979): "Moral Hazard and Observability," *Bell Journal of Economics*, 10, 74–91.

Hortaçsu, A. and C. Syverson (2015): "The Ongoing Evolution of US Retail: A Format Tug-of-War," *Journal of Economic Perspectives*, 29, 89–112.

Jullien, B. and P. Rey (2007): "Resale Price Maintenance and Collusion," *RAND Journal of Economics*, 38, 983–1001.

Kreps, D. M. and J. A. Scheinkman (1983): "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes," *Bell Journal of Economics*, 14, 326–337.

Krishnan, H. and R. A. Winter (2007): "Vertical Control of Price and Inventory," *American Economic Review*, 97, 1840–1857.

Marvel, H. P. and S. McCafferty (1984): "Resale Price Maintenance and Quality Certification," *RAND Journal of Economics*, 15, 346–359.

Marvel, H. P. and J. Peck (1995): "Demand Uncertainty and Returns Policies," *International Economic Review*, 36, 691–714.

Mirrlees, J. A. (1974): "Optimum Accumulation Under Uncertainty: The Case of Stationary Returns to Investment," in *Allocation Under Uncertainty: Equilibrium and Optimality*, ed. by J. H. Drèze, Springer, 36–50.

Montez, J. (2015): "Controlling Opportunism in Vertical Contracting when Production Precedes Sales," *RAND Journal of Economics*, 46, 650–670.

Montez, J. and N. Schutz (2021): "All-Pay Oligopolies: Price Competition with Unobservable Inventory Choices," *Review of Economic Studies*, 88, 2407–2438.

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

Nöldeke, G. and K. M. Schmidt (1995): "Option Contracts and Renegotiation: A Solution to the Hold-up Problem," *RAND Journal of Economics*, 26, 163–179.

Pasternack, B. A. (1985): "Optimal Pricing and Return Policies for Perishable Commodities," *Marketing Science*, 4, 166–176.

Petruzzi, N. C. and M. Dada (1999): "Pricing and the newsvendor problem: A review with extensions," *Operations Research*, 47, 183–194.

Rey, P. and J. Tirole (1986): "The Logic of Vertical Restraints," *American Economic Review*, 76, 921–939.

Rogerson, W. P. (2003): "Simple Menus of Contracts in Cost-Based Procurement and Regulation," *American Economic Review*, 93, 919–926.

Royden, H. L. and P. Fitzpatrick (2010): *Real Analysis (4th Edition)*, vol. 32, Pearson.

Schmidt, K. M. (2003): "Convertible Securities and Venture Capital Finance," *Journal of Finance*, 58, 1139–1166.

Shaffer, G. (1991): "Slotting Allowances and Resale Price Maintenance: A Comparison of Facilitating Practices," *RAND Journal of Economics*, 22, 120–135.

Townsend, R. M. (1979): "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory*, 21, 265–293.

Wang, S., H. Gurnani, and U. Subramanian (2020): "The Informational Role of Buyback Contracts," *Management Science,*, forthcoming.

Winter, R. A. (1993): "Vertical Control and Price Versus Nonprice Competition," *Quarterly Journal of Economics*, 108, 61–76.

Appendix

A.1 Proof of Lemma 4

Let $\{(a_j, b_j) : j \in J\}$ be a partition of Ω, where J is a countable index set, such that

$$\Omega = \bigcup_{j \in J} (a_j, b_j),$$

and for any $j, j' \in J$, $j \neq j'$, we have

$$(a_j, b_j) \cap (a_{j'}, b_{j'}) = \varnothing.$$

This partition is unique, so J is uniquely determined by Γ. Since V is nonincreasing on $(\omega, q - R(\omega))$, it must also be nonincreasing on any of these open intervals (a_j, b_j). Moreover, for any j and $\omega \in (a_j, b_j)$, $R(\omega) \geq q - b_j$, for any $\omega \notin \Omega$, $R(\omega) = q - \omega$.

Note that the partition of Ω can be either finite or countably infinite, so one has either $|J| < +\infty$ or $|J| = +\infty$, respectively. We will first discuss the case where $|J|$ is finite. In this case any admissible contract is weakly dominated by a local buyback contract.

