Discussion Paper Series – CRC TR 224

Discussion Paper No. 292
Project C 03

Optimal Retail Contracts With Return Policies

Ying-Ju Chen [1]
Zhengqing Gui [2]
Ernst-Ludwig von Thadden [3]
Xiaojian Zhao [4]

April 2021

[1] Hong Kong University of Science and Technology. Email: imchen@ust.hk
[2] Economics and Management School, Wuhan University. Email: zq.gui@whu.edu.cn
[3] University of Mannheim and CEPR. Email: vthadden@uni-mannheim.de
[4] Monash University. Email: xjzhao81@gmail.com

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Optimal Retail Contracts with Return Policies*

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

April 2021

Abstract

A central problem in vertical relationships is to minimize 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. Cash payments are bounded above by the retailer's revenue, while the return of unsold inventories is bounded above by the order quantity net of the actual quantity sold. While the majority of the papers in the literature takes the contractual forms as given and investigates the consequences that these contracts may lead to in various contexts, without assuming any functional form of contracts, we show that the optimal contract can be implemented by a buy-back contract: the manufacturer requests an upfront payment from the retailer and buys back the unsold inventories at the retailer's salvage value. The optimality of buy-back contracts is robust to several scenarios including competition between retailers.

Keywords: Retail contracts, return policies, buy-back contracts, incentive problems, limited liability.

JEL Classification: D82, D86, L42, L60.

---

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

†Hong Kong University of Science and Technology. Email: imchen@ust.hk.
‡Economics and Management School, Wuhan University. Email: zq.gui@whu.edu.cn.
§University of Mannheim and CEPR. Email: vthadden@uni-mannheim.de.
¶Monash University. Email: xjzhao81@gmail.com.

1 Introduction

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

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

In this paper, we present a model in which a manufacturer sells its products through a retailer who privately observes the retail demand. We argue that production usually precedes sales, thus the two parties have to write down the terms of the contract, including the quantity of products to be delivered to the retailer before any demand uncertainty is resolved. In fact, if we think of the retailer as a grocery, it must determine the storage level of each product it sells before consumers arrive. This assumption is also standard in the retail contracting literature (e.g., Deneckere, Marvel, and Peck, 1996, 1997; Montez, 2015). After the retail demand is materialized, the retailer makes a report of the demand to the manufacturer, and transfers are executed accordingly. A contract between the manufacturer and the retailer specifies the quantity to be produced and delivered, the retailer's cash payments to the manufacturer after sales, and the allocation of unsold inventories. The last two terms are contingent on the retailer's report. Information asymmetry essentially implies that the contract must be incentive-compatible: The retailer should find it optimal to truthfully report the retail demand after the demand is realized.

Moreover, the retailer is subject to limited liability, implying that the retailer's cash payments are bounded above by the sales revenue. This constraint is considered in the literature studying contracting problems in industrial organization (e.g., Brander and Lewis, 1986) and captures the fundamental feature of small and medium enterprises: They are typically resource-constrained and thus the only collateral that can be pledged is the business value they have created.[1] Also, the retailer's returns cannot exceed the total amount of unsold inventories, which imposes the feasibility constraint on the contract. It is further assumed that returning unsold inventories is inefficient.[2] This assumption, together with the retailer's limited liability, makes the return of unsold inventories a screening device. Consistent with our observation on small groceries, we also postulate that the manufacturer has full bargaining power.

Without assuming any functional form of contracts, very generally, we find that the optimal contract takes a rather simple form: The retailer transfers a fixed amount of cash to the manufacturer when the realized demand is high, and returns part of unsold inventories to the manufacturer when the realized demand is too low for the retailer to pay the fixed amount in full. The rationale for this result is the following: Facing the adverse selection problem, the manufacturer wants to elicit the retailer's private information on the retail demand, so the return of unsold inventory is used as a punishment when the reported demand is low. However, the manufacturer, who is assumed to have full bargaining power, also aims to minimize unsold inventories returned by the retailer, since it leads to efficiency loss. Therefore, return policies will be offered only when the reported demand is sufficiently low. From another perspective, the optimal contract can be implemented by the familiar buy-back contract, in which the manufacturer requests the retailer to make a sales-independent payment, and buys back unsold inventories at the retailer's salvage value. Although we derive buy-back contracts from a state-contingent contractual space, it does not imply that in reality people will actually solve a complex contracting problem. Those who implement buy-back contracts may not know that such contracts are optimal. This "as if" approach are widely used in economic theories. Therefore, our paper can be viewed as a micro-foundation of the commonly used buy-back contracts in the real world.

We further show that the quantity determined in the optimal contract is lower than the first-best level when the cost function is assumed to be linear. That is, supplies are rationed by the manufacturer in the presence of information asymmetry. Intuitively, in an economy with complete information, unsold inventories should best be kept by the retailer. Under information asymmetry, however, return polices serve as an incentive scheme, thus part of unsold inventories should be returned to the manufacturer. The manufacturer's marginal revenue from production is thereby less than that in the complete information economy.

We then extend our model to allow for competition between retailers. When an up-

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

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

stream manufacturer sells her products through several downstream retailers, the manufacturer may want to impose retail price control to maximize her total profits. In presence of information asymmetry, however, it is usually difficult for the manufacturer to observe the retail price or detect retailers' secret discounts, such as free add-ons, coupons or cash-back. Therefore, contracts with price maintenance conditions may not be enforceable in vertical relationships. We study whether the manufacturer can use retail contracts that only specifies cash payments and return polices to perform *de facto* price control. In this environment, we show that buy-back contracts are still optimal. Retail prices are equal to consumers' valuation of the product, implying zero consumers surplus. The total supply is increased compared to a monopoly case, but it is still less than the first-best level. Put differently, competition between retailers may push the market supply towards the efficient level, but is not effective in cutting down the retail price.

Our paper contributes to the growing literature on vertical relationships with asymmetric information and uncertain demand. The vast majority of this literature focuses on vertical restraints (e.g., Winter, 1993; Deneckere, Marvel, and Peck, 1996, 1997; Dana and Spier, 2001; Harstad and Mideksa, 2020), i.e., how (and why) the manufacturer controls the retail price, order quantity, or competition between retailers. In these papers, retail contracts are usually given by two-part tariffs, or revenue-sharing schemes, plus some specific restraints, such as the Resale Price Maintenance. However, a fundamental question of whether theses contractual forms are indeed optimal remains a problem. A related stream of literature studies the newsvendor problem (e.g., Pasternack, 1985; Marvel and Peck, 1995; Krishnan and Winter, 2007; Montez, 2015), in which the manufacturer proposes a contract to induce the retailer choosing the optimal price and inventory. However, in these models the retailer's payments are independent of the realized demand, so the adverse selection problem is assumed away. In a related literature (e.g., Rey and Tirole, 1986; Blair and Lewis, 1994), optimal retail contracts are derived under demand uncertainty, but they do not consider the prescribed newsvendor problem. Wang, Gurnani, and Subramanian (2020) examine the signaling role of buy-back contracts, while they take the buy-back form as exogenously given. The paper perhaps closest to the one presented here is by Arya and Mittendorf (2004), in which the manufacturer uses a return allowance to elicit the retailer's private information on demand. In their model the return policy is characterized by the price offered by the manufacturer, so the retailer's choice is all-or-nothing: Full return if the return allowance is higher than the retail price, zero return otherwise. In the present paper, the contract determines the quantity to be returned, so partial return polices are allowed.

Technically, our model is an ex post screening problem with hidden characteristics. When the type set is a continuum, as in our model, the standard methodology is to use control theory. This approach is pioneered by Guesnerie and Laffont (1984), and further developed by Hellwig (2010) who proposes a unified approach that only requires the compactness of the type set, and allows for mass points. However, the control-theoretic approach cannot be applied in the present paper. In our model, each type of the retailer's set of deviation is bounded by the limited liability and the feasibility constraint, and thus depends on the endogenous contract. Therefore, the retailer's incentive constraint cannot be simplified into a local differential 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. Relatedly, Gui, von Thadden, and Zhao (2019) provide a detailed discussion on how the presence of limited liability affects the analysis of incentive constraint in the financial contracting literature. In particular, overlooking the role of limited liability in specifying the incentive constraint may lead to an over-simplified analysis and sub-optimal contracts. This paper also shows that relaxing the limited liability constraint ex-post may be harmful to firms.

Our paper is part of a more general approach that tries to provide the foundation of observed economics or financial institutions as outcomes of optimal contracting (see, e.g., Nöldeke and Schmidt (1995) in the context of buyer-seller relationships, Aghion and Tirole (1997) in a model of hierarchical authority, or Schmidt (2003) for venture capital arrangements). Our paper follows a similar vein to justify the buy-back contract in the retail contract setting without imposing any functional form assumptions on the contract space. Some researchers have also demonstrated, in various contexts, that simple practical contracts seem to perform well even if they are known to be sub-optimal (e.g., Bower, 1993; Rogerson, 2003; Chu and Sappington, 2007). Unlike these papers, we show that the popular buy-back contract in practice may indeed be the optimal contract form.

