Supply Chain Frictions*

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

February 2024

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 quite 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 bargaining power. In a buyback contract, the manufacturer requests an upfront payment from the retailer and buys back the unsold inventory with a possibly variable 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: Supply chain frictions, buyback contracts, incentive compatibility, limited liability.

JEL Classification: D82, D86, L42, L60.

---

*This paper replaces an earlier paper entitled "Optimal Retail Contracts with Return Policies" with early work on some of the research presented here. We thank Simon Loertscher, Leslie Marx, Volker Nocke, Harry Pei, Patrick Rey, Nico Schutz, as well as seminar and conference participants at the 2022 PKU-NUS Annual International Conference on Quantitative Finance and Economics, the 2023 Markets, Contracts, and Organizations Conference (Canberra), the 2023 European Association for Research in Industrial Economics (EARIE) Annual Conference (Rome), and the 2023 Asia-Pacific Industrial Organization Conference (APIOC, Hong Kong) for stimulating discussions. 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. Email: zgui@nus.edu.sg.

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

¶Department of Economics, Monash Business School, Monash University. Email: xjzhao81@gmail.com.

1 Introduction

Supply chains are a major element of international trade and make decentralized manufacturing possible, often across considerable distances. Yet, exactly because of the large distances that typically separate the firms in a supply chain, these relationships are subject to inherent frictions. Two important such frictions are the mismatch between supply and demand and the information asymmetry between upstream and downstream participants. Supply-demand mismatch arises naturally because the interaction between upstream manufacturers and downstream users of their products typically is subject to unavoidable long lead times and demand uncertainty. Accordingly, production upstream must occur before downstream demand is realized. When demand is large, the downstream firm can only use up to the quantity delivered, and the excess demand is lost. When downstream demand is small, unsold upstream production may have second-best uses downstream or be salvaged upstream. As transactions with consumers take place downstream, the retailer directly observes the realized demand, but such information typically is not available upstream due to separation. This gives rise to the second friction, which is the asymmetry of information between the member firms of the supply chain. Thus, in a decentralized supply chain, these frictions are central to the working of supply chains and they naturally interact. This calls for a systematic investigation and is the objective of the present paper.

In this paper, we examine the simplest upstream-downstream relationship possible, namely a model in which a manufacturer sells its products through a retailer operating in a remote market.[1] Two key factors characterize this relationship. First, production precedes sales, thus the contracting parties have to write down the contractual terms before 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 manufacturer cannot directly observe the retailer's sales and revenue. As a result, a vertical contract should: (1) set the optimal trade volume, and (2) specify the appropriate contractual form, including state-contingent incentive-compatible prices, so as to induce the retailer to execute the contract as intended, enabling the gains from trade.

In practice, a number of contracts are used, including buyback, franchise, quantity flexibility, revenue sharing, service commitment, two-part tariff, wholesale price, and others. There is large literature investigating these contracts and their consequences on supply chain performance in various contexts (see Cachon (2003) and Shen et al. (2019) for extensive discussions). Existing studies typically focus on particular contract forms and examine their impact on mitigating supply-demand mismatch, supply chain inefficiencies, or information

[1] In its simplest form, this is the classical *news vendor problem* in Operations Management.

asymmetry. In fact, the survey paper Shen et al. (2019) mentions 455 papers and categorizes 131 papers based on the contract form investigated. While they provide various angles to understand the practical use of these contracts, the conceptual question *what is the optimal contract* under the typical frictions present in supply chains, however, has received less attention. As Cachon (2003) put it succinctly: "practice has been used as a motivation for theoretical work, but theoretical work has not found its way into practice".

In order to make progress on this front, we study a game with two dates. Contracting happens at date 0, which determines the quantity, an immediate cash transfer, and rules for the execution of the future interaction. Retail demand is realized at date 1, after which there can be further transfers and a return of unsold inventories. If date-1 cash transfers and return shipments are contingent on realized demand, they must be incentive-compatible. Hence, ex-post, the contracting parties face a tradeoff between cash and returns. These two channels differ in many aspects. Cash payments are bounded by the retailer's initial wealth plus his date-1 revenue, which is increasing in realized demand, but private information of the retailer. This limited liability constraint has been considered in the literature on contracts 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 create.[2] Return shipments, on the other hand, cannot exceed the total amount of leftover inventories, which is negatively related to ex-post demand. Finally, cash transfers are efficient, while returning unsold inventories typically is not.[3] This assumption, together with the retailer's limited liability, makes the return of unsold inventories an imperfect screening device. As a result, the manufacturer's objective is to minimize the use of returns by appropriate incentive-compatible contracts.

Without assuming any functional form of contracts, we find that the optimal contract takes a rather simple form. At date 0, in exchange for the shipment of goods, the retailer makes an upfront payment to the manufacturer. After the realization of demand at date 1, there is either no more obligation for transfers or shipments from the retailer (wholesale contracts), or, under a buyback contract, the retailer transfers state-dependent quantities of cash and inventory to the manufacturer. When the realized demand is high, the cash payment is a high, fixed target and there are no inventory returns, while the cash payment is low and returns are positive when the realized demand is too low for the retailer to pay the fixed cash target in full. The rationale for this result is the following. Facing the adverse

[2] 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.

[3] 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). This simplification is less realistic for non-perishable goods, such as clothes and electronic devices.

selection problem, the manufacturer wants to elicit the retailer's private information about sales, so the return of unsold inventory is used as a punishment when the reported demand is low. However, ex ante the manufacturer also aims to minimize inefficient inventory returns by the retailer. Therefore, return policies will be used only when the reported demand is sufficiently low. Notably, and differently from related contractual structures in corporate finance, which we discuss below, the associated buyback price may not be constant. Hence, optimal contracts can involve variable or constant ex-post pricing. The wholesale contract thus turns out to be a special case of a constant-price buyback contract.

Proposition 4 summarizes 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 observed in practice. Since the second-best optimal quantity under a buyback contract is strictly lower than the first-best level, the parties prefer wholesale contracts, but these cannot be implemented if the retailer does not have enough bargaining power.

This result is surprisingly robust to various extensions of our benchmark model. When the retail price is endogenous and influences the distribution of demand, Proposition 6 shows that the contract optimally reduces the price below 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 studies supply chains under general contracting subject to information asymmetry and limited liability. There is a lot of interesting prior work on specific contracts in such environments, for example on buyback/return contract in the presence of information asymmetry. Noteworthy are, in particular, Yue and Raghunathan (2007), Hsieh et al. (2008), Taylor and Xiao (2009), and Babich et al. (2012), which all assume that the retailer privately knows the demand distribution. Yue and Raghunathan (2007) compare two specific contract forms: no return and full return, where the latter guarantees that the buyback price equals the wholesale price. Hsieh et al. (2008) do not consider screening contracts but instead compare across three scenarios: the retailer shares demand information truthfully under an all-unit quantity discount contract with buyback, the retailer withholds demand information under an all-unit quantity discount contract with buyback, and the centralized supply chain. They focus on whether structural properties of ordering decisions preserve in these scenarios. In Taylor and Xiao (2009), the retailer can exert efforts to improve demand information accuracy, and only two exogenous contract forms are considered: rebates contract that compensates the retailer for each unit sold to end consumers, and returns contract that specifies a buyback price for each unsold unit. In Babich et al. (2012), the supplier designs a menu of contracts, each of which comprises a wholesale price, a buyback price, and a lump-sum transfer. Kumar and Srinivasan (2007) consider the case in which the retailer holds private risk preference and decides the order quantity and retail price. Our paper has very different focus and is distinct from the aforementioned papers by studying the general contract space without assuming any specific contract forms. Our proposed nonlinear return contract appears to be novel to the supply chain contracting literature[4] and may offer new guidance to practitioners for crafting optimal contracts.

The key feature of our model is to consider the frictions caused by information asymmetry and production-in-advance jointly. There is a large body of literature studying firms' inventory choices and production capacity, originating from Kreps and Scheinkman (1983) and then followed by many others (e.g., Davidson and Deneckere, 1986; Deneckere and Peck, 1995; Deneckere et al., 1996; Maggi, 1996). In a recent paper, Montez and Schutz (2021) apply techniques from all-pay auctions to study price competition where firms have private information about their inventory levels. Our study departs from these papers by arguing that market demand could also be the retailer's private information, which gives rise to the informational role of buyback contracts.[5] A number of papers introduce information asymmetry into vertical contracts (e.g., Rey and Tirole, 1986; Blair and Lewis, 1994; Arya and Mittendorf, 2004), but they do not specifically consider the tension between the retailer's private information about demand and inventory decisions.

[4]In the recent survey paper by Shen et al. (2019, Table 1), a return contract is defined by the right to "return the unsold products at the end of selling season at a return price". This *constant* return price specification is used in all the references therein.

[5]Wang et al. (2020) also examine the signaling role of buyback contracts, while they take the buyback contractual form as exogenous.

Technically, we model an ex-post screening problem with hidden characteristics. When the type set is a continuum, the standard methodology used in the literature is control theory, pioneered by Guesnerie and Laffont (1984) and further developed by Hellwig (2010). However, the control-theoretic approach cannot be applied in the present paper. In our model, each type of retailer's set of deviations 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 equation. 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.[6] We overcome this difficulty using an ironing approach in the spirit of Myerson (1981), and recently Loertscher and Muir (2022), but we differ from the Myersonian approach in two aspects. First, our incentive constraint does not imply the monotonicity of allocation rules, so our ironing process can be applied to problems with weaker incentive constraints than that of standard screening problems. Second, the "ironed" contract in our model need not be optimal and must be further optimized, while in the existing literature the "ironed" mechanism is already optimal.

The rest of this paper is organized as follows. Section 2 introduces the model setup. Section 3 provides the benchmark with symmetric information. Section 4 analyzes the model of asymmetric information and proves the optimality of buyback contracts. Section 5 discusses several key assumptions of the model. Section 6 extends our benchmark model to different environments. Section 7 concludes. Some 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 $[0, +\infty)$.[7] $F(\cdot; p)$ admits a density function f that is positive and bounded almost everywhere. In our baseline model, we assume that p is exogenous and observable. Therefore, we drop the reference to p in this section and the next.

[6] Relatedly, Gui et al. (2019) provide a detailed discussion on how the presence of limited liability affects the analysis of incentive constraints in the financial contracting literature. In particular, that paper shows that ignoring limited liability off the equilibrium path may lead to an over-simplified analysis and sub-optimal contracts.

[7] This formulation of stochastic outcomes, the "parameterized-distribution-function approach", was pioneered by Mirrlees (1974) and Holmström (1979). Early usages in the IO/price theory literature include Burns and Walsh (1981). See Section 6 for a brief structural discussion of this model.

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 $F(\cdot)$. This results in two distortions. First, production must take place prior to the realization of demand, thus there will be a supply-demand mismatch. In Section 6.2 we extend 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 a 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., $q > \omega$, the retailer can only sell up to ω. Thus realized sales satisfy

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

s, too, is unobservable to the manufacturer.

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 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.[8] 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 prior to 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.[9] This description captures many different types of retail contracts in practice.

[8] 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.

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

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 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. 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 $\hat{\omega}$.

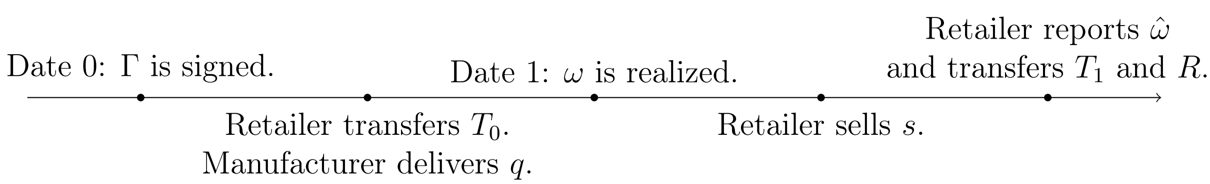


Figure 1: Timeline

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 feasibility, 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 for returns:

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

The second feasibility restriction is that 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:

$$T_0 \le W, \tag{FT$_0$}$$

$$T_1 \le W - T_0 + ps. \tag{FT$_1$}$$

In practice, liquidity constraints (feasibility with respect to payment) arise 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 (FT$_1$), because in practice liquidating leftover inventory typically takes time. In Section 5.1, 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 have the incentive to carry out sales $s(\omega)$ as planned.[10] Overall, this leads to

[10] Note that the last two points are distinct because sales $s(\omega)$ are unobservable upstream.

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

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

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

$$T_1(\hat{\omega}) \leq W - T_0 + p\hat{s} \tag{IC-FT$_1$}$$

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-FT$_1$) is special 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-FT$_1$). 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-FT$_1$) 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.[11] We will discuss how we circumvent this problem in Section 4 and what would happen if we dropped these restrictions to the incentive constraint in Section 5.2.

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 (or

[11] See the large literature following Guesnerie and Laffont (1984).

individual-rationality) constraints

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

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

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

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

To simplify the analysis, we restrict attention to schedules T_1 and R with finitely many discontinuities. This assumption allows us to simplify the characterization of contracts in Lemma 5 below when considering its Lagrangian dual problem. This or similar assumptions are standard, either explicitly or implicitly, in much of the literature on contracting with asymmetric information (Guesnerie and Laffont, 1984; Lacker and Weinberg, 1989).