Case 1. When $|J| < +\infty$, $\{(a_j, b_j) : j \in J\}$ is well-ordered by $\leq$, and so is the set of all the endpoints of these open intervals $\{a_j, b_j : j \in J\}$. We will use $\{\omega_i : i = 0, 1, \ldots, n\}$ to relabel $\{a_j, b_j : j \in J\} \cup \{0, q\}$, and let $0 = \omega_0 < \omega_1 < \cdots < \omega_n = q$.

Consider a local buyback contract $\hat{\Gamma}$ determined by $\{(\omega_i, t_i) : i = 0, 1, \ldots, n\}$, where for any $i \geq 1$, t_i is chosen to make $\mathrm{E}_\omega[V(\omega)|\omega_{i-1} < \omega < \omega_i] = \mathrm{E}_\omega[\hat{V}(\omega)|\omega_{i-1} < \omega < \omega_i]$. The retailer is thus indifferent between Γ and $\hat{\Gamma}$. The manufacturer's payoff has two possible cases.

(a) When $(\omega_{i-1}, \omega_i) \subseteq \Omega$, then $\hat{T}_1(\omega) \geq T_1(\omega)$ for any $\omega \in (\omega_{i-1}, \omega_i)$. We prove this by contradiction. Suppose that $W - T_0 + p\hat{\omega} \geq T_1(\hat{\omega}) > \hat{T}_1(\hat{\omega})$ for some $\hat{\omega} \in (\omega_{i-1}, \omega_i)$. By construction,

$$\begin{aligned} W - T_0 + p\hat{\omega} \geq T_1(\hat{\omega}) > \hat{T}_1(\hat{\omega}) \quad &\Longrightarrow \quad \hat{T}_1(\hat{\omega}) = W - T_0 + t_i < W - T_0 + p\hat{\omega}, \\ &\Longrightarrow \quad T_1(\hat{\omega}) > W - T_0 + t_i. \end{aligned}$$

For any $\omega \in (\hat{\omega}, \omega_i)$, (IC)-(IC-L) implies that either $W - T_0 + p\hat{\omega} < T_1(\omega)$, which means a type-$\hat{\omega}$ retailer cannot misreport ω; or $V(\hat{\omega}) \leq V(\omega)$, which means a type-$\hat{\omega}$ retailer does not want to misreport ω. We conclude from either cases that $V(\omega) > W - T_0 + t_i$.

For any $\omega \in (\omega_{i-1}, \hat{\omega})$, $V(\omega)$ is nonincreasing on (ω_{i-1}, ω_i), so we also have $V(\omega) \geq V(\hat{\omega}) > W - T_0 + t_i$. Thus $V(\omega) > W - T_0 + t_i$ for any $\omega \in (\omega_{i-1}, \omega_i)$, which violates our construction of $\hat{\Gamma}$.

(b) When $(\omega_{i-1}, \omega_i) \not\subseteq \Omega$, (ω_{i-1}, ω_i) is a subset of the interior of $[0, q]/\Omega$. For any $\omega \in (\omega_{i-1}, \omega_i)$, (FR) must bind.

(a) and (b) jointly imply that the manufacturer is weakly better off in $\hat{\Gamma}$.

Case 2. When $|J| = +\infty$, $\inf_{j \in J} |b_j - a_j| = 0$. For any $\delta > 0$ sufficiently small, the set $J_\delta = \{(a_j, b_j) : b_j - a_j \geq \delta, j \in J\}$ is nonempty, finite, and thus well-ordered. If we construct $\hat{\Gamma}_\delta$ using J_δ as that of Case 1, the retailer is still indifferent between Γ and $\hat{\Gamma}_\delta$. Moreover, as $\delta \to 0$, the manufacturer's payoff is approximated by $\hat{\Gamma}_\delta$. This process will finally give us a sequence $\delta_1, \delta_2, \ldots$, with $\lim_{n \to +\infty} \delta_n = 0$, and a sequence of local buyback contract $\hat{\Gamma}_{\delta_1}, \hat{\Gamma}_{\delta_1}, \ldots$, with $\lim_{n \to +\infty} \hat{\Gamma}_{\delta_n} = \Gamma$.