The rest of this paper is organized as follows. Section 2 introduces the model setup. Section 3 proves the optimality of buy-back contracts and solves for the optimal order quantity. Section 4 extends our benchmark model to study competitive retailers. Section 5 concludes and discusses some possible extensions. Most of the proofs are relegated to the Appendix.

2 Model

Consider a manufacturer (she) who sells its products through a retailer (he) to meet the retail demand. The retail price and demand are denoted by p and ω, respectively. Assume both contracting parties take p as exogenously given in this section. However, ω is materialized after the manufacturer determines its production quantity q. In other words, production must take place prior to the realization of demand, hence a supply-demand mismatch may arise.

The retailer observes ω freely because of his direct contact with consumers. Nonetheless, the manufacturer can only form priors about the distribution of ω, which we denote by $F(\omega)$. Assume $F(\omega)$ is defined on a bounded interval $[0, \bar{\omega}]$, and admits a continuously differentiable density function $f(\omega)$, with $0 < f(\omega) < +\infty$ for all ω. Therefore, the manufacturer has to rely on the retailer's report to execute any transfer between them. By applying the revelation principle, it is without loss to focus on direct mechanisms in which the retailer simply reports his demand (or "type") $\hat{\omega}$ and the transfer is executed correspondingly.

After observing ω, the retailer determines the quantity of products sold to the market, which is indexed by s. When there is supply shortage, i.e., $q < \omega$, the retailer can only sell up to the quantity q, and the excess demand is lost. When the demand falls short, i.e., $q > \omega$, the retailer can sell up to ω. In what follows, we will use $\omega^+ = \min\{\omega, q\}$ to denote the maximum quantity that the retailer can possibly sell. Thus, $0 \leq s \leq \omega^+$. s is also unobservable to the manufacturer.

Moreover, the retailer is able to salvage unsold inventories at a constant salvage value v_r per unit. If instead, the manufacturer possesses unsold inventories, her per unit salvage value is v_m. As discussed in Section 1, we assume that the retailer's salvage value is higher than the manufacturer's, i.e., $v_m < v_r < p$. Thus, it is more efficient for the retailer to keep unsold inventories.

The manufacturer offers the retailer a contract $\Gamma = (q, T, R)$ that specifies three terms: (1) the quantity q delivered to the retailer; (2) the cash transfer from the retailer to the manufacturer, $T(\hat{\omega})$, after selling her products on the market, where $\hat{\omega}$ is the demand reported by the retailer; and (3) the return shipment of unsold inventory, $R(\hat{\omega})$, again as a function of $\hat{\omega}$.

This contractual form captures many different types of retail contracts in practice.

Example 1 (Wholesale price). In a wholesale price contract, the manufacturer charges the retailer a constant wholesale price p_w per unit purchased. The corresponding transfers and returns are, respectively,

$$T(\omega) = p_w q,$$
$$R(\omega) = 0.$$

Example 2 (Buy-back). In a buy-back contract, the manufacturer charges the wholesale price and pays back the retailer b for each unsold unit. Therefore,

$$T(\omega) = p_w q - bR(\omega),$$
$$R(\omega) = q - s.$$

It also requires $b < p_w$ since the retailer should not profit from left over inventory.

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

$$\begin{aligned} T(\omega) &= p_w q + \phi p s, \\ R(\omega) &= 0, \end{aligned}$$

where ϕ is the manufacturer's share of the retailer's revenue.

All these contracts condition ex-post transfers and actions on the realized demand. If the demand is commonly observable, and the retailer faces no limited liability, all these contracts are enforceable. However, the retail demand can hardly be observed by the manufacturer, which corresponds to our assumption that ω is the retailer's private information. Therefore it is usually difficult to conduct such direct conditioning in practice.

We assume both contracting parties are risk-neutral. Given a contract Γ, the realized demand ω, the retailer's sales decision s and his report of the retail demand $\hat{\omega}$, his ex-post payoff is

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

where $T(\hat{\omega})$ and $R(\hat{\omega})$ are the actual cash transfer and return shipment based on the retailer's report $\hat{\omega}$, ps is the gross revenue from the realized sales, $v_r[q - s - R(\hat{\omega})]$ corresponds to the salvage value from the retailer's inventory on hand. Note that u_r increases with s, so the retailer will optimally choose $s = \omega^+$.

Accordingly, the manufacturer's payoff is

$$u_m(\Gamma, \hat{\omega}) = -c(q) + T(\hat{\omega}) + v_m R(\hat{\omega}),$$

where $c(q)$ is her cost of producing q units of good.

The social welfare $W(q)$ is given by

$$\begin{aligned} W(q) &= u_m(\Gamma, \hat{\omega}) + u_r(\Gamma, \omega, \hat{\omega}, s) \\ &= v_r q + (p - v_r)\omega^+ - (v_r - v_s)R(\hat{\omega}) - c(q), \end{aligned}$$

which is decreasing in $R(\hat{\omega})$ if q is predetermined.

The contracting problem becomes interesting because of the several restrictions faced by the retailer. First, the retailer cannot return more than the amount of unsold inventory he has and cannot re-order after strong demand. This implies the following feasibility constraint:

$$0 \leq R(\omega) \leq q - \omega^+. \tag{FC}$$

Second, the retailer cannot pay the manufacturer more than the amount of cash he has after the selling season. This implies the following limited liability (or liquidity, affordability) constraint:

$$T(\omega) \leq p\omega^+. \tag{LL}$$

We do not consider the salvage value of unsold inventory on the right-hand side of (LL) because in practice liquidating leftover inventory typically takes time. The insights of our analysis will not be truly affected even if we account for the value of unsold inventory when specifying (LL).

In the agency literature, (LL) is standard and widely used as an alternative to the assumption of risk-aversion. It can be interpreted as an extreme form of risk aversion, since it restricts the contract from reacting too strongly to demand fluctuations. In practice, it arises for various reasons such as the retailer's inability to raise money over and above what he has realized from sales, his option to quit the relationship ex-post, or legislation banning exploitative contracts.

Again from the revelation principle, contracts must be able to ensure that the retailer makes a truthful report of ω. This leads to the following incentive-compatibility constraint:

$$\begin{aligned} &u_r(\Gamma, \omega, \omega, \omega^+) \geq u_r(\Gamma, \omega, \hat{\omega}, s), \text{ for any } \omega, \hat{\omega} \text{ and } s \text{ such that} \\ &0 \leq s \leq \omega^+, \ 0 \leq R(\hat{\omega}) \leq q - s \text{ and } T(\hat{\omega}) \leq ps. \end{aligned} \tag{IC}$$

Note that as the type-ω retailer mis-reports to be type $\hat{\omega}$, the transfer and the return shipment change accordingly. We only require each type of retailer has no incentive to choose the contract designed for other types when his wealth, ps, and unsold inventory, $q-s$, can afford. The retailer may also sell less than ω^+ to fulfill the requirements for his deviation. In the presence of (FC) and (LL), our (IC) turns out to be type-dependent; the choice set for each type depends on the contract, which is endogenous. Incorporating (FC) and (LL) into (IC) makes it difficult for us to simplify the global incentive constraint into a local first-order condition, and apply the well-established control-theoretic approach to solve for the optimal contract, as in the mainstream literature of mechanism design with hidden types.[3] We will give a detailed discussion on how we circumvent this problem in Section 3.

Finally, the retailer has a certain level of reservation utility as his outside option, which is denoted by $\underline{u}_r$. Since the contract is offered by the manufacturer, the retailer's expected payoff from the contract must exceed $\underline{u}_r$. This gives rise to the retailer's individual-rationality constraint:

$$E_\omega u_r(\Gamma, \omega, \omega, \omega^+) \geq \underline{u}_r. \tag{IR$_r$}$$

[3]See the vast literature starting from Guesnerie and Laffont (1984).

Similarly, the manufacturer's individual-rationality constraint is given by:

$$E_\omega u_m(\Gamma, \omega, \omega, \omega^+) \geq 0. \tag{IR$_m$}$$

It means that the manufacturer's expected profit from the contract should be able to cover her production cost.

Hence a full statement of the contracting problem is

$$\begin{aligned} \max_{\Gamma} \quad & E_\omega u_m(\Gamma, \omega, \omega, \omega^+), \\ \text{subject to} \quad & \text{(FC), (LL), (IC), (IR}_r\text{), (IR}_m\text{).} \end{aligned}$$

We denote this problem by (P), and say that Γ is *optimal* if it solves (P).