We call a contract *feasible* if it satisfies the constraints (FS), (FR), (FT$_0$), and (FT$_1$), *admissible* if it satisfies all the constraints, and *optimal* if it is a solution to this problem. Moreover, if two contracts differ only on a set of states with zero measure, 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 $\Gamma = (q, T_0, s, T_1, R)$.

3 Symmetric Information Benchmark

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 other 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 $Q(q)$. Moreover, $Q'(q) = 1 - F(q)$ and $Q''(q) = -f(q)$. 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). □

4 Optimal Contract under Asymmetric Information

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 subject 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.

4.1 Implementation by wholesale contracts

Under a wholesale contract, there are no state-contingent transfers. In our framework, for an optimal contract to be wholesale the date-1 component must therefore satisfy

$$T_1(\omega) = T_1 \text{ and } R(\omega) = 0 \text{ for all } \omega.$$

In this case, the contract provides no costly incentives, 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 (IR_r) is feasible if and only if (6) holds. □

Lemma 1 characterizes the situations in which the retailer's initial liquidity constraint (FT_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 $\underline{q}$ 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.[12] In the following two subsections, we assume that $q > \underline{q}$ and investigate this second option.

4.2 The sales decision

In this subsection, we simplify the contracting problem by eliminating the in- and off-equilibrium sales decisions $s(\omega)$ and $\hat{s}$, respectively. To begin with we note that it is redundant to consider the choice of T_0 and T_1 separately, since (FT_0) and all payoffs involve only $T_0 + T_1(\omega)$ and T_1 is unbounded below. Hence, we can add any constant $C \leq W - T_0$ to T_0 and subtract it from $T_1(\omega)$ for all ω without changing any of the analysis and results. Without loss of generality, we assume from now on that $T_0 = W$.[13]

Next, denote by V the total ex-post value transfer to the manufacturer as evaluated by the retailer:

$$V(\omega) = T_1(\omega) + v_r R(\omega). \tag{8}$$

Using $T_0 = W$ and (8), the retailer's ex-post utility is

$$u_r(\omega, \hat{\omega}, s) = ps - T_1(\hat{\omega}) + v_r[q - s - R(\hat{\omega})] \tag{9}$$

$$= (p - v_r)s - V(\hat{\omega}) + v_r q. \tag{10}$$

As (10) shows, $u_r(\omega, \hat{\omega}, s)$ is strictly increasing in s for each ω, and the total transfer V is not affected by increasing s. Hence, it is optimal for the retailer ex-post to set s as large as possible. But as (9) shows, for a given deviation $\hat{\omega}$ the maximum sales volume that the retailer can choose is

$$\hat{s} = \min(q - R(\hat{\omega}), \min(\omega, q)) \tag{11}$$

$$= \min(q - R(\hat{\omega}), \omega). \tag{12}$$

By (11), if $R(\hat{\omega}) \leq q - \min(\omega, q)$, the maximum sales volume for deviation $\hat{\omega}$ is $\min(\omega, q)$, the maximum feasible volume given by (FS). If $R(\hat{\omega}) > q - \min(\omega, q)$, the maximum sales volume for deviation $\hat{\omega}$ is $q - R(\hat{\omega})$ because of the feasibility constraint for returns (FR).

The conceptual difficulty with the IC constraint (IC)-(IC-FT_1) is that the observable off-the-equilibrium choices $T_1(\hat{\omega})$ and $R(\hat{\omega})$ depend on the unobservable action $\hat{s}$. The preceding

[12] 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.

[13] If W is large or q small, as in the preceding subsection, this implies negative T_1, i.e. payments from the manufacturer to the retailer, in some states at date 1.

distinction allows to eliminate $\hat{s}$ from the IC constraint .

Lemma 2. *A feasible contract is incentive-compatible if and only if for any $\omega, \hat{\omega} \geq 0$ with $T_1(\hat{\omega}) \leq p \min(\omega, q)$,*

(IC1) $R(\hat{\omega}) \leq q - \min(\omega, q)$ *implies* $V(\omega) + (p - v_r)(\min(\omega, q) - s(\omega)) \leq V(\hat{\omega})$,

(IC2) $R(\hat{\omega}) > q - \min(\omega, q)$ *implies* $V(\omega) + (p - v_r)(q - R(\hat{\omega}) - s(\omega)) \leq V(\hat{\omega})$.

Proof. **"Only if":** Suppose that a feasible contract Γ satisfies (IC)-(IC-FT$_1$). Consider any $\omega, \hat{\omega} \geq 0$ with $T_1(\hat{\omega}) \leq p \min(\omega, q)$. If $R(\hat{\omega}) \leq q - \min(\omega, q)$, $\hat{s} = \min(\omega, q)$ satisfies (IC-FS)-(IC-FT$_1$) by definition. Inserting this $\hat{s}$ into (IC) gives (IC1). If $R(\hat{\omega}) > q - \min(\omega, q)$, let $\hat{s} = q - R(\hat{\omega})$. The assumption then reads $\hat{s} < \min(\omega, q)$, which implies (IC-FS). (IC-FR) holds trivially by construction. Finally, use feasibility (FR) and (FT$_1$), evaluated at $\hat{\omega}$, twice to obtain $T_1(\hat{\omega}) \leq ps(\hat{\omega}) \leq p(q - R(\hat{\omega})) = p\hat{s}$, which is (IC-FT$_1$). Inserting this $\hat{s}$ into (IC) then gives (IC2).

"If": Suppose that Γ satisfies (IC1) and (IC2) for any $\omega, \hat{\omega} \geq 0$ with $T_1(\hat{\omega}) \leq p \min(\omega, q)$. Consider ω, $\hat{\omega} \geq 0$, and $\hat{s}$ that satisfy the constraints (IC-FS)-(IC-FT$_1$). (IC-FS) and (IC-FT$_1$) imply $T_1(\hat{\omega}) \leq p \min(\omega, q)$. By the argument leading to (12), when $R(\hat{\omega}) \leq q - \min(\omega, q)$, the best possible ex-post deviation is $\hat{s} = \min(\omega, q)$. Since (IC) holds for this $\hat{s}$ by (IC1), it holds for all admissible $\hat{s}$. When $R(\hat{\omega}) > q - \min(\omega, q)$, the best possible ex-post deviation is $\hat{s} = q - R(\hat{\omega})$. Since (IC) holds for this $\hat{s}$ by (IC2), it holds for all admissible $\hat{s}$. Hence, Γ is incentive-compatible. □

Turning to the in-equilibrium sales decision $s(\omega)$, the argument leading up to (12) shows that, if there is a $\hat{\omega}$ such that $q - R(\hat{\omega}) \leq \omega \leq q$, with at least one inequality being strict, incentive-compatibility alone does not suffice to conclude that the retailer chooses $s = \min(\omega, q)$ ex post. However, as the following lemma shows, such a contract structure would not be optimal ex ante.

Lemma 3. *A contract is optimal only if $s(\omega) = \min(\omega, q)$ for 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 a positive measure. Ignoring some measure-theoretic fine points, consider an alternative contract $\hat{\Gamma}$ that for all $\omega \in A$ increases $s(\omega)$ to $s'(\omega) = s(\omega) + \varepsilon(\omega) <$

$\min(\omega, q)$, reduces $R(\omega)$ to $R'(\omega) = R(\omega) - \varepsilon(\omega) > 0$,[14] 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 (FT_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.

As any optimal contract has $s(\omega) = \min(\omega, q)$ for almost all ω, the piecewise continuity of T_1 and R implies that this is true pointwise. □

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 some $\hat{\omega}$. 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.

4.3 Implementing $q > \underline{q}$: Local buyback contracts

Lemmas 2 and 3 taken together now make it possible to identify several key features of optimal contracts that can be used to characterize them constructively. The following result follows directly from these two lemmas.

Lemma 4. *If T_1 and R are the ex-post components of an optimal contract, then:*

(O1) For any $\omega, \hat{\omega} \le q$ with $T_1(\hat{\omega}) \le p\omega$,

- *(a) if $R(\hat{\omega}) \le q - \omega$, then $V(\omega) \le V(\hat{\omega})$,*
- *(b) if $R(\hat{\omega}) > q - \omega$, then $V(\omega) \le V(\hat{\omega}) + (p - v_r)[R(\hat{\omega}) - (q - \omega)]$.*

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

(O2) For all $\omega \geq q$, $T_1(\omega) = T_1(q)$ and $R(\omega) = 0$.

Lemma 4 implies that in an optimal contract, any demand realization higher than q is irrelevant to the retailer's incentive problem, because beyond q payments are flat at $T_1(q)$ and returns are zero. Moreover, for any demand realization lower than q, an optimal contract only needs to prevent two types of deviation: First, if $\hat{\omega}$ is a feasible deviation under $s(\omega) = \omega$, then the retailer's total transfer $V(\omega)$ must be lower than that of $\hat{\omega}$. This is the case of (O1a) in Lemma 4. Second, if $\hat{\omega}$ is infeasible given the optimal sales at the realized state ω, then the retailer who wants to deviate must deliberately sell less, i.e., $\hat{s} = q - R(\hat{\omega})$, to fulfill the return requirements $R(\hat{\omega})$. In this case, $V(\omega)$ can be higher than $V(\hat{\omega})$, but is still bounded by a (affine) linear function of ω. This is the case of (O1b) in Lemma 4. But as we shall see momentarily, *(O1)* does not imply monotonicity.

Using Lemma 3 and Lemma 4, for any given q, we can relax the problem of finding the optimal date-1 component of the contract as follows:

$$\max_{T_1, R} \mathrm{E}_\omega[T_1(\omega) + v_m R(\omega)],$$

subject to

- the pointwise constraints

$$T_1(\omega) = T_1(q) \text{ and } R(\omega) = 0, \tag{O2}$$

  for all $\omega \geq q$, and

$$0 \leq R(\omega) \leq q - \omega, \tag{FR}$$
$$T_1(\omega) \leq p\omega, \tag{FT$_1$}$$

  for all $\omega \in [0, q]$,

- the incentive constraints (O1) of Lemma 4,

- and the retailer's participation constraint

$$(p - v_r)\, \mathrm{E}_\omega\left[\min(\omega, q)\right] - \mathrm{E}_\omega[V(\omega)] + v_r q \geq W + \underline{u}. \tag{IR$_r$}$$

In the relaxed problem, using the incentive constraints (O1) is still difficult because of the inherent off-the-equilibrium feasibility constraints. In classical adverse selection models, incentive-compatibility usually reduces to an envelope formula, which implies that the agent's indirect utility function is absolutely continuous on the state space (Milgrom and Segal, 2002; Hellwig, 2010). However, this is not the case in our model due to the presence of (FR) and (FT_1). To wit, consider the following date-1 component of a contract for some given quantity $q > \underline{q}$:

$$T_1(\omega) = \begin{cases} \alpha\omega & \omega < q/2, \\ pq/2 & \omega \geq q/2; \end{cases} \quad R(\omega) = \begin{cases} q - \omega & \omega < q/2, \\ 0 & \omega \geq q/2. \end{cases}$$

The retailer's total transfer function V is

$$V(\omega) = \begin{cases} \alpha\omega + v_r(q - \omega) & \omega < q/2, \\ pq/2 & \omega \geq q/2; \end{cases}$$

and his indirect utility function is

$$U_r(\omega) = u_r(\omega, \omega, \min(\omega, q)) = \begin{cases} (p - \alpha)\omega & \omega < q/2, \\ p\omega + v_r(q - \omega) - pq/2 & q/2 \leq \omega < q, \\ pq/2 & \omega \geq q. \end{cases}$$

If $\alpha \in [p - v_r, p]$, then this contract is admissible. Note that $R(\omega) > q/2$ for any $\omega \in [0, q/2)$, while $T_1(\omega) \geq pq/2$ for any $\omega \geq q/2$. The retailer with type $\omega \in [0, q/2)$ cannot exaggerate his type above $q/2$ due to the liquidity constraint for cash. The retailer with type $\omega \geq q/2$ does not want to understate his type below $q/2$ due to the incentive constraint. However, if $\alpha > p - v_r$, U_r has an upward jump at $q/2$. If also $\alpha > v_r$, the function V associated with T_1, R is non-monotonic and discontinuous. Figure 2 provides a graphical illustration of this example.

Moreover, unlike in classical models such as Guesnerie and Laffont (1984), incentive-compatibility does not imply the monotonicity of the decision function R. To see this, it suffices to consider a particularly simple case where V is constant on the whole state space, as depicted in Figure 3. Such a contract is clearly incentive-compatible because no type has an incentive to misreport other types (formally, both conditions of Lemma 2 are always satisfied). Given that $V(\omega) = T_1(\omega) + v_r R(\omega)$ is constant, the only constraint for T_1 is (FT_1) and for R is (FR). Hence, there is considerable leeway in choosing the shapes of T_1 and R.