A.2 Proof of Lemma 5

We omit the constant term in L:

$$L = \sum_{i=1}^{n} \int_{\omega_{i-1}}^{\omega_i} [T_1(\omega) + v_m R(\omega)] dF(\omega) + \int_q^{+\infty} [T_1(\omega) + v_m R(\omega)] dF(\omega)$$
$$- \lambda \left\{ \sum_{i=1}^{n} \int_{\omega_{i-1}}^{\omega_i} V(\omega) dF(\omega) + \int_q^{+\infty} V(\omega) dF(\omega) \right\}.$$

Our first step is to discuss the range of λ. For any $i = 1, 2, \ldots, n$, there must be one of the following three cases:

(a) If $t_i < p\omega_{i-1} + v_r(q - \omega_i)$, then for any $\omega \in (\omega_{i-1}, \omega_i]$, $p\omega \geq t_i - v_r(q - \omega_i)$, which means $T_1(\omega) = W - T_0 + t_i - v_r(q - \omega_i)$, $R(\omega) = q - \omega_i$. Therefore,

$$\frac{\partial L}{\partial t_i} = (1 - \lambda)[F(\omega_i) - F(\omega_{i-1})].$$

(b) If $p\omega_{i-1} + v_r(q - \omega_i) < t_i < p\omega_{i-1} + v_r(q - \omega_{i-1})$, then there exists a cutoff ω_i^1, determined by $p\omega_i^1 + v_r(q - \omega_i) = t_i$, such that

$$\omega < \omega_i^1 \Rightarrow T_1(\omega) = W - T_0 + p\omega, \ R(\omega) = \frac{t_i - p\omega}{v_r},$$
$$\omega \geq \omega_i^1 \Rightarrow T_1(\omega) = W - T_0 + t_i - v_r(q - \omega_i), \ R(\omega) = q - \omega_i.$$

Therefore,

$$\frac{\partial L}{\partial t_i} = (1-\lambda)[F(\omega_i) - F(\omega_{i-1})] - (1 - \frac{v_m}{v_r})[F(\omega_i^1) - F(\omega_{i-1})].$$

(c) If $p\omega_{i-1} + v_r(q - \omega_{i-1}) < t_i$, then there exist two cutoffs ω_i^1 and ω_i^2, determined by $p\omega_i^2 + v_r(q - \omega_i^2) = p\omega_i^1 + v_r(q - \omega_i) = t_i$, such that

$$\begin{aligned}
\omega < \omega_i^2 &\Rightarrow T_1(\omega) = W - T_0 + p\omega,\ R(\omega) = q - \omega, \\
\omega_i^2 \leq \omega < \omega_i^1 &\Rightarrow T_1(\omega) = W - T_0 + p\omega,\ R(\omega) = \frac{t_i - p\omega}{v_r}, \\
\omega \geq \omega_i^1 &\Rightarrow T_1(\omega) = W - T_0 + t_i - v_r(q - \omega_i),\ R(\omega) = q - \omega_i.
\end{aligned}$$

Therefore,

$$\frac{\partial L}{\partial t_i} = (1-\lambda)[F(\omega_i) - F(\omega_i^2)] - (1 - \frac{v_m}{v_r})[F(\omega_i^1) - F(\omega_i^2)].$$

It should also be noted that $\hat{L}$ is continuous at the two nondifferentiable points: $t_i = p\omega_{i-1} + v_r(q - \omega_i)$ and $t_i = p\omega_{i-1} + v_r(q - \omega_{i-1})$. By first-order conditions, we can prove that $\lambda \in (v_m/v_r, 1)$, which immediately tells us that (PC_r) binds in any optimal local buyback contract:

If $\lambda \leq v_m/v_r$, $\partial L/\partial t_i > 0$ for any t_i. Hence for any $i = 1, 2, \ldots, n$, $t_i \leq p\omega_i + v_r(q - \omega_i)$ must hold with equality. The retailer gets nothing from such a contract, which is a violation of (PC_r).