3 Analysis

In this section, we first characterize the optimal contract taking q as given, then calculate the manufacturer's expected payoff as a function of q, and finally derive the optimal quantity. For this purpose, we will sometimes write the contract as a pair (T, R), and write two parties' utility functions simply as $E_\omega u_r(T, R)$ and $E_\omega u_m(T, R)$, correspondingly. Also, we ignore (IR$_m$) at the moment, and will check whether it is satisfied for the optimal contract later on. Our analysis is restricted to $q \leq \bar{\omega}$.

Besides, it is convenient for us to say a contract (T, R) is *admissible* if it satisfies (FC), (LL), and (IC). For any given q, we also say an admissible contract (T, R) is q-*optimal* if it maximizes $E_\omega u_m(T, R)$ subject to (IR$_r$). Moreover, if two admissible contracts differ only in a zero-measure set, we say they are *equivalent* since both contracting parties are indifferent between them.

Finally, let

$$\Phi(\omega) = T(\omega) + v_r R(\omega),$$

which represents the retailer's total payout when he reports ω.

The rest of our analysis will be centered around a special form of contracts. It is defined piecewisely, and resembles a buy-back contract on each of its pieces.

Definition 1. *(T, R) is a* ***piecewise buy-back contract*** *(PBC) if there exists a sequence $\{(\omega_i, t_i)\}_{i=0,1,\ldots,n}$, where $n \in \mathbb{N}^+$, such that:*

(a) $0 = \omega_0 < \omega_1 < \cdots < \omega_n = q$, $t_1 < t_2 < \cdots < t_n$, and $t_i \leq p\omega_i + v_r(q - \omega_i)$;

(b) *For $i = 1$, $\omega \le \omega_1$, and any $i = 2, 3, \dots, n$, $\omega \in (\omega_{i-1}, \omega_i]$,*

$$\begin{cases} T(\omega) = p\omega, \quad R(\omega) = q - \omega, & \text{if } \omega \le \frac{t_i - v_r q}{p - v_r}, \\ T(\omega) = p\omega, \quad R(\omega) = \frac{t_i - p\omega}{v_r}, & \text{if } \frac{t_i - v_r q}{p - v_r} < \omega \le \frac{t_i - v_r(q-\omega_i)}{p}, \\ T(\omega) = t_i - v_r(q - \omega_i), \quad R(\omega) = q - \omega_i, & \text{if } \omega > \frac{t_i - v_r(q-\omega_i)}{p}; \end{cases}$$

For $\omega > q$, $T(\omega) = T(q)$, $R(\omega) = R(q)$.

Depending on the value of t_i, a PBC can have three different shapes on different intervals, which are depicted in Figure 1.

In each of the subfigures of Figure 1, the upward sloping dashed line represents (LL), and the downward sloping dashed line represents (FC). The shape of a PBC on $(\omega_{i-1}, \omega_i]$ is determined by t_i, ω_{i-1} and ω_i. In particular:

- When $t_i < p\omega_{i-1} + v_r(q - \omega_i)$, the contract exhibits Shape I;
- When $p\omega_{i-1} + v_r(q - \omega_i) \le t_i < p\omega_i + v_r(q - \omega_i)$, the contract exhibits Shape II;
- When $p\omega_i + v_r(q - \omega_i) \le t_i$, the contract exhibits Shape III.

Although distinct at first glance, the three shapes share a common feature that has a natural economic interpretation: The retailer is obligated to repay t_i to the manufacturer when $\omega \in (\omega_{i-1}, \omega_i]$ is realized, unless t_i exceeds his total wealth. Moreover, cash has priority in payments, but a minimum $q - \omega_i$ units of unsold inventory must be returned irrespective of t_i.

There are several important properties of a PBC, which can be verified immediately by our definition: For any $\omega \in (\omega_{i-1}, \omega_i]$,

- when $\omega \le \frac{t_i - v_r q}{p - v_r}$, both (LL) and (FC) bind, $\Phi(\omega) = p\omega + v_r(q - \omega)$;
- when $\frac{t_i - v_r q}{p - v_r} < \omega \le \frac{t_i - v_r(q-\omega_i)}{p}$,, only (LL) binds, $\Phi(\omega) = t_i$;
- when $\omega > \frac{t_i - v_r(q-\omega_i)}{p}$, both $T(\omega)$ and $R(\omega)$ are constant, $\Phi(\omega) = t_i$.

We then utilize these properties to provide conditions under which a PBC is admissible.

Proposition 1. *A PBC is admissible if and only if $u_r(T(\omega), R(\omega))$ is nondecreasing in ω.*

The condition stated in Proposition 1 serves to prevent the retailer from under–selling, i.e., it gives the retailer incentive to choose $s = \omega^+$. If he deviates to selling a quantity $\hat{\omega}$ less than ω^+, his utility would be weakly less according to the monotonicity of u_r.

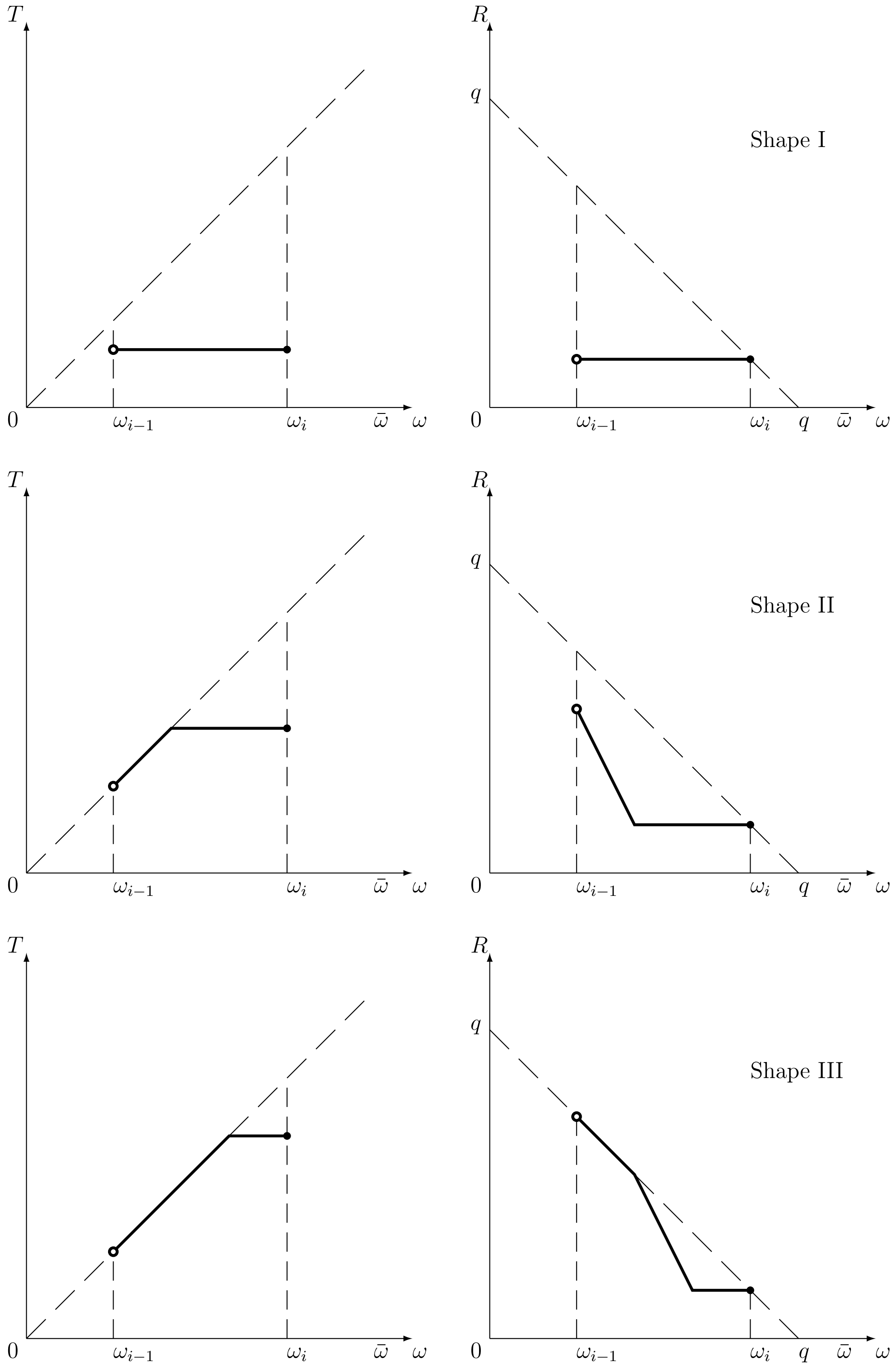

Figure 1: Three possible shapes of a PBC

3.1 Optimal contracts given the order quantity

When the type set is a continuum, as in our model, the standard technique for optimization problems with incentive constraint is control theory (e.g., Hellwig, 2010). The basic idea of this technique is to replace the global incentive constraint with local incentive constraints, thus prove the absolute continuity of the agent's indirect utility function. Hence the agent's indirect utility can be treated as the state variable, and an optimal solution can be obtained using the classical maximum principle.