We therefore propose a more indirect approach. The main idea is to show that any admissible contract that meets the necessary conditions in Lemma 3 and Lemma 4 is strictly

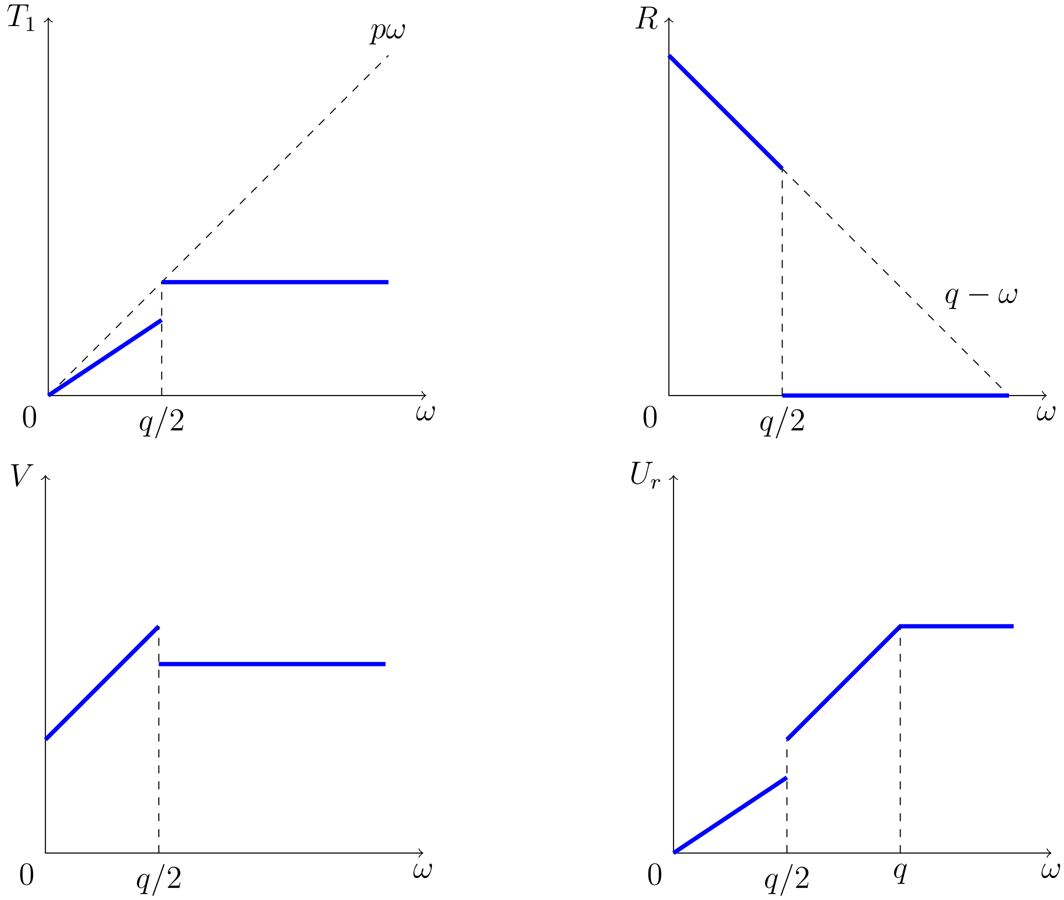


Figure 2: The non-monotonicity of V and discontinuity of the indirect utility function

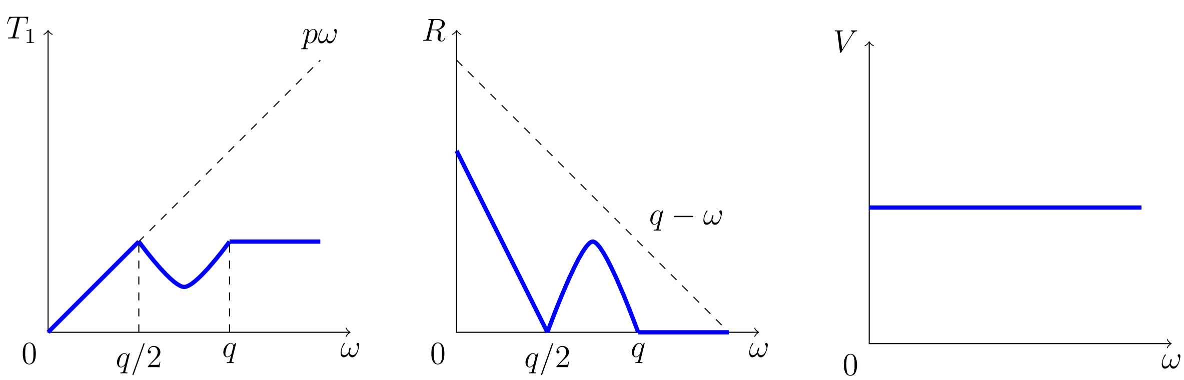


Figure 3: The non-monotonicity of T_1 and R.

dominated by an admissible contract of a specific form, if it does not already have this form. This implies that optimal contracts exist and are of this form. In order to construct such contracts, we apply the ironing technique used in traditional screening models (Baron and Myerson, 1982; Guesnerie and Laffont, 1984) and auction theory (Myerson, 1981), using some results from real analysis and the theory of convex functions (see, e.g., Rockafellar, 1970; Rudin, 1987).

Here is a sketch of our ironing approach. Since return shipments lead to efficiency loss, we want to substitute them with cash transfers at some intervals. If the retailer's expected utility remains unchanged after this process, then we get an improved contract. However, this improvement is subject to feasibility. Suppose that we reduce return shipments at $\hat{\omega}$ from $R(\hat{\omega})$ to $\hat{R}(\hat{\omega})$, and let ω be a state satisfying $p\omega \geq T_1(\hat{\omega})$ and $\hat{R}(\hat{\omega}) < q - \omega < R(\hat{\omega})$. In the original contract, the relationship between $V(\omega)$ and $V(\hat{\omega})$ is governed by (O1b) of Lemma 4, while in the improved contract, $V(\omega)$ and $V(\hat{\omega})$ should satisfy (O1a) of Lemma 4, which is tighter than (O1b).

Therefore, we start by ironing the total transfer function V to make it nondecreasing on the state space. This is followed by a restructuring of the function R to make it nonincreasing. The ironed contract has an important property: $p\omega \geq T_1(\hat{\omega})$ implies $V(\omega) \leq V(\hat{\omega})$. In other words, ω and $\hat{\omega}$ both satisfy (O1a) of Lemma 4 irrespective of R. We can then conveniently use cash transfers to replace return shipments on those ironed intervals.

Figure 4 provides a graphical illustration of the ironing procedure when $\mathcal{N} = \{1, 2\}$.

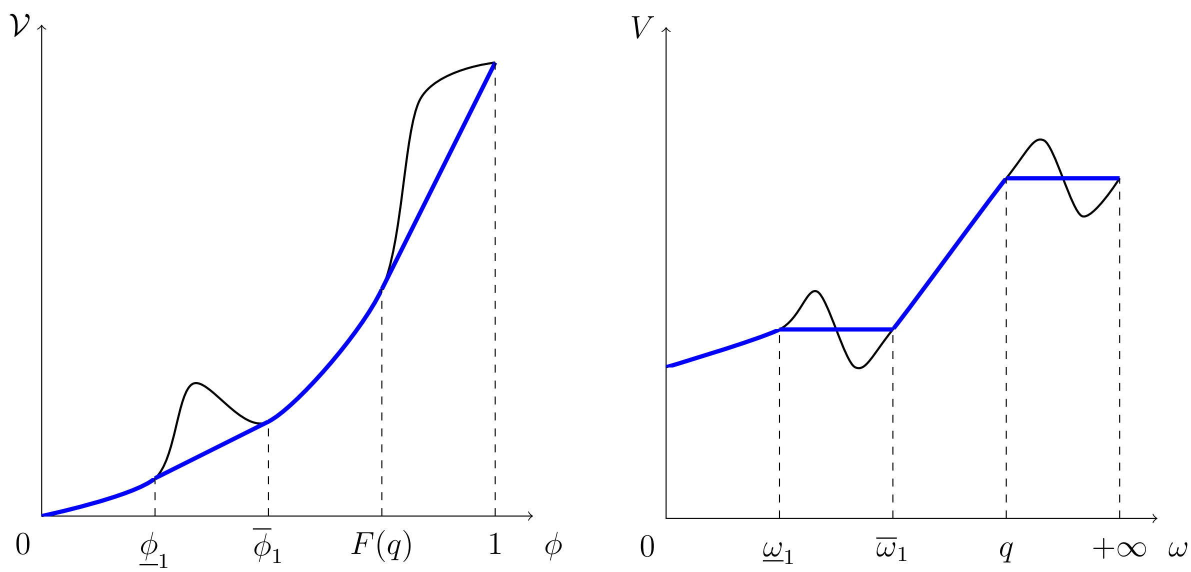


Figure 4: The ironing approach

In the ironed contract, there exist countably many intervals $[\underline{\omega}_n, \overline{\omega}_n)$, indexed by $n \in \mathcal{N}$, such that on each of them: (1) $V(\omega)$ equals a constant $t_n + v_r \max(q - \overline{\omega}_n, 0)$; and (2) there is a minimum level $\max(q - \overline{\omega}_n, 0)$ of return shipments. Intuitively, for realizations in $[\underline{\omega}_n, \overline{\omega}_n)$ the contract can be interpreted as setting a reference repayment level of t_n in cash and of $\max(q - \overline{\omega}_n, 0)$ of returns. If the retailer's cash is insufficient (which implies that there are more unsold units), the manufacturer buys back additional units at a marginal buyback price v_r. Hence, locally in $[\underline{\omega}_n, \overline{\omega}_n)$, a retailer who considers deviating marginally trades off cash payments and returns at the rate $1 : v_r$, as the unrestricted incentive constraint (IC) would require. In this sense, the new contract, which locally resembles a buyback contract on the interval $[\underline{\omega}_n, \overline{\omega}_n)$, provides optimal incentives. This gives rise to the following definition.

Definition 1. *A contract is a **local buyback contract** if there exist countably many disjoint intervals $[\underline{\omega}_n, \overline{\omega}_n)$ and an equal number of constants t_n, indexed by $n \in \mathcal{N}$, such that:*

(a) For any ω,

(a.1) if $\omega \in [\underline{\omega}_n, \overline{\omega}_n)$, then $T_1(\omega) = \min(p\omega, t_n)$, $R(\omega) = \max(q - \overline{\omega}_n, 0) + \max((t_n - p\omega)/v_r, 0)$,

(a.2) if $\omega \notin \bigcup_{n \in \mathcal{N}} [\underline{\omega}_n, \overline{\omega}_n)$, then $T_1(\omega) \leq p \min(\omega, q)$, $R(\omega) = \max(q - \omega, 0)$.

(b) For any $n \in \mathcal{N}$, $t_n \leq (p - v_r)\underline{\omega}_n + v_r \min(\overline{\omega}_n, q)$.

(c) $V(\omega)$ is nondecreasing and continuous.

One can immediately verify from Definition 1 that any local buyback contract is admissible. By (a) and (b), it satisfies (FT_1) and (FR). For incentive-compatibility, note that the retailer has no incentive to exaggerate his type because $V(\omega)$ is nondecreasing. If $\omega \in [\underline{\omega}_n, \overline{\omega}_n)$, the retailer cannot understate his type below $\underline{\omega}_n$, because $R(\hat{\omega}) > q - \underline{\omega}_n$ for all $\hat{\omega} < \underline{\omega}_n$, and he does not want to deviate to $\hat{\omega} \in [\underline{\omega}_n, \omega)$, because V is constant on $[\underline{\omega}_n, \overline{\omega}_n)$. If $\omega \notin \bigcup_{n \in \mathcal{N}} [\underline{\omega}_n, \overline{\omega}_n)$, he cannot understate his type because $R(\hat{\omega}) > \max(q - \omega, 0)$ for all $\hat{\omega} < \omega$.

Lemma 5 shows why these contracts are important.

Lemma 5. *If a contract implementing $q > \underline{q}$ is optimal, then it is a local buyback contract.*

Proof. Assume that Γ implementing $q > \underline{q}$ is optimal, and $\hat{\Gamma}$ is the corresponding local buyback contract constructed from the ironing procedure. By construction, the retailer's total transfer function in $\hat{\Gamma}$ is exactly $\hat{V}$. Since $\hat{\mathcal{V}}(1) = \mathcal{V}(1)$, the retailer's expected total transfer in $\hat{\Gamma}$ is identical to that of Γ.

Suppose that $T_1(\omega) > \hat{T}_1(\omega)$ for some $\omega \in [\underline{\omega}_n, \overline{\omega}_n)$. Then $T_1(\omega) > t_n$. There are three possible cases to be discussed.

Case 1: $V(\omega) \le k_n$. Then $R(\omega) < q - \overline{\omega}_n$ for some $\overline{\omega}_n < q$. Recall that in the right neighborhood of $\overline{\omega}_n$, V is strictly increasing, so there exists $\hat{\omega} > \overline{\omega}_n$ such that $\hat{\omega} \le q - R(\omega)$ and $V(\hat{\omega}) > k_n \ge V(\omega)$, a contradiction to (O1a) of Lemma 4.