If $\lambda \in (v_m/v_r, 1)$, $\partial L/\partial t_i > 0$ for any $t_i < p\omega_{i-1} + v_r(q - \omega_i)$, so we have $t_i \geq p\omega_{i-1} + v_r(q - \omega_i)$.

If $\lambda = 1$, L is negatively correlated with the retailer's expected return of inventory. The contract that maximizes L must have $R(\omega) = 0$, $T_1(\omega) \leq W - T_0$ for any ω, a contradiction to $q > \underline{q}$.

If $\lambda > 1$, $\partial L/\partial t_i < 0$ for any t_i, a contradiction since t_i has no lower bound.

The final step is to solve for the optimal local buyback contract using $\lambda \in (v_m/v_r, 1)$. For any $i = 1, 2, \ldots, n$, let

$$L_i = \int_{\omega_{i-1}}^{\omega_i} [T_1(\omega) + v_m R(\omega)] dF(\omega) - \lambda \int_{\omega_{i-1}}^{\omega_i} V(\omega) dF(\omega)$$

denote the retailer's Lagrangian conditional on $\omega \in (\omega_{i-1}, \omega_i)$. Then for $i \leq n-1$, $\partial L/\partial \omega_i = \partial L_i/\partial \omega_i + \partial L_{i+1}/\partial \omega_i$.

If $p\omega_{i-1} + v_r(q-\omega_i) < t_i < p\omega_{i-1} + v_r(q-\omega_{i-1})$,

$$\begin{aligned}
\frac{\partial L_i}{\partial \omega_i} &= [p\omega_i^1 + v_m(q-\omega_i)]F(\omega_i) - v_m[F(\omega_i) - F(\omega_i^1)] - \lambda t_i F(\omega_i), \\
\frac{\partial L_i}{\partial \omega_{i-1}} &= -[p\omega_{i-1} + \frac{v_m}{v_r}(t_i - p\omega_{i-1})]F(\omega_{i-1}) + \lambda t_i F(\omega_{i-1}) \\
&= [(\lambda - \frac{v_m}{v_r})t_i - (1 - \frac{v_m}{v_r})p\omega_{i-1}]F(\omega_{i-1}) \\
&< [(\lambda v_r - v_m)(q-\omega_{i-1}) - (1-\lambda)p\omega_{i-1}]F(\omega_{i-1}).
\end{aligned}$$

The last inequality comes from $t_i < p\omega_{i-1} + v_r(q-\omega_{i-1})$.

If $p\omega_{i-1} + v_r(q-\omega_{i-1}) < t_i$,

$$\begin{aligned}
\frac{\partial L_i}{\partial \omega_i} &= [p\omega_i^1 + v_m(q-\omega_i)]F(\omega_i) - v_m[F(\omega_i) - F(\omega_i^1)] - \lambda t_i F(\omega_i), \\
\frac{\partial L_i}{\partial \omega_{i-1}} &= -[p\omega_{i-1} + v_m(q-\omega_{i-1})]F(\omega_{i-1}) + \lambda[p\omega_{i-1} + v_r(q-\omega_{i-1})]F(\omega_{i-1}) \\
&= [(\lambda v_r - v_m)(q-\omega_{i-1}) - (1-\lambda)p\omega_{i-1}]F(\omega_{i-1}).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{\partial L}{\partial \omega_i} &< [p\omega_i^1 - (1-\lambda)p\omega_i + \lambda v_r(q-\omega_i)]F(\omega_i) - v_m[F(\omega_i) - F(\omega_i^1)] - \lambda t_i F(\omega_i) \\
&= (1-\lambda)[t_i - p\omega_i - v_r(q-\omega_i)]F(\omega_i) - v_m[F(\omega_i) - F(\omega_i^1)] \\
&< 0.
\end{aligned}$$

The second equality comes from $p\omega_i^1 + v_r(q-\omega_i) = t_i$, and the last inequality comes from $t_i < p\omega_i + v_r(q-\omega_i)$.