However, the control-theoretic approach cannot be applied in the present paper. Since the absolute continuity of the agent's indirect utility function is crucial if one attempts to rewrite the contracting problem into an optimal control problem, it essentially requires that the global incentive constraint be substituted by local ones. When the agent is able to mimic any other type irrespective of his own type or the contract, the substitution of incentive constraints, as shown in the adverse selection literature, is valid. Nevertheless, in our model, (IC) interacts with (FC) and (LL), implying that the set of states that the retailer is able to misreport depends on the realized demand and the endogenous contract. Therefore, it is possible that the retailer is only able to mimic a subset of types, or even cannot mimic other type.

As an example, if (IC) is specified as

$$u_r(\Gamma, \omega, \omega, \omega^+) \geq u_r(\Gamma, \omega, \hat{\omega}, s), \text{ for any } \omega, \hat{\omega} \text{ and } s,$$

then one can use local incentive constraints to replace (IC); u_r is thus absolutely continuous.[4] For the (IC) presented in Section 2, if (IC) binds at ω, the retailer with any $\hat{\omega} < \omega$ cannot misreport ω because he cannot afford the cash payment specified in the contract when ω is reported. Hence, the retailer's indirect utility function may have a jump at ω. The possible discontinuities in contracts prevent us from using control theory.

Instead, we apply a step-by-step constructive method to show that any admissible contract is either weakly dominated or approximated by a PBC. Such methodology is standard in the early literature of the Costly State Verification model (e.g., Gale and Hellwig, 1985). We begin by proving that, for any admissible contract, we can find countably many disjoint intervals, on which the retailer's total payout is nonincreasing. Lemma 1 is a prerequisite result for such a conjecture.

Lemma 1. *If (T, R) is admissible, then for any $\omega \leq q$, $\Phi(\omega)$ is nonincreasing on $(\omega, q - R(\omega))$ unless $R(\omega) = q - \omega$.*

[4] Actually, assuming that (IC) holds for any ω and $\hat{\omega}$ that lie 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.

The proof of Lemma 1 is basically an application of (IC). Based on Lemma 1, we let

$$S_{(T,R)} = \bigcup_{\omega\in[0,q]} (\omega, q - R(\omega)), \tag{1}$$

which represents the union of all the open intervals $(\omega, q-R(\omega))$ generated from ω. [5] By definition, $S_{(T,R)}$ is an open set, so there exist countably many nonempty disjoint open intervals, denoted by $\{(a_j, b_j) : j \in J_{(T,R)}\}$, where $J_{(T,R)}$ is a countable index set, such that

$$S_{(T,R)} = \bigcup_{j\in J_{(T,R)}} (a_j, b_j),$$

and for any $j, j' \in J_{(T,R)}$, $j \neq j'$, we have

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

In other words, $\{(a_j, b_j) : j \in J_{(T,R)}\}$ is a partition of $S_{(T,R)}$. This partition is unique, so $J_{(T,R)}$ is uniquely determined by (T, R). Since $\Phi(\omega)$ is nonincreasing on $(\omega, q - R(\omega))$, it must also be nonincreasing on any of these open intervals (a_j, b_j). Moreover, our definition implies that for any j and $\omega \in (a_j, b_j)$, $R(\omega) \geq q - b_j$. This property is similar to that of a PBC, and thus plays a key role when we conjecture a PBC in the proof of Lemma 2 and Lemma 3.

Note that the partition of $S_{(T,R)}$ can be either finite or countably infinite, so one has either $|J_{(T,R)}| < +\infty$ or $|J_{(T,R)}| = +\infty$, correspondingly. We will first discuss the case when $|J_{(T,R)}|$ is finite. In this case any admissible contract is weakly dominated by a PBC.

Lemma 2. *If (T, R) is admissible, and $|J_{(T,R)}| < +\infty$, then there exists a PBC, $(\hat{T}, \hat{R})$, such that:*

(a) $E_\omega u_m(\hat{T}, \hat{R}) \geq E_\omega u_m(T, R)$;

(b) $E_\omega u_r(\hat{T}, \hat{R}) = E_\omega u_r(T, R)$.

Moreover, the inequality in (a) is strict unless $(\hat{T}, \hat{R})$ is a PBC.

The proof of Lemma 2 involves several steps. We first transform the contract on any $(a_j, b_j) \subseteq S_{(T,R)}$ so that it exhibits a buy-back structure, with the retailer's expected total payout conditional on (a_j, b_j) unchanged. This transformation makes the manufacturer better off because the buy-back structure increases the retailer's cash payments given that $\Phi(\omega)$ is nonincreasing on (a_j, b_j). Then the same transformation is performed on any interval that is not a subset of $S_{(T,R)}$. Since (FC) binds at any state that belongs to

[5] S may be empty, but it will not affect the conjecture we make in Lemma 2 and Lemma 3.

the interior of $[0,q]/S_{(T,R)}$, the transformation will never make the manufacturer worse off as long as the retailer's expected payoff is unchanged. After these two steps and some other small adjustments, we will finally get a PBC that outperforms the original contract.

Lemma 3. *If (T,R) is admissible, and $|J_{(T,R)}| = +\infty$, then for any $\varepsilon > 0$, there exists a PBC, $(\hat{T},\hat{R})$, such that:*

(a) $E_\omega u_m(\hat{T},\hat{R}) > E_\omega u_m(T,R) - \varepsilon$*;*

(b) $E_\omega u_r(\hat{T},\hat{R}) = E_\omega u_r(T,R)$*.*

When the partition of $S_{(T,R)}$ is infinite, we cannot expect the intervals in $\{(a_j,b_j) : j \in J_{(T,R)}\}$ are well-ordered, so the conjecture in the proof of Lemma 2 may not be applicable. To circumvent this difficulty, we choose an arbitrarily small $\delta > 0$, and treat all the intervals in the partition of $S_{(T,R)}$ whose lengths are smaller than δ as if they have a binding (FC). Therefore, there will be only finitely many intervals left in the partition of $S_{(T,R)}$, thus we can conjecture a PBC in the same way as shown in the proof of Lemma 2.

Clearly, this conjecture makes the retailer strictly worse off, because the return of unsold inventory is raised on some intervals to make (FC) bind. The manufacturer is strictly better off, because the retailer pays strictly more than the original contract. However, the retailer's utility loss is bounded above by an increasing function of δ. Hence for any given $\varepsilon > 0$, we can always find a δ so that the retailer's utility loss from the newly constructed PBC is smaller than ε. Finally, we compensate his utility loss by reducing t_1 in the new PBC. Lemma 3 is thereby proved.

Proposition 2 directly comes from Lemmas 2 and 3, so we state it without any separate proof.

Proposition 2. *If (T,R) is admissible, then for any $\varepsilon > 0$, there exists a PBC, $(\hat{T},\hat{R})$, such that:*

(a) $E_\omega u_m(\hat{T},\hat{R}) > E_\omega u_m(T,R) - \varepsilon$*; i*

(b) $E_\omega u_r(\hat{T},\hat{R}) = E_\omega u_r(T,R)$*.*

According to Proposition 2, any admissible contract can be either weakly dominated, or approximated by a sequence of PBCs. As a consequence, if there exists a PBC that maximizes $E_\omega u_m(T,R)$ subject to (IR_r) and (IR_m) among all the PBCs, it will outperform all the admissible contracts. If such PBC is admissible, then it is an optimal contract to the manufacturer's problem. Proposition 3 formally confirms the result.

Proposition 3. *If, among all the PBCs, (T,R) is admissible and maximizes $E_\omega u_m(T,R)$ subject to (IR_r), then it is q-optimal.*

Our next step is to find the optimal PBC and see whether it is indeed admissible. Note that, given a PBC, both contracting parties' expected payoffs are pinned down by $\{\omega_i, t_i : i = 0, 1, \dots, n\}$. Thus the standard technique for constrained optimization problems can be applied.