Case 2: $V(\omega) > k_n$ for some $\overline{\omega}_n \le q$. Then there exists $\hat{\omega} \in (\omega, \overline{\omega}_n]$ such that $V(\hat{\omega}) \le k_n < V(\omega)$. By (O1a) of Lemma 4, $T_1(\hat{\omega}) > p\omega \ge T_1(\omega) > t_n$, which implies that $R(\hat{\omega}) < q - \overline{\omega}_n$. The state $\hat{\omega}$ fits the assumption of Case 1 and will again lead to the same contradiction.

Case 3: $V(\omega) > k_n$ for some $\overline{\omega}_n = +\infty$. Then $T_1(\omega) > t_n = k_n$. Since $\lim_{\omega \to +\infty} V(\omega) = k_n$, there exists $\hat{\omega} > \omega$ such that $V(\hat{\omega}) = T_1(\hat{\omega}) < T_1(\omega) \le V(\omega)$, a contradiction to (O1a) of Lemma 4.

Therefore, $T_1(\omega) \le \hat{T}_1(\omega)$ for any $\omega \in [\underline{\omega}_n, \overline{\omega}_n)$. Since $\mathrm{E}_\omega[\hat{V}(\omega) | \underline{\omega}_n \le \omega < \overline{\omega}_n] = \mathrm{E}_\omega[V(\omega) | \underline{\omega}_n \le \omega < \overline{\omega}_n]$, we have $\mathrm{E}_\omega[\hat{R}(\omega) | \underline{\omega}_n \le \omega < \overline{\omega}_n] \le \mathrm{E}_\omega[R(\omega) | \underline{\omega}_n \le \omega < \overline{\omega}_n]$. Moreover, for any $\omega \notin \bigcup_{n \in \mathcal{N}} [\underline{\omega}_n, \overline{\omega}_n)$, Γ and $\hat{\Gamma}$ are identical. In expectation, $\hat{\Gamma}$ uses (weakly) more cash and (weakly) less return shipments than Γ. Thus, $\hat{\Gamma}$ is also optimal, and there must be $T_1(\omega) = \hat{T}_1(\omega) = \min(p\omega, t_n)$ for almost all $\omega \in [\underline{\omega}_n, \overline{\omega}_n)$.

If $T_1(\omega) = p\omega$, then $R(\omega) < \max(q - \omega, 0)$. Otherwise, by (O1b) of Lemma 4, both (FT_1) and (FR) bind for any states lower than ω, V is therefore strictly increasing from 0 to ω, and ω cannot be an interior point of an ironed interval. Now that (FT_1) is binding while (FR) is slack at ω, V is weakly decreasing in the neighborhood of ω. If $T_1(\omega) = t_n$, then V is constant in the neighborhood of ω. Thus, on the whole interval $[\underline{\omega}_n, \overline{\omega}_n)$, V is weakly decreasing. Note that $V(\overline{\omega}_n+) \ge k_n$. If $V(\overline{\omega}_n-) < k_n$, then V has an upward jump at $\overline{\omega}_n$, a contradiction to (O1b) of Lemma 4. Hence, $V(\overline{\omega}_n-) \ge k_n$. Recall that $\hat{V}$ equals the constant k_n on $[\underline{\omega}_n, \overline{\omega}_n)$. The monotonicity of V must imply that V also equals the constant k_n on $[\underline{\omega}_n, \overline{\omega}_n)$. In summary, Γ is equivalent to $\hat{\Gamma}$. □

Figure 5 depicts how we construct the local buyback contract used in Lemma 5. In the top panel, the black line is an arbitrary admissible contract. The blue lines demonstrate our ironing approach, which results in the local buyback contract in the bottom panel.

Remark 1. *Section 4.3 illustrates the key difference between our constructive approach and other ironing approaches in the literature. The ironing approach in Myerson (1981) and many other papers in screening, monopoly pricing, and auction serves to ensure that the "ironed" mechanism has a monotone allocation rule, which is both necessary and sufficient for incentive-compatibility. However, in our model, incentive-compatibility itself does not imply monotonicity neither of T_1 and R, nor of the indirect utility U_r, nor their continuity. We use ironing to get a monotonic V, but additional structures of T_1 and R (the local buyback structure) are still needed for incentive-compatibility. In this sense, our approach can be applied to problems with incentive constraints weaker than that of the standard screening*

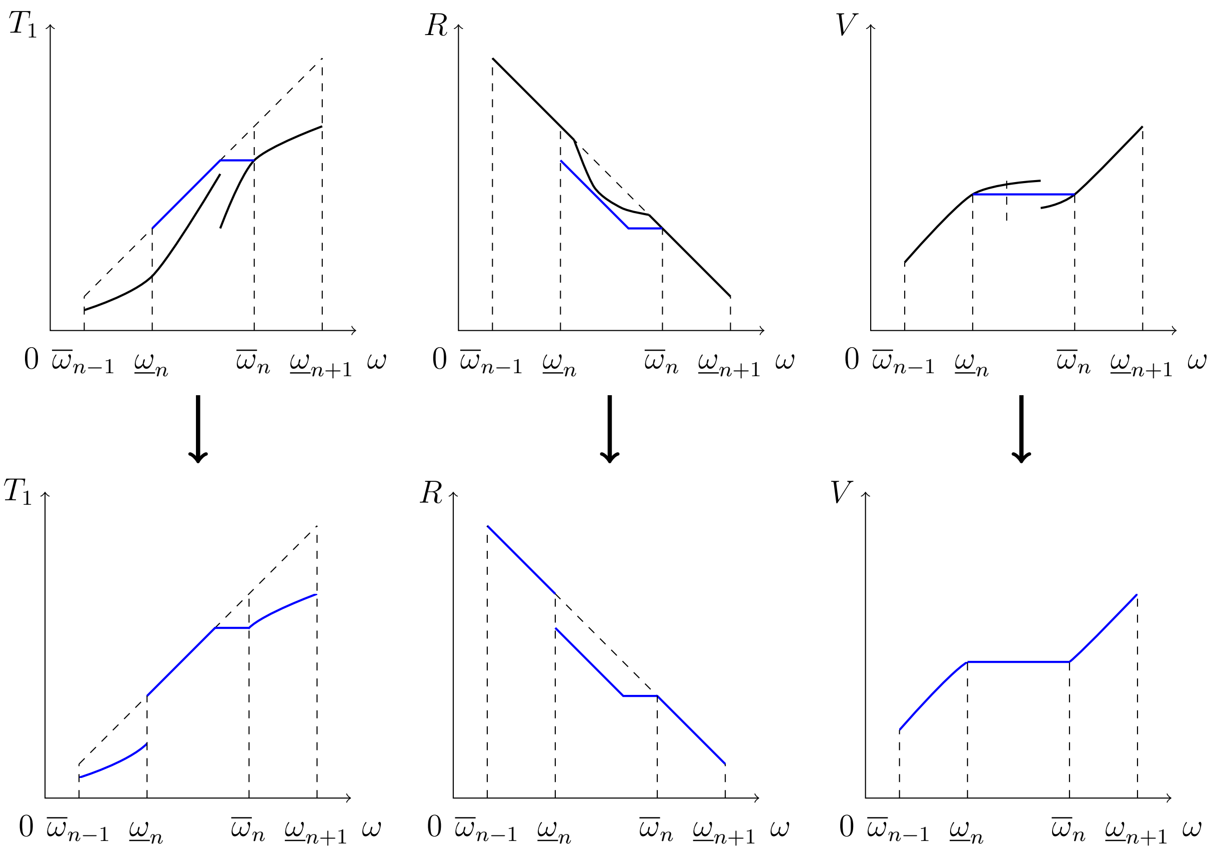


Figure 5: Constructing a local buyback contract

problems.

4.4 Implementing $q > \underline{q}$: Buyback contracts

Lemma 5 transforms the problem of finding optimal contracts into the problem of finding optimal local buyback contracts. In any local buyback contract, T_1 and R can have discontinuities only at $\underline{\omega}_n$. Therefore, the assumption that T_1 and R have only finitely many discontinuities restricts our attention to local buyback contracts with $|\mathcal{N}| = N < +\infty$. The analysis is further simplified if one assumes that the distribution of retail demand has a nondecreasing hazard rate, i.e., that $f(\omega)/[1 - F(\omega)]$ is nondecreasing. From now on we make this assumption, which, while not without loss of generality, is standard in the mechanism design literature and is satisfied by many commonly used distributions (see, e.g., Bagnoli and Bergstrom, 2005).

When $|\mathcal{N}|$ is finite, without loss of generality denote $\mathcal{N} = \{1, 2, \ldots, N\}$. For convenience, we rank the set of cutoffs $\{\underline{\omega}_n, \overline{\omega}_n | 1 \leq n \leq N\}$ by ascending order of n, and denote $\overline{\omega}_0 = 0$.

Thus, $N = n(q)$, and

$$\overline{\omega}_{n-1} \leq \underline{\omega}_n \leq \overline{\omega}_n \leq \underline{\omega}_{n+1} \text{ for any } 1 \leq n \leq N-1. \tag{13}$$

The following result completes our characterization of local buyback contracts, by "filling in the holes" between the buyback intervals.

Lemma 6. *If a local buyback contract is optimal, then for each $\omega \in [\overline{\omega}_n, \underline{\omega}_{n+1})$ and $0 \leq n \leq N-1$,*

$$T_1(\omega) = t_n + p(\omega - \overline{\omega}_n). \tag{14}$$

Proof. (O1b) of Lemma 4 requires that $V'(\omega) \leq p - v_r$, and Definition 1 requires that $R(\omega) = q - \omega$, so (14) is equivalent to $V'(\omega) \leq p - v_r$ being binding on any $[\overline{\omega}_n, \underline{\omega}_{n+1})$. Suppose that (14) does not hold for some n. We can reduce $\underline{\omega}_{n+1}$ to $\underline{\omega}'_{n+1}$ and increase $T_1(\omega)$ to make (14) hold on $[\overline{\omega}_n, \underline{\omega}'_{n+1})$. Here, $\underline{\omega}'_{n+1}$ is chosen to keep t_{n+1} unchanged and $V(\omega)$ continuous.[15] The resulting contract is still a local buyback contract which uses more cash and less return shipments in expectation. The retailer is worse off, but he can be compensated by the manufacturer with lump-sum transfers. The manufacturer is strictly better off even after the compensation. Therefore, a binding (14) is necessary for optimality. □

Figure 6 graphically illustrates our discussion. In the top panel, the blue line is an arbitrary local buyback contract. The red lines depict how we increase $T_1(\omega)$ to make (14) bind, which yields the improved contract in the bottom panel.

Using the continuity of V, as well as (13) and (14), we can write $\{t_n\}_{2 \leq n \leq N}$ as functions of the cutoffs $\{\underline{\omega}_n, \overline{\omega}_n\}_{1 \leq n \leq N}$ and t_1. To see this, one can compute that

$$V(\omega) = \begin{cases} t_n + v_r(q - \overline{\omega}_n) & \omega \in [\underline{\omega}_n, \overline{\omega}_n),\ 0 \leq n \leq N-1, \\ t_n + v_r(q - \overline{\omega}_n) + (p - v_r)(\omega - \overline{\omega}_n) & \omega \in [\overline{\omega}_n, \underline{\omega}_{n+1}),\ 0 \leq n \leq N-1, \\ t_N & \omega \in [\underline{\omega}_N, +\infty). \end{cases}$$

Continuity implies that

$$t_n + v_r(q - \overline{\omega}_n) + (p - v_r)(\underline{\omega}_{n+1} - \overline{\omega}_n) = t_{n+1} + v_r(q - \overline{\omega}_{n+1}),\ 1 \leq n \leq N-1,$$
$$t_{N-1} + v_r(q - \overline{\omega}_{N-1}) + (p - v_r)(\underline{\omega}_N - \overline{\omega}_{N-1}) = t_N,$$

[15] If $n = N$, we can simply raise $T_1(\omega)$ on $[\overline{\omega}_N, +\infty)$ and the argument continues to hold.