$\partial L/\partial \omega_i < 0$ implies that the boundary constraint derived in the proof of $\lambda \in (v_m/v_r, 1)$, i.e., $t_i \geq p\omega_{i-1} + v_r(q-\omega_i)$, must bind. However, in this case $\omega_i^1 = \omega_{i-1}$, and

$$\lim_{\omega_i^1 \to \omega_{i-1}+} \frac{\partial L}{\partial t_i} = (1-\lambda)[F(\omega_i) - F(\omega_{i-1})] > 0.$$

Thus $t_i \leq p\omega_i + v_r(q-\omega_i)$ binds with equality, which essentially means $\omega_i = \omega_{i-1}$. Hence, a local buyback contract is optimal only if $n = 1$ and (PC_r) binds, which essentially implies that it is a buyback contract.

A.3 Proof of Proposition 5

When the retailer can get cash from salvaging unsold inventories, the manufacturer chooses q, T_0, and T_1 to maximize

$$-cq + T_0 + T_1,$$

subject to (L_0), (LL_1'), and (PC_r), where (PC_r) can be simplified as

$$W - T_0 + (p - v_r)Q(q) + v_r q - T_1 \geq W + \underline{u}.$$

The Lagrangian of this problem is

$$\begin{aligned} L(q, T_0, T_1) = &-cq + T_0 + T_1 + \lambda_0(W - T_0) + \lambda_1(W - T_0 + v_r q - T_1) \\ &+ \mu[-T_0 + \int_0^{+\infty} p\omega + v_r(q - \omega)dF(\omega) - T_1 - \underline{u}]. \end{aligned}$$

First-order necessary conditions are

$$\frac{\partial L}{\partial q} = -c + \lambda_1 v_r + \mu[p - (p - v_r)F(q)] = 0, \tag{32}$$

$$\frac{\partial L}{\partial T_0} = 1 - \lambda_0 - \lambda_1 - \mu = 0, \tag{33}$$

$$\frac{\partial L}{\partial T_1} = 1 - \lambda_1 - \mu = 0. \tag{34}$$

By (33) and (34), $\lambda_0 = 0$ and $\lambda_1 + \mu = 1$, but $\mu = 0$ violates (32). Thus, there must be $\mu > 0$.

If (LL_1') is slack, then $\lambda_1 = 0$, $\mu = 1$, (32) becomes identical to the first-order condition in the first best. As a result, the manufacturer will offer $q = q^{FB}$, and receive expected payoff

$$\mathrm{E}_\omega\, u_m(\omega) = S(q^{FB}).$$

The range of $\underline{u}$ is determined by a binding (PC_r) and a slack (LL_1'), i.e.,

$$\underline{u} > (p - v_r)Q(q^{FB}) - W.$$

If (LL_1') binds, then (PC_r) implies that

$$\mathrm{E}_\omega\, u_m(\omega) = W - (c - v_r)q, \tag{35}$$

where q is determined by a binding (PC_r), i.e.,

$$(p - v_r)Q(q) = W + \underline{u}.$$

Note that the right-hand side of (35) is bounded above by W, which implies that when $\underline{u} \to 0$ and $W \to 0$, $\mathrm{E}_\omega\, u_m(\omega) \to 0$. However, in the benchmark model, when $\underline{u} = W = 0$, $\mathrm{E}_\omega\, u_m(\omega) = S(q^{FB})$. Hence, there exists a cutoff $\hat{u} \leq (p - v_r)Q(q^{FB})$ such that when $W + \underline{u} < \hat{u}$, the manufacturer is worse off when the retailer can salvage cash from unsold inventories.

A.4 Proof of Proposition 6

In order to compare the optimal price and quantity with the first best, we denote by $\mathcal{P}(q)$ the solution of (29) given q, and $\mathcal{Q}(p)$ the solution of (4) given p. Thus, $p^{FB} = \mathcal{P}(q^{FB})$, and $q^{FB} = \mathcal{Q}(p^{FB})$. Moreover,

$$\begin{aligned}
S_p(q; \mathcal{P}(q)) = 0 \ &\Rightarrow\ \mathcal{P}'(q) S_{pp}(q; \mathcal{P}(q)) + S_{pq}(q; \mathcal{P}(q)) = 0, \\
S_q(\mathcal{Q}(p); p) = 0 \ &\Rightarrow\ \mathcal{Q}'(p) S_{qq}(\mathcal{Q}(p); p) + S_{pq}(\mathcal{Q}(p); p) = 0.
\end{aligned}$$

By (27), $\mathcal{P}'(q) > 0$ and $\mathcal{Q}'(p) > 0$.