Let L be the Lagrangian of the manufacturer's optimization problem, and λ be the Lagrangian multiplier of (IR$_r$). If (T, R) maximizes $E_\omega u_m(T, R)$ subject to (IR$_r$) and (IR$_m$), $\{\omega_i, t_i : i = 0, 1, \dots, n\}$ should maximize

$$L = E_\omega u_m(T, R) + \lambda[E_\omega u_r(T, R) - \underline{u}_r],$$

subject to (IR$_r$) and (IR$_m$), as well as several boundary constraints:

$$0 = \omega_0 \leq \omega_1 \leq \cdots \leq \omega_n = q, \tag{2}$$

$$t_1 \leq t_2 \leq \cdots \leq t_n, \tag{3}$$

$$t_i \leq p\omega_i + v_r(q - \omega_i) \text{ for any } i = 1, 2, \dots, n, \tag{4}$$

and the complementary slackness constraint:

$$\lambda \geq 0, \quad \lambda[E_\omega u_r(T, R) - \underline{u}_r] = 0. \tag{5}$$

Here we use weak inequalities instead of strict ones in (2), (3), and (4), because we want to allow for the possibility that the optimal PBC has less than n pieces. This happens when some ω_i and ω_{i+1} coincide, or $t_i = p\omega_i + v_r(q - \omega_i)$ for some i.

In fact, the objective function L can be expressed as some constant plus $\hat{L}$, where

$$\begin{aligned}\hat{L} = &\sum_{i=1}^{n} \int_{\omega_{i-1}}^{\omega_i} [T(\omega) + v_m R(\omega)] dF(\omega) + \int_q^{\bar{\omega}} [T(\omega) + v_m R(\omega)] dF(\omega) \\ &- \lambda\{\sum_{i=1}^{n} \int_{\omega_{i-1}}^{\omega_i} [T(\omega) + v_r R(\omega)] dF(\omega) + \int_q^{\bar{\omega}} [T(\omega) + v_r R(\omega)] dF(\omega)\}.\end{aligned}$$

By applying first-order necessary conditions, we characterize the optimal PBC, and show that it is admissible. Therefore by Proposition 3, the optimal PBC is a q-optimal contract.

Proposition 4. *The only PBC that is q-optimal has $n = 1$ and a binding (IR$_r$).*

According to Definition 1, our q-optimal contract (T^*, R^*) is given by

$$\begin{cases} T^*(\omega) = p\omega, \quad R^*(\omega) = q - \omega, & \text{if } \omega < \frac{t^* - v_r q}{p - v_r}, \\ T^*(\omega) = p\omega, \quad R^*(\omega) = \frac{t^* - p\omega}{v_r}, & \text{if } \frac{t^* - v_r q}{p - v_r} \leq \omega < \frac{t^*}{p}, \\ T^*(\omega) = t^*, \quad R^*(\omega) = 0, & \text{if } \omega \geq \frac{t^*}{p}, \end{cases} \tag{6}$$

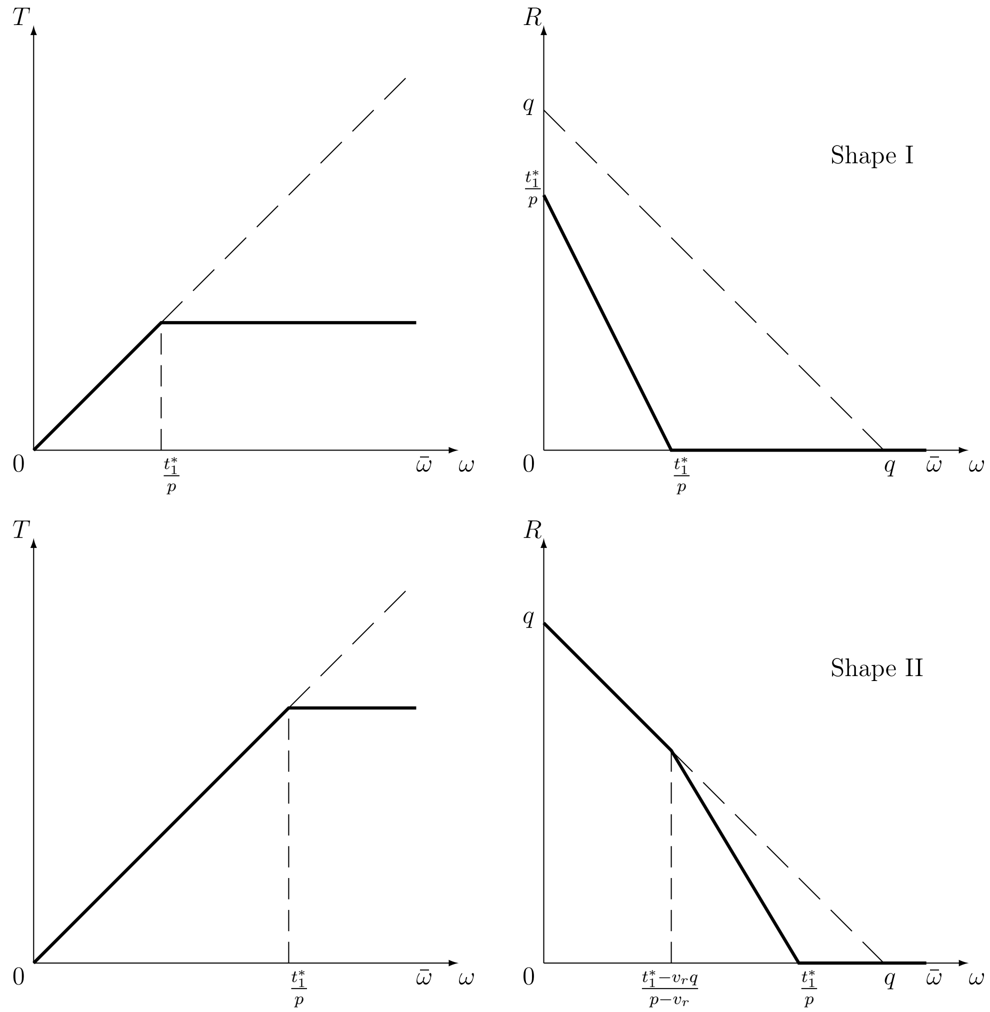

Figure 2: Two possible shapes of (T^*, R^*).

where t_1 is determined by a binding (IR_r). Likewise, (T^*, R^*) has two possible shapes, depending on different values of t^*. Figure 2 is a graphical illustration of (T^*, R^*), which corresponds to Shape II and Shape III in Figure 1 with $n = 1$.

Proposition 4, as well as its mathematical representation (6), has a nice economic interpretation. First, the manufacturer provides the retailer with a fixed quantity q at the wholesale price t_1/q. If he sells more than t_1/p, then t_1 is repaid to the manufacturer and he salvages unsold inventories. If he sells less than t_1/p, some of unsold inventories must be returned to the manufacturer so as to make his total payout equal t_1. This means that the supplier effectively pays him a unit price v_r for his returns. The only exception occurs when the demand is considerably small such that the retailer cannot fulfill the obligation of repaying t_i even if everything he has is transferred to the manufacturer. Hence, our q-optimal contract is indeed a buy-back contract.

Corollary 1. *Our q-optimal contract can be implemented by a buy-back contract with a wholesale price t_1/q and a buy-back price v_r.*

Thus, our analysis provides a foundation of retail contracts. In the supply chain contracting literature, a pre-dominant paradigm is to compare amongst various contracts observed in practice (e.g., Cachon, 2003; Chen, 2003). While these comparisons generate useful managerial implications, a potential caveat is that the contracts considered may be sub–optimal. By taking a different approach, our analysis speaks directly to the question of contract optimality. Remarkably, even though salvaging unsold inventory is more efficient at the retailer, the manufacturer buys back some of them in order to alleviate the ex-post adverse selection problem.

Moreover, Proposition 5 shows that (T^*, R^*) is unique in the sense that any q-optimal contract is equivalent to (T^*, R^*).

Proposition 5. *An admissible contract is q-optimal only if it is equivalent to (T^*, R^*).*

Proposition 5 is proved by contradiction. Suppose that there is another contract (T, R) that gives the two contracting parties the same expected payoffs as (T^*, R^*). By Lemma 2 and Lemma 3, $|J_{(T,R)}| = +\infty$, otherwise (T, R) will be outperformed by a PBC. However, when $|J_{(T,R)}| = +\infty$, we can arbitrarily choose an endpoint ω_i of the open intervals in the partition of $S_{(T,R)}$, and divide $[0, \bar{\omega}]$ into two parts: $[0, \omega_i]$ and $(\omega_i, \bar{\omega}]$. Now we treat $[0, \omega_i]$ as if it is the type set, and conjecture in the same way as in the proof of Lemma 2, Lemma 3, and Proposition 4. Consequently, on the "truncated" type set $[0, \omega_i]$, (T, R) is weakly dominated by a PBC with $n = 1$. The same result applies to $(\omega_i, \bar{\omega}]$. Therefore, on the complete type set $[0, \bar{\omega}]$, (T, R) is weakly dominated by a PBC with $n = 2$, which is a violation of Proposition 4.