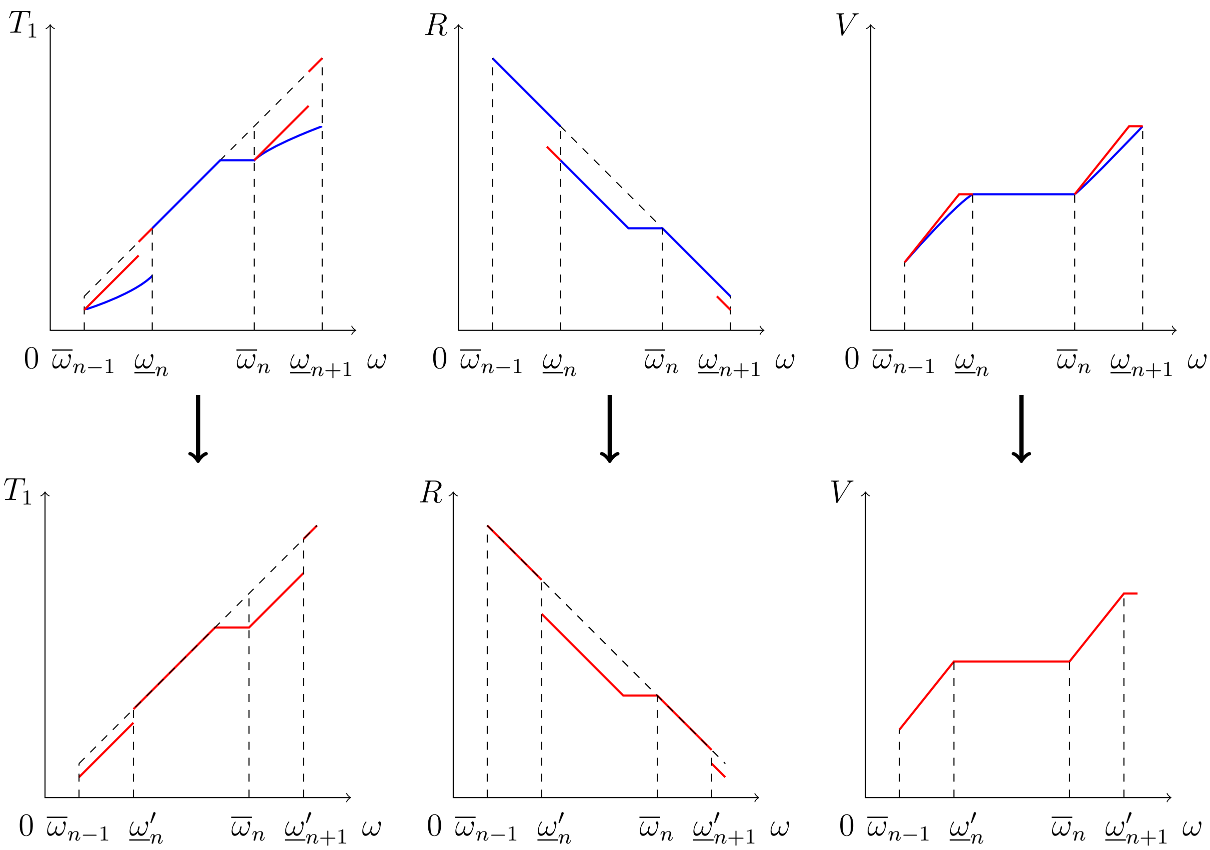


Figure 6: Improving a local buyback contract

which, by recursive substitution, can be simplified to

$$\begin{aligned} t_n &= t_1 - p\overline{\omega}_1 + (p - v_r)\underline{\omega}_n + v_r\overline{\omega}_n - (p - v_r)\sum_{j=2}^{n-1}(\overline{\omega}_j - \underline{\omega}_j),\ 2 \le n \le N-1, \\ t_N &= t_1 - p\overline{\omega}_1 + (p - v_r)\underline{\omega}_N + v_r q - (p - v_r)\sum_{j=2}^{N-1}(\overline{\omega}_j - \underline{\omega}_j). \end{aligned} \tag{15}$$

Thus, (b) and (c) of Definition 1 boil down to the recursive relationship (15) plus the following constraint for t_1:

$$t_1 \le (p - v_r)\underline{\omega}_1 + v_r\overline{\omega}_1, \text{ with equality if } \underline{\omega} > 0. \tag{16}$$

As a result, both contracting parties' expected payoffs in any local buyback contract can be pinned down by the cutoffs $\{\underline{\omega}_n, \overline{\omega}_n\}_{1\le n\le N}$, t_1, and the quantity q. Thus, standard techniques 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 the retailer's participation constraint (IR_r). If Γ maximizes $\text{E}_\omega[u_m(\omega)]$ subject to (IR_r), the sequence $\{\underline{\omega}_n, \overline{\omega}_n\}_{1\leq n\leq N}$, the parameters t_1, q, and λ must jointly be a stationary point of the Lagrangian

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

subject to (13)-(16), as well as the complementary slackness constraints for (IR_r):

$$\lambda \geq 0, \quad \lambda(\text{E}_\omega[u_r(\omega, \omega, \min(\omega, q))] - W - \underline{u}) = 0. \tag{17}$$

Omitting constant terms in L, we have:

$$L = \int_0^{+\infty} [(1-\lambda)T_1(\omega) - (\lambda v_r - v_m)R(\omega)]\mathrm{d}F(\omega) - cq.$$

We start by discussing the range of λ. If $\lambda \leq v_m/v_r$, then L is strictly increasing in $\text{E}_\omega[R(\omega)]$ (unless $\lambda = v_m/v_r$) and $\text{E}_\omega[T_1(\omega)]$. If $\lambda \geq 1$, then L is strictly decreasing in $\text{E}_\omega[T_1(\omega)]$ (unless $\lambda = 1$) and $\text{E}_\omega[R(\omega)]$. In both cases, the Lagrangian has no interior stationary point. Therefore, $\lambda \in (v_m/v_r, 1)$, which immediately tells us that (IR_r) binds in any optimal contracts.

By further examining first-order necessary conditions, we find that the Lagrangian has an interior stationary point only when $N = 1$. We name the resulting contract a *buyback contract* and make a formal definition below.

Definition 2. *A contract is a* ***buyback contract*** *if there exist two constants,* $\underline{\omega} \in [0, q]$ *and* $t \geq 0$*, such that:*

(a) for any ω*,*

$$T_1(\omega) = \min(p\omega, t);$$

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

(b) $t \leq p\underline{\omega} + v_r(q - \underline{\omega})$*, with equality if* $\underline{\omega} > 0$*.*

Put simply, Definition 2 comes from taking $n = 1$ and $\overline{\omega}_1 = q$ in Definition 1. The optimality of buyback contracts is formally stated in Proposition 2 and proved in Appendix A.2.

Proposition 2. *A contract implementing* $q > \underline{q}$ *is optimal only if it is a buyback contract with a binding* (IR_r)*.*

Proof. See Appendix A.2. □

To see the intuition, consider increasing both $\underline{\omega}_1$ and $\overline{\omega}_1$ by $\varepsilon > 0$ sufficiently small. According to the recursive relationship (15), this increases t_1 by εp while keeping all other t_ns unchanged. By Definition 1, these increments will reduce return shipments by ε and increase cash repayments by εp on the interval $[t_1/p, \overline{\omega}_1)$, and will increase return shipments by $\varepsilon(p - v_r)/v_r$ on the interval $[\underline{\omega}_1, t_1/p)$. The gross effect on the Lagrangian is, approximately,

$$\varepsilon[p(1-\lambda) + (\lambda v_r - v_m)][F(\overline{\omega}_1) - F(t_1/p)] - \frac{\varepsilon(p - v_r)(\lambda v_r - v_m)}{v_r}[F(t_1/p) - F(\underline{\omega}_1)].$$

Note that t_1/p is a convex combination of $\underline{\omega}_1$ and $\overline{\omega}_1$. Therefore, when F is not too concave, which is ensured by nondecreasing hazard rate, the gross effect will be positive. This pushes $\overline{\omega}_1$ to the rightmost and suggests that the optimal local buyback contract should have $N = 1$ and $\overline{\omega}_1 = q$.

It is interesting to distinguish two types of buyback contracts. With slightly abuse of notations, we write T_1 as a function of R, denoted by $T_1 = T_1(R)$. One can therefore define the *buyback price* $b(R)$ as

$$b(R) = \frac{t - T_1(R)}{R}.$$

Intuitively, $b(R)$ is the menu of per-unit prices offered by the manufacturer should the retailer return R units of unsold inventories. The price is set to ensure that the retailer can repay exactly t by cash after returning R. Depending on the buyback price scheme, a buyback contract may exhibit two different structures.

Constant buyback price. If $\underline{\omega} = 0$, then $T_1(\omega) = p\omega$ and $R(\omega) = (t - p\omega)/v_r$ for any $\omega < t/p$. The buyback price is a constant $b(R) = v_r$ for all R. In this case, the manufacturer's expected utility is

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

where t is determined by a binding (IR_r),

$$\begin{aligned} \mathrm{E}_\omega\, u_r(\omega, \omega, \min(\omega, q)) &= S(q) + cq - t \\ &= W + \underline{u}. \end{aligned} \tag{18}$$

Hence,

$$\mathrm{E}_\omega\, u_m(\omega) = S(q) - \underline{u} - \left(1 - \frac{v_m}{v_r}\right) p(t/p - Q\,(t/p)). \tag{19}$$

Here, $(1 - v_m/v_r)[t - pQ(t/p)]$ is the expected efficiency loss. If the retailer is obligated to make an upfront payment t at date 1, he has to order at least t/p units, which generates an expected revenue $pQ(t/p)$. Thus, $t - pQ(t/p)$ 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.

Variable buyback price. If $\underline{\omega} > 0$, then when $\omega \in [\underline{\omega}, t/p)$, the buyback price is still v_r, but when $\omega < \underline{\omega}$, $T_1(\omega) = p\omega$, $R(\omega) = q - \omega$, so $T_1(R) = p(q - R)$. The buyback price is therefore

$$b(R) = \frac{t - p(q-R)}{R} = p - \frac{pq - t}{R}.$$

In this case, the manufacturer's expected utility is

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

where t is determined by a binding (IR_r),

$$\begin{aligned}\mathrm{E}_\omega\, u_r(\omega, \omega, \min(\omega, q)) &= S(q) + cq - \int_0^{\underline{\omega}} [p\omega + v_r(q-\omega)]\mathrm{d}F(\omega) - \int_{\underline{\omega}}^{+\infty} t\mathrm{d}F(\omega) \\ &= W + \underline{u}. \end{aligned} \tag{20}$$

Hence,

$$\mathrm{E}_\omega\, u_m(\omega) = S(q) - \underline{u} - \left(1 - \frac{v_m}{v_r}\right) \left[p(t/p - Q\,(t/p)) - (p - v_r)(\underline{\omega} - Q(\underline{\omega}))\right]. \tag{21}$$

Comparing (19) and (21), we see that the efficiency loss is now related to $\underline{\omega}$. That is because when $\omega \leq \underline{\omega}$, 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 what is expected in the buyback contract with a constant price.

The cutoff between the case of constant price and that of variable price can be derived from taking $\underline{\omega} \to 0$ (and thus $t \to v_r q$) on the right-hand side of (18) (or equivalently the right-hand side of (20)). That is, let $\overline{q}$ be the solution for

$$S(q) + (c - v_r)q = W + \underline{u}. \tag{22}$$

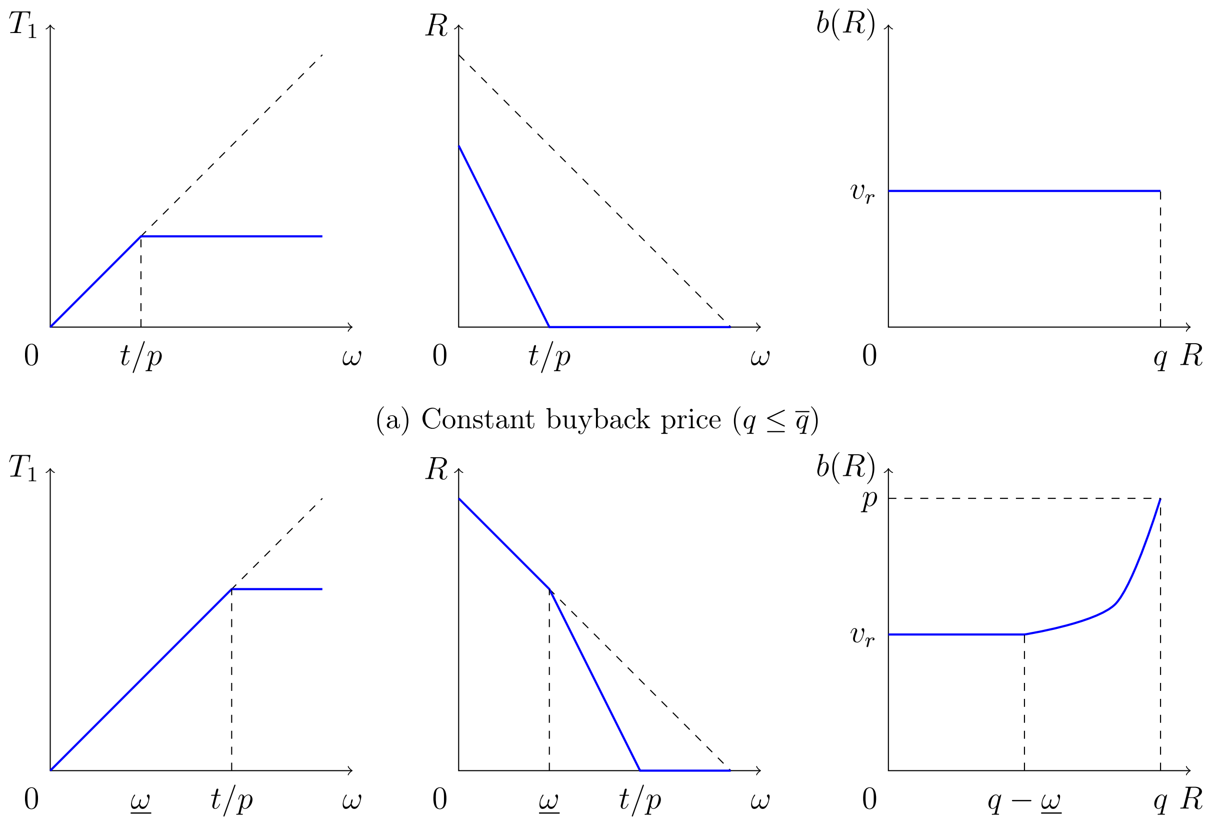
Then, $\underline{\omega} = 0$ if and only if $q \leq \overline{q}$, and $\underline{\omega} > 0$ if and only if $q > \overline{q}$. It is immediate from (6) and (22) that $\underline{q} < \overline{q}$.