When $0 < t \leq v_r q$, first-order derivatives of (30) are given by

$$\begin{aligned}
\frac{\partial\, \mathrm{E}_\omega\, u_m(\omega)}{\partial p} &= S_p(q; p) - (1 - \frac{v_m}{v_r})\{t_p - Q(\overline{\omega}; p) - pQ_p(\overline{\omega}; p) - [1 - F(\overline{\omega}; p)](t_p - \overline{\omega})]\} \\
&= S_p(q; p) - (1 - \frac{v_m}{v_r})\{F(\overline{\omega}; p)t_p - Q(\overline{\omega}; p) - pQ_p(\overline{\omega}; p) + [1 - F(\overline{\omega}; p)]\overline{\omega}\}, \quad (36) \\
\frac{\partial\, \mathrm{E}_\omega\, u_m(\omega)}{\partial q} &= S_q(q; p) - (1 - \frac{v_m}{v_r})F(\overline{\omega}; p)t_q \\
&= [1 - (1 - \frac{v_m}{v_r})F(\overline{\omega}; p)][p - (p - v_r)F(q; p)] - c. \quad (37)
\end{aligned}$$

We first look at (36). It can be verified that $t_p = S_p(q; p)$. Thus,

$$\left.\frac{\partial\, \mathrm{E}_\omega\, u_m(\omega)}{\partial p}\right|_{p=\mathcal{P}(q)} = (1 - \frac{v_m}{v_r})\{Q(\overline{\omega}; p) + pQ_p(\overline{\omega}; p) - [1 - F(\overline{\omega}; p)]\overline{\omega}\}\Big|_{p=\mathcal{P}(q)}.$$

Moreover, $S_{pq}(q; p) > 0$ implies

$$S_p(q; p) > S_p(\overline{\omega}; p) = Q(\overline{\omega}; p) + (p - v_r)Q_p(\overline{\omega}; p).$$

Therefore,

$$\frac{\partial \operatorname{E}_\omega u_m(\omega)}{\partial p}\bigg|_{p=\mathcal{P}(q)} < (1-\frac{v_m}{v_r})\{v_r Q_p(\overline{\omega};p) - [1-F(\overline{\omega};p)]\overline{\omega}\}\bigg|_{p=\mathcal{P}(q)} < 0.$$

By the concavity of $S(q;p)$, we must have $p^* < \mathcal{P}(q^*)$. Next, according to (37),

$$\frac{\partial \operatorname{E}_\omega u_m(\omega)}{\partial q}\bigg|_{q=\mathcal{Q}(p)} = -(1-\frac{v_m}{v_r})F(\overline{\omega};p)[p-(p-v_r)F(q;p)]\bigg|_{q=\mathcal{Q}(p)} < 0,$$

which implies $q^* < Q(p^*)$. Finally, by the concavity of S,

$$S(q^{FB};p^{FB}) - S(q^*;p^*) \leq S_p(q^*;p^*)(p^{FB}-p^*) + S_q(q^*;p^*)(q^{FB}-q^*). \tag{38}$$

The left-hand side of (38) is positive. In the right-hand side, $p^* < \mathcal{P}(q^*)$ and $q^* < Q(p^*)$ imply that $S_p(q^*;p^*) > 0$ and $S_q(q^*;p^*) > 0$. Hence, $p^* < p^{FB}$ and $q^* < q^{FB}$.