3.2 The optimal order quantity

Now we are ready to solve for the optimal quantity q provided by the manufacturer, using the manufacturer's expected payoffs in any q-optimal contract. To simplify the analysis, we impose the following assumption on the cost function.

Assumption 1. *$c(q) = cq$, with $v_r \leq c \leq p$.*

That is, we assume the cost function $c(q)$ is linear in q, with the marginal cost of production c less than the retail price, but higher than the retailer's salvage value of the unsold inventory.

As a benchmark, consider the first-best quantity for the contracting problem q^{FB}; it maximizes the expected social welfare, denoted by $E_\omega W(q)$, with the retailer salvaging any unsold inventory. That is, q^{FB} maximizes

$$E_\omega W(q) = \begin{cases} \displaystyle\int_0^q [p\omega + v_r(q-\omega)]dF(\omega) + \int_q^{\bar{\omega}} pqdF(\omega) - cq & \text{if } q < \bar{\omega}, \\ \displaystyle\int_0^{\bar{\omega}} [p\omega + v_r(q-\omega)]dF(\omega) - cq & \text{if } q \geq \bar{\omega}. \end{cases}$$

Note that

$$\frac{\partial E_\omega W(q)}{\partial q} = \begin{cases} p - c - (p - v_r)F(q) & \text{if } q < \bar{\omega}, \\ v_r - c & \text{if } q \geq \bar{\omega}. \end{cases}$$

Our assumption $v_r \leq c \leq p$ ensures that $E_\omega W(q)$ is quasiconcave, thus there exists a unique solution $q^{FB} \in [0, \bar{\omega}]$, where

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

We further assume the first-best social welfare is able to cover the retailer's reservation utility $\underline{u}_r$, i.e.,

$$E_\omega W(q^{FB}) = \int_0^{q^{FB}} (p - v_r)\omega dF(\omega) > \underline{u}_r. \tag{7}$$

Clearly, if (7) is violated, the set of contract that satisfies both (IR_m) and (IR_r) is either empty or a singleton, and the problem thus becomes straightforward.

In the presence of information asymmetry, the q-optimal contract derived in Proposition 4 has two different shapes, depending on the value of t^* and its relationship with $v_r q$. t^* is again determined by a binding (IR_r). In particular, let $\underline{q}$ be the solution for

$$\int_0^q [p\omega + v_r(q - \omega)]dF(\omega) + \int_q^{\bar{\omega}} pqdF(\omega) = \underline{u}_r; \tag{8}$$

and $\overline{q}$ be the solution for

$$\int_0^q [p\omega + v_r(q - \omega)]dF(\omega) + \int_q^{\bar{\omega}} pqdF(\omega) - v_r q = \underline{u}_r. \tag{9}$$

Then, $t^* > 0$ if and only if $q > \underline{q}$. When $\underline{q} < q < \overline{q}$, $t^* < v_r q$; when $q \geq \overline{q}$, $t^* \geq v_r q$. We will then provide separate analysis for each of these two cases, and use first-order conditions for the manufacturer's utility maximization problem to characterize q^*.

When $t^* < v_r q$, the manufacturer's expected utility is

$$E_\omega u_m(T^*, R^*) = \int_0^{\frac{t^*}{p}} [p\omega + \frac{v_m}{v_r}(t^* - p\omega)]dF(\omega) + \int_{\frac{t^*}{p}}^{\bar{\omega}} t^* dF(\omega) - cq,$$

where t^* is determined by

$$\int_0^q [p\omega + v_r(q - \omega)]dF(\omega) + \int_q^{\bar{\omega}} pqdF(\omega) - t^* = \underline{u}_r. \tag{10}$$

The first-order condition is

$$[1-(1-\frac{v_m}{v_r})F(\frac{t^*}{p})][p-(p-v_r)F(q)]=c. \tag{11}$$

The left-hand side of (11) decreases with q, and when $q \to \underline{q}$, it converges to $p-(p-v_r)F(\underline{q})$. By (8), when $q=\underline{q}$, $t^*=0$, so we must have $\underline{q} \leq q^{FB}$, which implies

$$p-(p-v_r)F(\underline{q}) \geq p-(p-v_r)F(q^{FB})=c.$$

Therefore, (11) has a solution on $[\underline{q},\overline{q}]$ if and only if

$$[1-(1-\frac{v_m}{v_r})F(\frac{v_r\overline{q}}{p})][p-(p-v_r)F(\overline{q})] \leq c.$$

When $t^* > v_r q$, the manufacturer's expected utility is

$$E_\omega u_m(T^*,R^*)=\int_0^{\omega_1'}[p\omega+v_m(q-\omega)]dF(\omega)+\int_{\omega_1'}^{\frac{t^*}{p}}[p\omega+\frac{v_m}{v_r}(t^*-p\omega)]dF(\omega)$$
$$+\int_{\frac{t^*}{p}}^{\bar{\omega}} t^* dF(\omega)-cq,$$

where $\omega_1'=\frac{t^*-v_r q}{p-v_r}$, and t^* is determined by

$$\int_{\omega_1'}^{q}[p\omega+v_r(q-\omega)]dF(\omega)+\int_q^{\bar{\omega}} pq dF(\omega)-t^*[1-F(\omega_1')]=\underline{u}_r. \tag{12}$$

The first-order condition is

$$\{1-F(\omega_1')-(1-\frac{v_m}{v_r})[F(\frac{t^*}{p})-F(\omega_1')]\}\{p-(p-v_r)\frac{F(q)-F(\omega_1')}{1-F(\omega_1')}\}+v_m F(\omega_1')=c. \tag{13}$$

Similarly, the left-hand side of (13) decreases with q, and when $q \to \overline{q}$, it converges to the left-hand side of (11); when $q \to \bar{\omega}$ and $\underline{u}_r \to 0$, it converges to v_m. Therefore, the first-order derivative of $E_\omega u_m(T^*,R^*)$ is continuous at the cutoff $\overline{q}$, and there exists a cutoff u_0 such that when $\underline{u}_r \leq u_0$, first-order conditions (11) and (13) have a unique solution q^*.

We can also compare q^* with q^{FB}. If q^* solves (11), then

$$p-(p-v_r)F(q^*)>c \quad \Rightarrow \quad F(q^*)<\frac{p-c}{p-v_r}.$$

If q^* solves (13), then

$$[1-F(\omega_1')][p-(p-v_r)\frac{F(q)-F(\omega_1')}{1-F(\omega_1')}]+v_rF(\omega_1')>c \quad \Rightarrow \quad F(q^*)<\frac{p-c}{p-v_r}.$$

In both cases, $q^* < q^{FB}$. Our discussion is summarized in Proposition 6.

Proposition 6. *Under Assumption 1, there exists a cutoff u_0 such that when $\underline{u}_r \le u_0$, (11) and (13) have a unique solution $q^* < q^{FB}$.*

For completeness, one has to plug q^* into $E_\omega u_m(T^*, R^*)$ to verify (IR_m). This is straightforward since $q^* \ge \underline{q}$.

4 Competitive Retailers

In this section, we extend our benchmark model to allow for competition between retailers. It is commonly observed 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 as he extract profits from all the retailers contracting with him. However, retailers usually compete with each other and attract customers by cutting down retail prices; this is against the manufacturer's private interest. 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, Marvel, and Peck, 1996; Jullien and Rey, 2007; Asker and Bar-Isaac, 2014) and intensively discussed in legal practice.[6]

While whether RPM is legal under antitrust laws remains controversial, we argue that it is usually very costly for the manufacturer to verify retailers' price-cutting behaviors due to the severe information asymmetry between these two parties. Retailers sometimes offer implicit discounts to attract customers, such as free add-ons, extra services, vouchers, or cash rebates. These are all similar to price-cutting, but are rather difficult to get detected, especially when the manufacturer has no information about the retail demand. Thus, the following questions become important in the competitive retail contracting environment: What is the optimal contract when the manufacturer sells her products through several retailers? Can the manufacturer mitigate competition when there is information asymmetry that makes RPM infeasible? We will address these questions using a simple model with one manufacturer and two retailers.

Consider an environment that is identical to the one described in Section 2 except that now there are two symmetric retailers, indexed by subscript $i \in \{1, 2\}$. The timing

[6] See, e.g., *Leegin Creative Leather Products, Inc. v. PSKS, Inc., dba Kay's Kloset...Kay's Shoes* (`https://www.justice.gov/atr/case/leegin-creative-leather-products-inc-v-psks-inc-dba-kays-klosetkays-shoes`).

of the game is revised as follows. First, the manufacturer offers contracts to two retailers, respectively. The contract specifies the quantity to be delivered q_i, the cash payment function T_i, and the amount of products to be returned R_i. The last two components are contingent on retailer i's report $\hat{\omega}_i$. Then, both retailers decide whether to accept their corresponding contracts. If contracts are accepted, they choose their retail prices p_i simultaneously. Finally, the retail demand ω is realized and allocated according to efficient rationing (Kreps and Scheinkman, 1983). When retailers post the same price, the demand is split evenly. Both retailers can freely observe ω, but the manufacturer only knows its distribution. Hence, the manufacturer maximizes her profits from two contracts subject to all the constraints listed in in Section 2.