The preceding discussion is summarized in Proposition 3 and graphically illustrated in Figure 7.

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

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

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



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

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

Figure 7: Two types of buyback contracts

Remark 2. *Section 4.4 demonstrates another distinction between our constructive approach and other ironing approaches in the literature. Just like us, these approaches all start with taking convex closure of certain components of a contract. We use this construction to further simplify and explicitly characterize the optimal contract under mild assumptions. This makes us different from Myerson (1981) and Baron and Myerson (1982) where the ironed contract is indeed optimal. The main method used in this section is linear programming, which is different from the optimal control approach employed by Guesnerie and Laffont (1984).*

4.5 Optimal order quantity

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 (19). The derivative of $\mathrm{E}_\omega u_m(\omega)$ is

$$\begin{aligned}\frac{\mathrm{d}\,\mathrm{E}_\omega u_m(\omega)}{\mathrm{d}q} &= S'(q) - \left(1 - \frac{v_m}{v_r}\right)[1 - Q'(t/p)]\, t'(q) \\ &= S'(q) - \left(1 - \frac{v_m}{v_r}\right) F(t/p)[v_r + (p - v_r)(1 - F(q))]. \end{aligned} \tag{23}$$

In the second equality, we use $t'(q) = S'(q) + c = v_r + (p - v_r)(1 - F(q))$, which comes from (18). When $q \to \underline{q}$ from the right, we have $t \to 0$, 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.

If the manufacturer offers a buyback contract that implements $q > \overline{q}$, her payoff is determined by (19). The derivative of $\mathrm{E}_\omega u_m(\omega)$ is

$$\begin{aligned}\frac{\mathrm{d}\,\mathrm{E}_\omega u_m(\omega)}{\mathrm{d}q} &= S'(q) - \left(1 - \frac{v_m}{v_r}\right)[v_r + (p - v_r)Q'(\underline{\omega})\underline{\omega}'(q) - Q'(t/p)t'(q)] \\ &= S'(q) - \left(1 - \frac{v_m}{v_r}\right)\left[v_r F(t/p) + (p - v_r)\frac{(F(t/p) - F(\underline{\omega}))(1 - F(q))}{1 - F(\underline{\omega})}\right]. \end{aligned} \tag{24}$$

In the second equality, we use:

$$\underline{\omega}'(q) = \frac{t'(q) - v_r}{p - v_r},$$
$$t'(q) = \frac{S'(q) + c - v_r F(\underline{\omega})}{1 - F(\underline{\omega})} = v_r + (p - v_r)\frac{1 - F(q)}{1 - F(\underline{\omega})},$$

both of which come from (20).

We denote by q^* the optimal quantity that the manufacturer wants to implement using a buyback contract. Then, q^* should solve the first-order necessary condition of the manufacturer's payoff, which is given by either (23) or (24). In both cases, $S'(q^*) > 0$, which implies $q^* < q^{FB}$ from (4).

In a nutshell, when (6) holds, optimal contracts are wholesale contracts with the first-best quantity. When (6) fails, optimal contracts are buyback contracts. There must be an 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 do 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 variable.

It is also interesting to study the comparative statics regarding the retailer's initial cash holding W and reservation utility $\underline{u}$. Note that q^{FB} depends on neither W nor $\underline{u}$. By (6) and (22), an increase in $W + \underline{u}$ leads to a decrease in both $\underline{q}$ and $\overline{q}$. Moreover, when $q \in (\underline{q}, \overline{q}]$, the first-order condition implied by (23) can be reformulated as:

$$\left[1 - \left(1 - \frac{v_m}{v_r}\right) F(t/p)\right] [p - (p - v_r)F(q)] = c.$$

When $W + \underline{u}$ increases, t decreases, q must also increase.

If we interpret $W + \underline{u}$ as the retailer's bargaining power, then Proposition 4 characterizes its relationship with the structure of optimal supply chain contracts. 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 variable pricing. The manufacturer has to increase the buyback price in low-demand states 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 for 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 to alleviate the ex-post adverse selection problem.

5 Discussion

5.1 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 (FT_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 $s = \min(\omega, q)$. (FT_1) becomes

$$T_1(\omega) \leq W - T_0 + p\min(\omega, q) + v_r(q - \min(\omega, q)), \tag{LL_1'}$$

and (IC)-(IC-FT$_1$) becomes

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

Clearly, (SIC$'$) suggests that $T_1(\omega)$ is constant for all ω, and by (LL$_1'$),

$$T_1(\omega) \le W - T_0 + v_r q. \tag{25}$$

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. (25) further implies that the manufacturer's expected utility satisfies

$$\mathrm{E}_\omega\, u_m(\omega) = -cq + T_0 + \mathrm{E}_\omega\, T_1(\omega) \le W - cq + v_r q. \tag{26}$$

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

$$\mathrm{E}_\omega\, u_m(\omega) \le -cq + T_0 + W - T_0 + t = W - cq + t. \tag{27}$$

Comparing (26) and (27), 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.*

5.2 Incentive-compatibility

Another important feature of our model is that the retailer's incentive-compatibility constraint incorporates his limited liability and feasibility constraints.[16] Such model specification actually implies that (FT$_1$) 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

[16] 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.

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 (FT_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

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

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 $\underline{u} < u_0$.

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 (FT_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. 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.[17] For the (IC) presented in Section 2, if (FT_1) binds at ω, the retailer with any $\hat{\omega} < \omega$ cannot misreport ω because

[17] 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.

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 Proposition 2. This proof technique has been applied in the early literature of Costly State Verification (e.g., Gale and Hellwig, 1985), but is much more involved in this paper.

6 Extensions

6.1 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 $Q \geq 0$ 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 price p. Note that we will use subscripts p and q to denote partial derivatives throughout this section. In line with traditional models, such as the state-space model sketched above, we assume that $F(\cdot;p)$ follows the first-order dominance order, i.e., for any p and $\omega > 0$, $F_p(\omega;p) > 0$. This assumption ensures that retail demand is more likely to be realized at a higher level when the price is lower. To avoid unbounded solutions, we also assume that $\lim_{p\to\infty} F(\omega;p) = 1$. That is, when the price is sufficiently high, the retail demand vanishes to zero.

The expected feasible demand and social surplus are as in Section 3,

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

respectively. To avoid excessive technical details, we assume that the social surplus is concave in the space of feasible p and q, and $S_{pq}(q;p) \geq 0$. Assuming that p is observable and contractible, the definition of retail contract must be extended to $\Gamma = (p, q, T_0, s, T_1, R)$.

When information is symmetric, the first-order conditions for maximizing surplus (3) are

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

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

When there is asymmetric information, we maintain the assumption in Section 4.4 that F has a nondecreasing hazard rate, i.e., given any p, $f(\omega;p)/[1 - F(\omega;p)]$ is nondecreasing in ω. Since p is observable and implemented before the realization of demand, it does not change the fact that optimal contracts must be either wholesale or buyback. Following an argument similar to Lemma 1, for any p, let $\underline{q}(p)$ be the largest solution of

$$S(\underline{q}(p);p) + c\underline{q}(p) \leq W + \underline{u}. \tag{30}$$

Intuitively, $\underline{q}(p)$ is the maximum quantity that can be implemented by a wholesale contract under price p. As a result, the first-best quantity q^{FB} can be implemented by a wholesale contract under the first-best price p^{FB} if and only if $q^{FB} \leq \underline{q}(p^{FB})$.

Now we keep p as fixed. To implement any $q > \underline{q}(p)$, the manufacturer has to turn to a buyback contract. In this case, (19) becomes

$$\mathrm{E}_\omega\, u_m(\omega) = S(q;p) - \underline{u} - \left(1 - \frac{v_m}{v_r}\right) p\left[t/p - Q\left(t/p;p\right)\right], \tag{31}$$

and (21) becomes

$$\mathrm{E}_\omega\, u_m(\omega) = S(q;p) - \underline{u} - \left(1 - \frac{v_m}{v_r}\right) \left\{p\left[t/p - Q\left(t/p;p\right)\right] - (p - v_r)[\underline{\omega} - Q(\underline{\omega};p)]\right\}. \tag{32}$$

Assuming that $\mathrm{E}_\omega u_m(\omega)$ given by (31) and (32) is also concave, 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 buyback 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.

6.2 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 8.

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 date-1 residual demand 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).

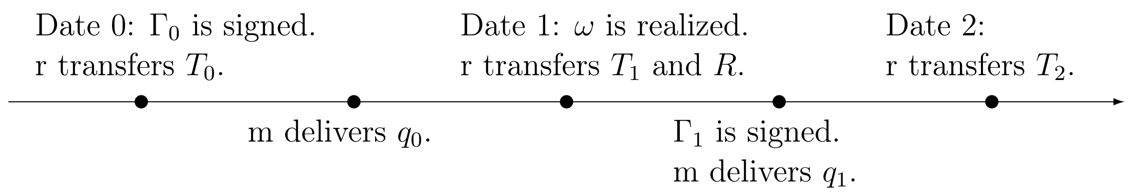


Figure 8: Timeline of the model with reordering.

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 component 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.

6.3 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.[18] 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 $j \in \{1, 2, \dots, n\}$. 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

[18] 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.

T_1^j and the return shipment R^j. The last two components are contingent on retailer j's report $\hat{\omega}^j$.[19] 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 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.

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

7 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, thus 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 variable. 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

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.

Babich, V., H. Li, P. Ritchken, and Y. Wang (2012): "Contracting with asymmetric demand information in supply chains," *European Journal of Operational Research*, 217, 333–341.

Bagnoli, M. and T. Bergstrom (2005): "Log-Concave Probability and its Applications," *Economic Theory*, 26, 445–469.

Baron, D. P. and R. B. Myerson (1982): "Regulating a Monopolist with Unknown Costs," *Econometrica*, 50, 911–930.

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

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.

Davidson, C. and R. Deneckere (1986): "Long-Run Competition in Capacity, Short-Run Competition in Price, and the Cournot Model," *RAND Journal of Economics*, 17, 404–415.

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.

Deneckere, R. and J. Peck (1995): "Competition Over Price and Service Rate When Demand is Stochastic: A Strategic Analysis," *RAND Journal of Economics*, 26, 148–162.

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.

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.

Hsieh, C.-C., C.-H. Wu, and Y.-J. Huang (2008): "Ordering and pricing decisions in a two-echelon supply chain with asymmetric demand information," *European Journal of Operational Research*, 190, 509–525.

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.

Kumar, S. and G. Srinivasan (2007): "Manufacturer's pricing strategies in a single-period framework under price-dependent stochastic demand with asymmetric risk-preference information," *Journal of the Operational Research Society*, 58, 1449–1458.

Lacker, J. M. and J. A. Weinberg (1989): "Optimal Contracts under Costly State Falsification," *Journal of Political Economy*, 97, 1345–1363.

Loertscher, S. and E. V. Muir (2022): "Monopoly Pricing, Optimal Randomization, and Resale," *Journal of Political Economy*, 130, 566–635.

Maggi, G. (1996): "Strategic Trade Policies with Endogenous Mode of Competition," *American Economic Review*, 86, 237–58.

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.

Milgrom, P. and I. Segal (2002): "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70, 583–601.

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.

Myerson, R. B. (1981): "Optimal auction design," *Mathematics of Operations Research*, 6, 58–73.

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

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.

Rockafellar, R. T. (1970): *Convex Analysis*, Princeton University Press.

Rudin, W. (1987): *Real and Complex Analysis*, McGraw-Hill.

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

Shen, B., T.-M. Choi, and S. Minner (2019): "A review on supply chain contracting with information considerations: information updating and information asymmetry," *International Journal of Production Research*, 57, 4898–4936.

Taylor, T. A. and W. Xiao (2009): "Incentives for retailer forecasting: Rebates vs. returns," *Management Science*, 55, 1654–1669.

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.

Yue, X. and S. Raghunathan (2007): "The impacts of the full returns policy on a supply chain with information asymmetry," *European Journal of Operational Research*, 180, 630–647.

Appendix

A.1 The Ironing Approach

Assume that Γ implements $q > \underline{q}$ and satisfies the necessary conditions in Lemma 3 and Lemma 4. Let V be the associated total transfer as defined in (8). We first construct a new function $\hat{V}$ by mapping V into the quantile space. For any $\phi \in [0,1]$, let $\mathcal{V}(\phi)$ be the accumulated total transfer for all types below $F^{-1}(\phi)$.[20] That is,

$$\mathcal{V}(\phi) = \int_0^{F^{-1}(\phi)} V(\omega)\mathrm{d}F(\omega) = \int_0^{\phi} V(F^{-1}(\hat{\phi}))\mathrm{d}\hat{\phi}.$$

By construction, $\mathcal{V}(\phi)$ is increasing, absolutely continuous, and admits a (Radon–Nikodym) derivative $\mathcal{V}'(\phi) = V(F^{-1}(\phi))$. Using $\phi = F(\omega)$, we have $\mathcal{V}'(F(\omega)) = V(\omega)$.