When $t > v_r q$, first-order derivatives of (31) are given by

$$\frac{\partial \operatorname{E}_\omega u_m(\omega)}{\partial p} = \frac{v_m}{v_r}S_p(q;p) + (1-\frac{v_m}{v_r})\{Q(\overline{\omega};p) + pQ_p(\overline{\omega};p) + [1-F(\overline{\omega};p)](t_p - \overline{\omega})\}, \tag{39}$$

$$\frac{\partial \operatorname{E}_\omega u_m(\omega)}{\partial q} = \frac{v_m}{v_r}[p-(p-v_r)F(q;p)] - c + (1-\frac{v_m}{v_r})p[1-F(\overline{\omega};p)](\frac{t_q - v_r}{p-v_r}). \tag{40}$$

Similarly, we start from (39). Note that the definition of t can be rewritten as

$$(p-v_r)Q(\underline{\omega};p) + v_r q = (p-v_r)Q(q;p) + v_r q - W - \underline{u}. \tag{41}$$

The right-hand side of (41) is exactly the expression of t from (15). Therefore,

$$S_p(\underline{\omega};p) + [1-F(\underline{\omega};p)](t_p - \underline{\omega}) = S_p(q;p).$$

Plugging this into (39), and evaluating the derivative at $p = \mathcal{P}(q)$ give us

$$\begin{aligned}\frac{\partial \operatorname{E}_\omega u_m(\omega)}{\partial p}\bigg|_{p=\mathcal{P}(q)} &= (1 - \frac{v_m}{v_r})\{S_p(\overline{\omega};p) + v_r Q_p(\overline{\omega};p) - \frac{1-F(\overline{\omega};p)}{1-F(\underline{\omega};p)}S_p(\underline{\omega};p) \\ &\quad - [1-F(\overline{\omega};p)](\overline{\omega}-\underline{\omega})\}\bigg|_{p=\mathcal{P}(q)} \\ &< (1-\frac{v_m}{v_r})[\frac{1-F(\overline{\omega};p)}{1-F(\underline{\omega};p)}]\{S_p(\overline{\omega};p) - S_p(\underline{\omega};p) \\ &\quad - [1-F(\underline{\omega};p)](\overline{\omega}-\underline{\omega})\}\bigg|_{p=\mathcal{P}(q)},\end{aligned}$$

where second inequality comes from $S_p(\overline{\omega};p) < 0$ and $Q_p(\overline{\omega};p) < 0$. Furthermore,

$$S_{pq}(q;p) = 1 - F(q;p) - (p - v_r)F_p(q;p) < 1 - F(q;p).$$

Therefore,

$$S_p(\overline{\omega};p) - S_p(\underline{\omega};p) = \int_{\underline{\omega}}^{\overline{\omega}} S_{pq}(q;p)dq < \int_{\underline{\omega}}^{\overline{\omega}} [1-F(\underline{\omega};p)]dq = [1-F(\underline{\omega};p)](\overline{\omega}-\underline{\omega}),$$

which implies

$$\frac{\partial \operatorname{E}_\omega u_m(\omega)}{\partial p}\bigg|_{p=\mathcal{P}(q)} < 0,$$

and $p^* < \mathcal{P}(q^*)$. Next, (17) also gives us

$$t_q = v_r + (p - v_r)\frac{1-F(q;p)}{1-F(\underline{\omega};p)}.$$

Plugging this into (40), and evaluating the derivative at $q = \mathcal{Q}(p)$ give us

$$\frac{\partial \operatorname{E}_\omega u_m(\omega)}{\partial q}\bigg|_{q=\mathcal{Q}(p)} = (1-\frac{v_m}{v_r})\{p[1-F(q;p)]\frac{1-F(\overline{\omega};p)}{1-F(\underline{\omega};p)} - c\}\bigg|_{q=\mathcal{Q}(p)} < 0,$$

which also implies $q^* < Q(p^*)$. Hence, we have $p^* < p^{FB}$ and $q^* < q^{FB}$.

A.5 Proof of Proposition 7

Part (b) of the proposition is proved by sequential rationality. As discussed in Section 5.2, given the structure of Γ_1, the optimal Γ_0 should still be a buyback contract. Hence, we only need to analyze p_0 and q_0.