It is worth noting that, different from our benchmark model, retail prices are endogenously determined by retailers in this oligopoly market. Therefore, customers' willingness to pay for the product must be bounded above. We assume that customers share the same valuation for the product, which is denoted by v, and $v > c$. A retailer may have incentive to undercut his competitor's price so that customers will first purchase from herself. However, such incentive is limited by his obligation to return unsold inventories to the manufacturer. Since prices are determined before the realization of demand, if a retailer lowers his price to increase sales and revenue, he may be unable to return enough inventories pre-specified in the contract. The manufacturer thus can mitigate competition between downstream retailers through return policies.

In a first-best economy, there is no information asymmetry between the manufacturer and retailers. Hence, the manufacturer can enforce both retailers to choose $p_1 = p_2 = v$. By symmetry, the first-best quantity vector (q_1^{FB}, q_2^{FB}) satisfies $q_1^{FB} = q_2^{FB}$. Similar to our discussion in Section 3.2, q_1^{FB} maximizes

$$E_\omega W(q_1) = \begin{cases} \int_0^{2q_1} [v\omega + v_r(2q_1 - \omega)]dF(\omega) + \int_{2q_1}^{\bar{\omega}} 2vq_1 dF(\omega) - 2cq_1 & \text{if } 2q_1 < \bar{\omega}, \\ \int_0^{\bar{\omega}} [v\omega + v_r(2q_1 - \omega)]dF(\omega) - 2cq_1 & \text{if } 2q_1 \geq \bar{\omega}. \end{cases}$$

Note that

$$\frac{\partial E_\omega W(q_1)}{\partial q_1} = \begin{cases} 2[v - c - (v - v_r)F(2q_1)] & \text{if } q < \bar{\omega}, \\ 2[v_r - c] & \text{if } q \geq \bar{\omega}. \end{cases}$$

Our assumption $v_r \leq c \leq v$ ensures that $E_\omega W(q_1)$ is quasiconcave, thus there exists a unique solution $q_1^{FB} \in [0, \bar{\omega}]$, where

$$F(2q_1^{FB}) = \frac{v - c}{v - v_r}.$$

This expression is the same as the characterization of q^{FB} in Section 3.2 except that now

p is replaced by v. Therefore, the manufacturer earns monopoly profit.

When there is information asymmetry, equilibrium contracts are characterized in Proposition 7.

Proposition 7. *Under Assumption 1, there exists a cutoff u_1 such that when $\underline{u}_r \le u_1$, optimal retail contracts are identical for both retailers. They are buy-back contracts and exhibit Shape II in Figure 2. In particular, for $i = 1, 2$,*

$$\begin{cases} T_i^*(\omega) = \frac{1}{2}v\omega, \quad R_i^*(\omega) = q_i - \frac{1}{2}\omega, & \text{if } \omega < \frac{2(t_i^* - v_r q_i)}{v - v_r}, \\ T_i^*(\omega) = \frac{1}{2}v\omega, \quad R_i^*(\omega) = \frac{t_i^* - \frac{1}{2}v\omega}{v_r}, & \text{if } \frac{t_i^* - v_r q_i}{v - v_r} \le \omega < \frac{t_i^*}{v}, \\ T_i^*(\omega) = t_i^*, \quad R_i^*(\omega) = 0, & \text{if } \omega \ge \frac{t_i^*}{v}, \end{cases} \tag{14}$$

where $t_i^* > v_r q_i$ is determined by a binding (IR_r). Optimal order quantities satisfy $\frac{1}{2}q^* < q_1^* = q_2^* < q_1^{FB}$. Retailers both propose $p_1 = p_2 = v$.

Several remarks can be made on Proposition 7. First, buy-back contracts are robust with competitive retailers. By Proposition 4, the optimality of buy-back contracts is not sensitive to how prices or quantities are determined, as long as they are fixed before the realization of demand.

Second, the total quantity delivered by the manufacturer is larger than the optimal quantity in a single-retailer market, but is still less than the first-best quantity. In other words, introducing a competitive retailer increases the overall supply, while there is still inefficient rationing in the market. This result is in line with the literature starting from Kreps and Scheinkman (1983), which shows that price competition and quantity precommitment yield Cournot outcomes.

Finally, both retailers set their prices equal to consumers' valuation of the product, implying that consumers have no surplus. Due to information asymmetry, the manufacturer cannot observe retail prices, therefore RPM is infeasible in our model. However, the manufacturer can still use retail contracts to incentivize downstream retailers to increase their prices and avoid competition. The key mechanism is the use of return polices. In a Shape-II buy-back contract, the manufacturer may specify a cash payment function so that each retailer can fulfill such payments if he set retail prices to v. Moreover, a retailer has to return all his cash and unsold inventories to the manufacturer when the demand realization is sufficiently low. If any retailer wants to undercut his competitor, he has to sell more than one half of the realized demand to increase revenue, but then he will unable to return the specified amount of unsold products. Therefore, no one has incentive to deviate on price.

Our analysis in this section shows that, even if information asymmetry prevents the manufacturer from implementing price control, she can still maintain a monopoly price in the retail market through retail contracts. Competition between retailers may increase the market supply, but cannot achieve the first-best level. Consumers get zero surplus, which is the same as that of a monopoly.

5 Discussion

In this paper, we show that the optimal retail contract takes the form of a buy-back contract when the retailer privately observes the realized demand and the order quantity is sufficiently high, thereby providing a microeconomic foundation for the buy-back contract. Moreover, when the cost function is linear, the optimal order quantity is shown to be lower than the first-best quantity. While we prove that buy-back contracts are optimal when there are competitive retailers, it is worth noting that the optimality of buy-back contracts is remarkably robust in other dimensions.

Renegotiation-proofness. The retailer has no incentive to renegotiate and will indeed return his inventories as specified, because the manufacturer compensates him fully. The manufacturer, on the other hand, will have an incentive to renegotiate the contract because she pays more for the returns than they are worth to her. But it is the retailer who ships the returns, because he alone knows the realized demand. Thus, the optimal contract is renegotiation-proof.

Initial cash holdings. In our benchmark model, we implicitly assume that the retailer has no cash at hand before the selling season. Relaxing this assumption will make the retailer better off because the probability of returning unsold inventories is reduced. However, the spirit of (IC) is invariant to the retailer's initial cash holdings, and our main finding, i.e., the optimality of buy-back contracts, is not prone to this assumption.

Effort/price-dependent demand. Finally, it can be seen that the buy-back contract keeps being optimal if the retailer can enhance demand by costly marketing efforts, or if the retail price is determined at the contracting stage and the demand distribution is price-dependent.

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

References

Aghion, Philippe and Jean Tirole. 1997. “Formal and Real Authority in Organizations.” *Journal of Political Economy* 105 (1):1–29.

Arya, Anil and Brian Mittendorf. 2004. “Using Return Policies to Elicit Retailer Information.” *RAND Journal of Economics* 35 (3):617–630.

Asker, John and Heski Bar-Isaac. 2014. “Raising Retailers’ Profits: On Vertical Practices and the Exclusion of Rivals.” *American Economic Review* 104 (2):672–686.

Blair, Benjamin F and Tracy R Lewis. 1994. “Optimal Retail Contracts with Asymmetric Information and Moral Hazard.” *RAND Journal of Economics* 25 (2):284–296.

Bower, Anthony G. 1993. “Procurement Policy and Contracting Efficiency.” *International Economic Review* 34 (4):873–901.

Brander, James A and Tracy R Lewis. 1986. “Oligopoly and Financial Structure: The Limited Liability Effect.” *American Economic Review* 76 (5):956–970.

Cachon, Gérard P. 2003. “Supply Chain Coordination with Contracts.” In *Handbooks in Operations Research and Management Science*, vol. 11, edited by Antonius G de Kok and Stephen C Graves. Elsevier, 227–339.

Chen, Fangruo. 2003. “Information Sharing and Supply Chain Coordination.” In *Handbooks in Operations Research and Management Science*, vol. 11, edited by Antonius G de Kok and Stephen C Graves. Elsevier, 341–421.

Chu, Leon Yang and David EM Sappington. 2007. “Simple Cost-Sharing Contracts.” *American Economic Review* 97 (1):419–428.

Dana, James D and Kathryn E Spier. 2001. “Revenue Sharing and Vertical Control in the Video Rental Industry.” *Journal of Industrial Economics* 49 (3):223–245.