Denote by $\hat{\mathcal{V}}$ the lower convex envelope of $\mathcal{V}$, which is the largest convex function below $\mathcal{V}$, formally defined by

$$\hat{\mathcal{V}} = \sup\{\mathcal{U}|\mathcal{U} \text{ is convex and } \mathcal{U}(\phi) \leq \mathcal{V}(\phi) \text{ for all } \phi \in [0,1]\}.$$

Since $\hat{\mathcal{V}}$ is convex, it is absolutely continuous and admits a nondecreasing (Radon–Nikodym) derivative $\hat{\mathcal{V}}'$. The "ironed total transfer function" is defined as

$$\hat{V}(\omega) = \lim_{\hat{\omega} \to \omega+} \hat{\mathcal{V}}'(F(\hat{\omega})). \tag{A.1}$$

Clearly, $\hat{V}$ is right-continuous and nondecreasing. Furthermore, $\hat{\mathcal{V}}(1) = \mathcal{V}(1)$, so $\hat{V}$ has the same expectation as V with respect to the probability measure given by F. In what follows, we establish several properties of $\hat{V}$ that will allow us to construct an alternative contract that dominates Γ.

Note that by construction, there exists countably many disjoint intervals $[\underline{\phi}_n, \overline{\phi}_n)$, indexed by $n \in \mathcal{N}$, such that $\hat{\mathcal{V}}(\phi)$ is linear on every $[\underline{\phi}_n, \overline{\phi}_n)$ and is strictly convex otherwise. Since by (O2) of Lemma 4, $\mathcal{V}$ is linear on $[F(q), 1]$, there exists one $n \in \mathcal{N}$, say $n(q)$, such that: (1) $[F(q), 1) \subseteq [\underline{\phi}_{n(q)}, \overline{\phi}_{n(q)})$; and (2) for all $n \neq n(q)$, $\overline{\phi}_n \leq F(q)$. By applying the inverse mapping F^{-1}, which maps $[0,1]$ to the extended real line, we define $\underline{\omega}_n = F^{-1}(\underline{\phi}_n)$ and $\overline{\omega}_n = F^{-1}(\overline{\phi}_n)$ for all n. Thus, $\underline{\omega}_{n(q)} \leq q$, $\overline{\omega}_{n(q)} = +\infty$, and for $n \neq n(q)$, $\underline{\omega}_n \leq \overline{\omega}_n \leq q$.

Several observations follow from this construction. First, on each $[\underline{\phi}_n, \overline{\phi}_n)$, $\hat{\mathcal{V}}(\phi)$ is linear, so $\hat{\mathcal{V}}'(\phi)$ is constant, and we denote this constant value by k_n. By (A.1), for any $\omega \in [\underline{\omega}_n, \overline{\omega}_n)$,

[20] Recall that F is an atomless distribution, so $F^{-1} : [0,1] \mapsto [0, +\infty)$ is well-defined.

$\hat{V}(\omega) = k_n$. Since $\mathcal{V}$ is enveloped by $\hat{\mathcal{V}}$ from below,

$$k_n \le V(\underline{\omega}_n) \le p\underline{\omega}_n + v_r(q - \underline{\omega}_n). \tag{A.2}$$

Second, for any $\phi \notin \bigcup_{n\in\mathcal{N}}[\underline{\phi}_n, \overline{\phi}_n)$, $\hat{\mathcal{V}}(\phi) = \mathcal{V}(\phi)$. By convexity, for almost all $\omega \notin \bigcup_{n\in\mathcal{N}}[\underline{\omega}_n, \overline{\omega}_n)$, $\hat{V}(\omega) = V(\omega)$.

Third, for any $\omega \notin \bigcup_{n\in\mathcal{N}}[\underline{\omega}_n, \overline{\omega}_n)$, $R(\omega) = \max(q-\omega, 0)$. Suppose contrary to the assertion that $R(\omega) < q - \omega$ for some $\omega < q$. Let $\hat{\omega}$ be a type satisfying $\omega \le \hat{\omega} \le q - R(\omega)$. Then $T(\omega) \le p\hat{\omega}$ and $R(\omega) \le q - \hat{\omega}$, which, by (O1a) of Lemma 4, implies that $V(\hat{\omega}) \le V(\omega)$. Since the analysis applies to any $\hat{\omega} \in [\omega, q - R(\omega)]$, $\mathcal{V}$ is concave on $[F(\omega), F(q - R(\omega))]$. Passing to the convex envelope of $\mathcal{V}$ shows that $\phi = F(\omega)$ must belong to some $[\underline{\phi}_n, \overline{\phi}_n)$ on which $\hat{\mathcal{V}}$ is linear, a contradiction.

Finally, $\hat{V}$ is continuous. Suppose contrary to the assertion that $\hat{V}(\omega-) < \hat{V}(\omega) = \hat{V}(\omega+)$ at some ω, where $\hat{V}(\omega-)$ and $\hat{V}(\omega+)$ represent the left and right limit of $\hat{V}$ at ω, respectively. Then $\hat{\mathcal{V}}$ has subdifferential $[\hat{V}(\omega-), \hat{V}(\omega+)]$ at $\phi = F(\omega)$, implying that $\hat{\mathcal{V}}$ has a kink at ϕ. Hence, $\mathcal{V}(\phi) = \hat{\mathcal{V}}(\phi)$. Since $\mathcal{V}$ is enveloped by $\hat{\mathcal{V}}$ from below, $V(\omega-) \le \hat{V}(\omega-) < \hat{V}(\omega+) \le V(\omega+)$. However, by (O1b) of Lemma 4, for any $\hat{\omega} < \omega$,

$$V(\omega) \le V(\hat{\omega}) + (p - v_r)[R(\hat{\omega}) - (q - \omega)] \le V(\hat{\omega}) + (p - v_r)(\omega - \hat{\omega}).$$

When $\hat{\omega}$ converges to ω from the left, $V(\omega) \le V(\omega-)$. Similarly, for any $\hat{\omega} > \omega$,

$$V(\hat{\omega}) \le V(\omega) + (p - v_r)(\hat{\omega} - \omega).$$

When $\hat{\omega}$ converges to ω from the right, $V(\omega+) \le V(\omega) \le V(\omega-)$, a contradiction.

Using the function $\hat{V}$ and the set of cutoffs $\{\underline{\omega}_n, \overline{\omega}_n\}$, we are ready to construct an alternative contract $\hat{\Gamma}$, defined as[21]

$$\hat{T}_1(\omega) = \begin{cases} \min(p\omega, t_n) & \omega \in [\underline{\omega}_n, \overline{\omega}_n), \\ V(\omega) - v_r \max(q - \omega, 0) & \omega \notin \bigcup_{n\in\mathcal{N}}[\underline{\omega}_n, \overline{\omega}_n), \end{cases} \tag{A.3}$$

$$\hat{R}(\omega) = \begin{cases} \max(q - \overline{\omega}_n, 0) + \max((t_n - p\omega)/v_r, 0) & \omega \in [\underline{\omega}_n, \overline{\omega}_n), \\ \max(q - \omega, 0) & \omega \notin \bigcup_{n\in\mathcal{N}}[\underline{\omega}_n, \overline{\omega}_n), \end{cases} \tag{A.4}$$

[21] To save notation, we allow the operators min and max to take arguments on the extended reals. That is, for any $z \in \mathbb{R}$, $\min(z, +\infty) = z$ and $\max(z, -\infty) = z$.

where $t_n = k_n - v_r \max(q - \overline{\omega}_n, 0)$. This implies

$$t_n \leq (p - v_r)\underline{\omega}_n + v_r \min(\overline{\omega}_n, q) \text{ for any } n \in \mathcal{N}. \tag{A.5}$$

Clearly, $\hat{\Gamma}$ as constructed from (A.3)-(A.5) is a local buyback contract.

A.2 Proof of Proposition 2

Let $L(x, y)$ be the integral part of the Lagrangian conditional on the interval (x, y), and define $t_0 = 0$. For any $1 \leq n \leq N - 1$,

$$\begin{aligned} L(\underline{\omega}_n, \overline{\omega}_n) = &\int_{\underline{\omega}_n}^{t_n/p} \left[(1-\lambda)p\omega - (\lambda v_r - v_m)\left(q - \overline{\omega}_n + \frac{t_n - p\omega}{v_r}\right)\right] \mathrm{d}F(\omega) \\ &+ \int_{t_n/p}^{\overline{\omega}_n} [(1-\lambda)t_n - (\lambda v_r - v_m)(q - \overline{\omega}_n)] \mathrm{d}F(\omega), \end{aligned}$$

and for $n = N$,

$$L(\underline{\omega}_N, \overline{\omega}_N) = \int_{\underline{\omega}_N}^{t_N/p} \left[\left(1 - \frac{v_m}{v_r}\right) p\omega - \left(\frac{\lambda v_r - v_m}{v_r}\right) t_N\right] \mathrm{d}F(\omega) + \int_{t_N/p}^{\overline{\omega}_N} [(1-\lambda)t_N] \mathrm{d}F(\omega).$$

For any $1 \leq n \leq N + 1$,

$$L(\overline{\omega}_{n-1}, \underline{\omega}_n) = \int_{\overline{\omega}_{n-1}}^{\underline{\omega}_n} [(1-\lambda)(t_{n-1} + p(\omega - \overline{\omega}_{n-1})) - (\lambda v_r - v_m)(q - \omega)] \mathrm{d}F(\omega).$$

First, we compute partial derivatives for each $L(x, y)$, ignoring the relationship between t_ns and all the cutoffs. For any $1 \leq n \leq N - 1$,

$$\begin{aligned} \frac{\partial L(\underline{\omega}_n, \overline{\omega}_n)}{\partial \underline{\omega}_n} &= -\left[(1-\lambda)p\underline{\omega}_n - (\lambda v_r - v_m)\left(q - \overline{\omega}_n + \frac{t_n - p\underline{\omega}_n}{v_r}\right)\right] f(\underline{\omega}_n), \\ \frac{\partial L(\underline{\omega}_n, \overline{\omega}_n)}{\partial \overline{\omega}_n} &= [(1-\lambda)t_n - (\lambda v_r - v_m)(q - \overline{\omega}_n)] f(\overline{\omega}_n) + (\lambda v_r - v_m)[F(\overline{\omega}_n) - F(\underline{\omega}_n)], \\ \frac{\partial L(\underline{\omega}_n, \overline{\omega}_n)}{\partial t_n} &= (1-\lambda)[F(\overline{\omega}_n) - F(t_n/p)] - \frac{\lambda v_r - v_m}{v_r}[F(t_n/p) - F(\underline{\omega}_n)], \\ \frac{\partial L(\overline{\omega}_{n-1}, \underline{\omega}_n)}{\partial \underline{\omega}_n} &= [(1-\lambda)(t_{n-1} + p(\underline{\omega}_n - \overline{\omega}_{n-1})) - (\lambda v_r - v_m)(q - \underline{\omega}_n)] f(\underline{\omega}_n), \\ \frac{\partial L(\overline{\omega}_n, \underline{\omega}_{n+1})}{\partial \overline{\omega}_n} &= -[(1-\lambda)t_n - (\lambda v_r - v_m)(q - \overline{\omega}_n)] f(\overline{\omega}_n) - (1-\lambda)p[F(\underline{\omega}_{n+1}) - F(\overline{\omega}_n)], \\ \frac{\partial L(\overline{\omega}_n, \underline{\omega}_{n+1})}{\partial t_n} &= (1-\lambda)[F(\underline{\omega}_{n+1}) - F(\overline{\omega}_n)], \end{aligned}$$

and for $n = N$,

$$\frac{\partial L(\underline{\omega}_N, \overline{\omega}_N)}{\partial \underline{\omega}_N} = -\left[\left(1 - \frac{v_m}{v_r}\right) p\underline{\omega}_N - \left(\frac{\lambda v_r - v_m}{v_r}\right) t_N\right] f(\underline{\omega}_N),$$
$$\frac{\partial L(\underline{\omega}_N, \overline{\omega}_N)}{\partial t_N} = (1-\lambda)[1 - F(t_N/p)] - \frac{\lambda v_r - v_m}{v_r}[F(t_N/p) - F(\underline{\omega}_N)].$$