We use $u_m(\omega|\Gamma_0)$ and $u_m(\omega|\Gamma_1)$ to represent the manufacturer's utility from contracts Γ_0 and Γ_1, respectively. Then, given the contracts characterized in parts (a) and (b) of the proposition, the manufacturer's objective becomes

$$\begin{aligned} \mathrm{E}_\omega\, u_m(\omega) &= \mathrm{E}_\omega\, u_m(\omega|\Gamma_0) + \mathrm{E}_\omega\, u_m(\omega|\Gamma_1) \\ &= \mathrm{E}_\omega\, u_m(\omega|\Gamma_0) + \beta \int_{q_0}^{+\infty} p_0(\omega - q_0) dF(\omega; p_0). \end{aligned}$$

First-order derivatives are

$$\frac{\partial\, \mathrm{E}_\omega\, u_m(\omega)}{\partial p_0} = \frac{\partial\, \mathrm{E}_\omega\, u_m(\omega|\Gamma_0)}{\partial p_0} + \beta \int_{q_0}^{+\infty} \frac{\partial p_0 f(\omega; p_0)}{\partial p_0}(\omega - q_0) d\omega, \tag{42}$$

$$\frac{\partial\, \mathrm{E}_\omega\, u_m(\omega)}{\partial q_0} = \frac{\partial\, \mathrm{E}_\omega\, u_m(\omega|\Gamma_0)}{\partial q_0} - \beta p_0[1 - F(q_0; p_0)]. \tag{43}$$

It is straightforward to see that $p_0 > p^*$ and $q_0 < q^*$. When $\beta \to 0$, the model boils down to the benchmark model in Section 2. When $\beta \to 1$, (43) implies that $q_0 \to 0$, therefore from (42) we have $p_0 \to p^m$.

A.6 Proof of Proposition 8

Since retailers are symmetric, it suffices to prove the proposition when $n = 2$. First, we show $p^1 = p^2$ by contradiction. Suppose that $p^1 < p^2$. Then increasing p_1 will not change the distribution of ω as demand is determined by the higher price p_2. If the manufacturer increases p_1 and the date-1 cash repayment T_1^1 uniformly so that the retailer is indifferent, she can ensure a higher payoff from Γ^1 without affecting her payoff from Γ^2. Therefore, the manufacturer optimally offers $p^1 = p^2$. It is then straightforward to see that Γ^1 and Γ^2 are identical. Moreover, they are both buyback or wholesale contracts, because by Proposition 4, the optimality of buyback contracts is robust to any distribution of demand.

Let t^1 be the additional upfront payment determined by Γ^1. Then, when $0 < t^1 \le v_r q^1$, the manufacturer's expected payoff is

$$\begin{aligned} \mathrm{E}_\omega\, u_m(\omega) &= 2\{W - cq^1 + \int_0^{\varpi^1} [\frac{1}{2}p^1\omega + \frac{v_m}{v_r}(t^1 - \frac{1}{2}p^1\omega)] dF(\omega; p^1) + \int_{\varpi^1}^{+\infty} t^1 dF(\omega; p^1)\} \\ &= 2(W - cq^1) + \{\int_0^{\varpi^1} [p^1\omega + \frac{v_m}{v_r}(2t^1 - p^1\omega)] dF(\omega; p^1) + \int_{\varpi^1}^{+\infty} 2t^1 dF(\omega; p^1)\}, \end{aligned} \tag{44}$$

where $\overline{\omega}^1 = 2t^1/p^1$, and t^1 is determined by a binding (PC_r),

$$\begin{aligned} t^1 &= \int_0^{2q^1} [\frac{1}{2}p^1\omega + v_r(q^1 - \frac{1}{2}\omega)]dF(\omega|p^1) + \int_{2q^1}^{+\infty} p^1q^1 dF(\omega|p^1) - W - \underline{u} \\ &= \frac{1}{2}(p^1 - v_r)Q(p^1, 2q^1) + v_r q^1 - W - \underline{u}. \end{aligned} \tag{45}$$

Comparing (44) with (16) and (45) with (15), we can conclude that the manufacturer's expected utility is equivalent to that from our benchmark model where the only retailer has reservation utility $2(W + \underline{u})$. Hence, the proposition is proved.