Next, partial derivatives for the Lagrangian can also be computed. For $1 \le n \le N-1$,

$$\begin{aligned}
\frac{\partial L}{\partial \underline{\omega}_n} &= \frac{\partial L(\underline{\omega}_n, \overline{\omega}_n)}{\partial \underline{\omega}_n} + \frac{\partial L(\overline{\omega}_{n-1}, \underline{\omega}_n)}{\partial \underline{\omega}_n} \\
&= -\left[(1-\lambda)(p\overline{\omega}_{n-1} - t_{n-1}) + (\lambda v_r - v_m)\left(\overline{\omega}_n - \underline{\omega}_n - \frac{t_n - p\underline{\omega}_n}{v_r}\right)\right] f(\underline{\omega}_n), \\
\frac{\partial L}{\partial \overline{\omega}_n} &= \frac{\partial L(\underline{\omega}_n, \overline{\omega}_n)}{\partial \overline{\omega}_n} + \frac{\partial L(\overline{\omega}_n, \underline{\omega}_{n+1})}{\partial \overline{\omega}_n} \\
&= -(1-\lambda)p[F(\underline{\omega}_{n+1}) - F(\overline{\omega}_n)] + (\lambda v_r - v_m)[F(\overline{\omega}_n) - F(\underline{\omega}_n)], \\
\frac{\partial L}{\partial t_n} &= \frac{\partial L(\underline{\omega}_n, \overline{\omega}_n)}{\partial t_n} + \frac{\partial L(\overline{\omega}_n, \underline{\omega}_{n+1})}{\partial t_n} \\
&= (1-\lambda)[F(\underline{\omega}_{n+1}) - F(t_n/p)] - \frac{\lambda v_r - v_m}{v_r}[F(t_n/p) - F(\underline{\omega}_n)],
\end{aligned}$$

and for $n = N$,

$$\begin{aligned}
\frac{\partial L}{\partial \underline{\omega}_N} &= \frac{\partial L(\underline{\omega}_N, \overline{\omega}_N)}{\partial \underline{\omega}_N} + \frac{\partial L(\overline{\omega}_{N-1}, \underline{\omega}_N)}{\partial \underline{\omega}_N} \\
&= -\left[(1-\lambda)(p\overline{\omega}_{N-1} - t_{N-1}) + (\lambda v_r - v_m)\left(q - \underline{\omega}_N - \frac{t_N - p\underline{\omega}_N}{v_r}\right)\right] f(\underline{\omega}_N), \\
\frac{\partial L}{\partial t_N} &= \frac{\partial L(\underline{\omega}_N, \overline{\omega}_N)}{\partial t_N}.
\end{aligned}$$

Then, we can derive total derivatives by accounting for the recursive relationship (15). If $\underline{\omega}_1 > 0$, then $\partial L/\partial \underline{\omega}_1 = 0$, $t_1 = (p - v_r)\underline{\omega}_1 + v_r\overline{\omega}_1$, and

$$\frac{\mathrm{d}L}{\mathrm{d}\underline{\omega}_1} = (p - v_r)\sum_{n=1}^{N} \frac{\partial L}{\partial t_n};$$
$$\frac{\mathrm{d}L}{\mathrm{d}\overline{\omega}_1} = \frac{\partial L}{\partial \overline{\omega}_1} + p\frac{\partial L}{\partial t_1} - (p - v_r)\sum_{n=1}^{N} \frac{\partial L}{\partial t_n}.$$

Therefore,

$$\frac{\mathrm{d}L}{\mathrm{d}\underline{\omega}_1} + \frac{\mathrm{d}L}{\mathrm{d}\overline{\omega}_1} = [(1-\lambda)p + (\lambda v_r - v_m)][F(\overline{\omega}_n) - F(t_n/p)] \\ - \frac{(p-v_r)(\lambda v_r - v_m)}{v_r}[F(t_n/p) - F(\underline{\omega}_n)]. \tag{A.6}$$

Finally, observe that

$$\frac{\partial L}{\partial q} = -(\lambda v_r - v_m)F(q) + (1-\lambda)p[1-F(q)] - c.$$

Since $\partial L/\partial q = 0$ is necessary for optimality, we have

$$1 - F(q) = \frac{c + \lambda v_r - v_m}{(1-\lambda)p + \lambda v_r - v_m}. \tag{A.7}$$

By nondecreasing hazard rate, for any $\omega \in [t_1/p, \overline{\omega}_1]$,

$$f(\omega) \geq \frac{[1-F(\omega)]f(t_1/p)}{1-F(t_1/p)} \geq \frac{[1-F(q)]f(t_1/p)}{1-F(t_1/p)},$$

which implies that

$$F(\overline{\omega}_1) - F(t_1/p) = \int_{t_1/p}^{\overline{\omega}_1} f(\omega)\mathrm{d}\omega \geq \frac{(\overline{\omega}_1 - t_1/p)[1-F(q)]f(t_1/p)}{1-F(t_1/p)}. \tag{A.8}$$

Similarly, for any $\omega \in [\underline{\omega}_1, t_1/p]$,

$$f(\omega) \leq \frac{f(t_1/p)[1-F(\omega)]}{1-F(t_1/p)} \leq \frac{f(t_1/p)}{1-F(t_1/p)},$$

which implies that

$$F(t_1/p) - F(\underline{\omega}_1) = \int_{\underline{\omega}_1}^{t_1/p} f(\omega)\mathrm{d}\omega \leq \frac{(t_1/p - \underline{\omega}_1)f(t_1/p)}{1-F(t_1/p)}. \tag{A.9}$$

Plugging (A.7), (A.8), and (A.9) into (A.6) yields

$$\frac{\mathrm{d}L}{\mathrm{d}\underline{\omega}_1} + \frac{\mathrm{d}L}{\mathrm{d}\overline{\omega}_1} \geq \frac{c(p-v_r)(\overline{\omega}_1 - \underline{\omega}_1)f(t_1/p)}{p[1-F(t_1/p)]} > 0.$$

That is, either $\mathrm{d}L/\mathrm{d}\underline{\omega}_1 > 0$, which contradicts $\underline{\omega}_1 < \overline{\omega}_1$, or $\mathrm{d}L/\mathrm{d}\overline{\omega}_1 > 0$, which implies $N = 1$ and $\overline{\omega}_1 = q$.

If $\underline{\omega}_1 = 0$, then there must be $\mathrm{d}L/\mathrm{d}\underline{\omega}_1 \le 0$. A similar argument will also give us $\mathrm{d}L/\mathrm{d}\overline{\omega}_1 > 0$.

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 (FT_0), (LL'_1), and (IR_r), where (IR_r) can be simplified as

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

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

$$\frac{\partial L}{\partial q} = -c + \lambda_1 v_r + \mu[p - (p - v_r)F(q)] = 0, \tag{A.10}$$

$$\frac{\partial L}{\partial T_0} = 1 - \lambda_0 - \lambda_1 - \mu = 0, \tag{A.11}$$

$$\frac{\partial L}{\partial T_1} = 1 - \lambda_1 - \mu = 0. \tag{A.12}$$

By (A.11) and (A.12), $\lambda_0 = 0$ and $\lambda_1 + \mu = 1$, but $\mu = 0$ violates (A.10). Thus, there must be $\mu > 0$.

If (LL'_1) is slack, then $\lambda_1 = 0$, $\mu = 1$, (A.10) 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 (IR_r) and a slack (LL_1'), i.e.,

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

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

$$\text{E}_\omega\, u_m(\omega) = W - (c - v_r)q, \tag{A.13}$$

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

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

Note that the right-hand side of (A.13) is bounded above by W, which implies that when $\underline{u} \to 0$ and $W \to 0$, $\text{E}_\omega\, u_m(\omega) \to 0$. However, in the benchmark model, when $\underline{u} = W = 0$, $\text{E}_\omega\, u_m(\omega) = S(q^{FB})$. Hence, there exists a cutoff $\hat{u} \le (p - v_r)Q(q^{FB})$ 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

First, note that $S(q;p)$ becomes negative for sufficiently large p and q, so it is without loss to solve the manufacturer's optimization problem under the assumption that p and q are both bounded. In this case, the optimal (p,q) must be an interior stationary point of $\text{E}_\omega\, u_m(\omega)$. Hence we can use first-order necessary conditions to quantify (p,q).

When $0 < t \le v_r q$, the first-order conditions of (31) imply that:

$$\begin{aligned}
S_q(q;p) &= \left(1 - \frac{v_m}{v_r}\right) F(t/p;p)t_q, \\
S_p(q;p) &= \left(1 - \frac{v_m}{v_r}\right) [F(t/p;p)t_p + (t/p)Q_q(t/p;p) - pQ_p(t/p;p) - Q(t/p;p)].
\end{aligned}$$

By (18), $t_q = S_q(q;p) + c = p - (p - v_r)F(q;p) > 0$, so $S_q(q;p) > 0$. Also, $t_p = S_p(q;p) = Q(q;p) + (p - v_r)Q_p(q;p)$, so

$$\begin{aligned}
S_p(q;p) &= \left(1 - \frac{v_m}{v_r}\right) [F(t/p;p)S_p(q;p) + (t/p)Q_q(t/p;p) - S_p(t/p;p) - v_r Q_p(t/p;p)] \\
&\ge \left(1 - \frac{v_m}{v_r}\right) [F(t/p;p)S_p(q;p) - S_p(q;p) + (t/p)Q_q(t/p;p) - v_r Q_p(t/p;p)].
\end{aligned}$$

where the inequality comes from $S_{pq} \geq 0$. Thus, $S_p(q;p) \geq 0$.

When $t > v_r q$, the first-order conditions of (32) imply that:

$$S_q(q;p) = \left(1 - \frac{v_m}{v_r}\right)[F(t/p;p) - F(\underline{\omega};p)]t_q,$$

$$S_p(q;p) = \left(1 - \frac{v_m}{v_r}\right)[(t/p - t_p)Q_q(t/p;p) - pQ_p(t/p;p) - Q(t/p;p) + (t_p - \underline{\omega})Q_q(\underline{\omega};p) + (p - v_r)Q_p(\underline{\omega};p) + Q(\underline{\omega};p)].$$

By (20),

$$t_q = \frac{S_q(q;p) + c - v_r F(\underline{\omega};p)}{1 - F(\underline{\omega};p)} = v_r + (p - v_r)\frac{1 - F(q;p)}{1 - F(\underline{\omega};p)} > 0,$$

$$t_p = \frac{S_p(q;p) - \int_0^{\underline{\omega}} \omega \mathrm{d}F(\omega;p)}{1 - F(\underline{\omega};p)} = \underline{\omega} + \frac{S_p(q;p) - Q(\underline{\omega};p)}{Q_q(\underline{\omega};p)}.$$

Thus, $S_q(q;p) > 0$, and $S_p(q;p)$ can be further simplified as:

$$S_p(q;p) = \left(1 - \frac{v_m}{v_r}\right)[(t/p - t_p)Q_q(t/p;p) - S_p(t/p;p) - v_r Q_p(t/p;p) + (t_p - \underline{\omega})Q_q(\underline{\omega};p) + S_p(\underline{\omega};p)].$$

Furthermore,

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

which implies

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

Therefore,

$$S_p(q;p) \geq \left(1 - \frac{v_m}{v_r}\right)[(t/p - t_p)Q_q(t/p;p) - v_r Q_p(t/p;p)].$$

Assume that $S_p(q;p) \leq 0$, then $t_p \leq \underline{\omega} \leq t/p$, which again implies $S_p(q;p) \geq 0$, a contradiction. Hence, $S_p(q;p) \geq 0$.

Finally, recall that $S_{pq}(q;p) > 0$, so $S_q(q;p) > 0$ and $S_p(q;p) > 0$ jointly imply that $q^* < q^{FB}$ and $p^* < p^{FB}$.

A.5 Proof of Proposition 7

Part (b) of the proposition is proved by sequential rationality. As discussed in Section 6.2, 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

$$\frac{\partial\,\mathrm{E}_\omega\, u_m(\omega)}{\partial p_0} = \frac{\partial\,\mathrm{E}_\omega\, u_m(\omega|\Gamma_0)}{\partial p_0} + \beta\int_{q_0}^{+\infty}\frac{\partial p_0 f(\omega;p_0)}{\partial p_0}(\omega - q_0)d\omega, \tag{A.14}$$

$$\frac{\partial\,\mathrm{E}_\omega\, u_m(\omega)}{\partial q_0} = \frac{\partial\,\mathrm{E}_\omega\, u_m(\omega|\Gamma_0)}{\partial q_0} - \beta p_0[1 - F(q_0;p_0)]. \tag{A.15}$$

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 $\beta \to 1$, (A.15) implies that $q_0 \to 0$, therefore from (A.14) 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 $0 < t^1 \leq v_r q^1$, the manufacturer's expected payoff is

$$\begin{aligned}
\mathrm{E}_\omega\, u_m(\omega) &= 2\{W - cq^1 + \int_0^{\overline{\omega}^1} [\frac{1}{2}p^1\omega + \frac{v_m}{v_r}(t^1 - \frac{1}{2}p^1\omega)]dF(\omega;p^1) + \int_{\overline{\omega}^1}^{+\infty} t^1 dF(\omega;p^1)\} \\
&= 2(W - cq^1) + \{\int_0^{\overline{\omega}^1} [p^1\omega + \frac{v_m}{v_r}(2t^1 - p^1\omega)]dF(\omega;p^1) + \int_{\overline{\omega}^1}^{+\infty} 2t^1 dF(\omega;p^1)\},
\end{aligned} \tag{A.16}$$

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

$$\begin{aligned}
t^1 &= \int_0^{2q^1} [\frac{1}{2}p^1\omega + v_r(q^1 - \frac{1}{2}\omega)]dF(\omega|p^1) + \int_{2q^1}^{+\infty} p^1q^1 dF(\omega|p^1) - W - \underline{u} \\
&= \frac{1}{2}(p^1 - v_r)Q(p^1, 2q^1) + v_r q^1 - W - \underline{u}.
\end{aligned} \tag{A.17}$$

Comparing (A.16) with (19) and (A.17) with (18), 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.