Deneckere, Raymond, Howard P Marvel, and James Peck. 1996. “Demand Uncertainty, Inventories, and Resale Price Maintenance.” *Quarterly Journal of Economics* 111 (3):885–913.

———. 1997. “Demand Uncertainty and Price Maintenance: Markdowns as Destructive Competition.” *American Economic Review* 87 (4):619–641.

Gale, Douglas and Martin F Hellwig. 1985. “Incentive-Compatible Debt Contracts: The One-Period Problem.” *Review of Economic Studies* 52 (4):647–663.

Guesnerie, Roger and Jean-Jacques 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 (3):329–369.

Gui, Zhengqing, Ernst-Ludwig von Thadden, and Xiaojian Zhao. 2019. "Incentive-Compatibility, Limited Liability and Costly Liquidation in Financial Contracting." *Games and Economic Behavior* 118:412–433.

Harstad, Bård and Torben Mideksa. 2020. "Coalitions under Vertical Contracts: A Theory for Mergers, Collusion, and Conservation." Working paper.

Hellwig, Martin F. 2000. "Financial Intermediation with Risk Aversion." *Review of Economic Studies* 67 (4):719–742.

———. 2001. "Risk Aversion and Incentive Compatibility with Ex Post Information Asymmetry." *Economic Theory* 18 (2):415–438.

———. 2010. "Incentive Problems with Unidimensional Hidden Characteristics: A Unified Approach." *Econometrica* 78 (4):1201–1237.

Hortaçsu, Ali and Chad Syverson. 2015. "The Ongoing Evolution of US Retail: A Format Tug-of-War." *Journal of Economic Perspectives* 29 (4):89–112.

Jullien, Bruno and Patrick Rey. 2007. "Resale Price Maintenance and Collusion." *RAND Journal of Economics* 38 (4):983–1001.

Kreps, David M and Jose A Scheinkman. 1983. "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes." *Bell Journal of Economics* 14 (2):326–337.

Krishnan, Harish and Ralph A Winter. 2007. "Vertical Control of Price and Inventory." *American Economic Review* 97 (5):1840–1857.

Marvel, Howard P and Stephen McCafferty. 1984. "Resale Price Maintenance and Quality Certification." *RAND Journal of Economics* 15 (3):346–359.

Marvel, Howard P and James Peck. 1995. "Demand Uncertainty and Returns Policies." *International Economic Review* 36 (3):691–714.

Montez, João. 2015. "Controlling Opportunism in Vertical Contracting when Production Precedes Sales." *RAND Journal of Economics* 46 (3):650–670.

Nöldeke, Georg and Klaus M Schmidt. 1995. "Option Contracts and Renegotiation: A Solution to the Hold-up Problem." *RAND Journal of Economics* 26 (2):163–179.

Pasternack, Barry Alan. 1985. "Optimal Pricing and Return Policies for Perishable Commodities." *Marketing Science* 4 (2):166–176.

Rey, Patrick and Jean Tirole. 1986. "The Logic of Vertical Restraints." *American Economic Review* 76 (5):921–939.

Rogerson, William P. 2003. "Simple Menus of Contracts in Cost-Based Procurement and Regulation." *American Economic Review* 93 (3):919–926.

Schmidt, Klaus M. 2003. "Convertible Securities and Venture Capital Finance." *Journal of Finance* 58 (3):1139–1166.

Shaffer, Greg. 1991. "Slotting Allowances and Resale Price Maintenance: A Comparison of Facilitating Practices." *RAND Journal of Economics* 22 (1):120–135.

Townsend, Robert M. 1979. "Optimal Contracts and Competitive Markets with Costly State Verification." *Journal of Economic Theory* 21 (2):265–293.

Wang, Shouqiang, Haresh Gurnani, and Upender Subramanian. 2020. "The Informational Role of Buyback Contracts." *Management Science,* forthcoming.

Winter, Ralph A. 1993. "Vertical Control and Price Versus Nonprice Competition." *Quarterly Journal of Economics* 108 (1):61–76.

A Proofs

A.1 Proof of Proposition 1

The "only if" part is straightforward, so we only need to prove the "if" part. When $\omega \leq q$ For any $i = 1, 2, \ldots, n$, $\omega \in (\omega_{i-1}, \omega_i]$,

$$\omega \leq \frac{t_i - v_r q}{p - v_r} \quad \Rightarrow \quad T(\omega) = p\omega, \quad 0 \leq R(\omega) = q - \omega; \tag{15}$$

$$\frac{t_i - v_r q}{p - v_r} < \omega \leq \frac{t_i - v_r(q - \omega_i)}{p} \quad \Rightarrow \quad T(\omega) = p\omega, \quad 0 \leq q - \omega_i < R(\omega) = \frac{t_i - p\omega}{v_r} \leq q - \omega; \tag{16}$$

$$\omega > \frac{t_i - v_r(q - \omega_i)}{p} \quad \Rightarrow \quad T(\omega) = t_i - v(q - \omega_i) \leq p\omega, \quad 0 \leq R(\omega) = q - \omega_i \leq q - \omega. \tag{17}$$

Thus (FC) and (LL) hold.

To verify (IC), we note that, the retailer has no incentive to sell less than the realized demand because $u_r(T(\omega), R(\omega))$ is nondecreasing. Then, we rule out the retailer's incentive to misreport given that there is no under-selling.

The retailer's total payout for any $i = 1, 2, \ldots, n$ and $\omega \in (\omega_{i-1}, \omega_i]$ is

$$\Phi(\omega) = \begin{cases} p\omega + v_r(q - \omega) & \text{if } p\omega < t_i - v_r(q - \omega), \\ t_i & \text{if } p\omega \geq t_i - v_r(q - \omega), \end{cases}$$

which is nondecreasing on $(\omega_{i-1}, \omega_i]$. Then, consider $\hat{\omega}_1$, $\hat{\omega}_2$ and some $i \in \{1, 2, \ldots, n\}$.

If $\omega_{i-1} < \hat{\omega}_1 < \hat{\omega}_2 \leq \omega_i$, then $\Phi(\hat{\omega}_1) \leq \Phi(\hat{\omega}_2)$ because $\Phi(\omega)$ is nondecreasing on $(\omega_{i-1}, \omega_i]$. We only need to study the case when $\Phi(\hat{\omega}_1) < \Phi(\hat{\omega}_2)$. Note that

$$\begin{aligned} \Phi(\hat{\omega}_1) < \Phi(\hat{\omega}_2) \quad &\Rightarrow \quad \Phi(\hat{\omega}_1) = p\hat{\omega}_1 + v_r(q - \hat{\omega}_1) < \Phi(\hat{\omega}_2) = t_i, \\ &\Rightarrow \quad R(\hat{\omega}_1) = q - \hat{\omega}_1, \end{aligned}$$

which means (FC) binds at $\hat{\omega}_1$. Therefore, a type-$\hat{\omega}_2$ retailer cannot misreport $\hat{\omega}_1$.

If $\omega_{i-2} < \hat{\omega}_1 \leq \omega_{i-1} < \hat{\omega}_2 \leq \omega_i$, then from (15), (16), and (17), we know that $R(\hat{\omega}_1) \geq q - \omega_{i-1} > q - \hat{\omega}_2$, so a type-$\hat{\omega}_2$ retailer cannot mimic $\hat{\omega}_1$. Moreover, we have either

$$\Phi(\hat{\omega}_2) = t_i \quad \Rightarrow \quad \Phi(\hat{\omega}_2) = t_i > t_{i-1} \geq \Phi(\hat{\omega}_1),$$

which means a type-$\hat{\omega}_1$ retailer does not want to misreport $\hat{\omega}_2$; or

$$\Phi(\hat{\omega}_2) = p\hat{\omega}_2 + v_r(q - \hat{\omega}_2) < t_i \quad \Rightarrow \quad T(\hat{\omega}_2) = p\hat{\omega}_2,$$

which means (LL) binds at $\hat{\omega}_2$, i.e., a type-$\hat{\omega}_1$ retailer is unable to misreport $\hat{\omega}_2$.

It is also straightforward to check the admissibility for $\omega = 0$ and $\omega > q$.

A.2 Proof of Lemma 1

Suppose that $R(\omega) < q - \omega$ for some $\omega \leq q$, i.e., $(\omega, q - R(\omega))$ is nonempty. Let $\omega < \hat{\omega}_1 < \hat{\omega}_2 < q - R(\omega)$. Since $q - \hat{\omega}_1 > R(\omega)$, a type-$\hat{\omega}_1$ retailer can misreport ω, implying that $\Phi(\hat{\omega}_1) \leq \Phi(\omega)$. A similar argument also gives us $\Phi(\hat{\omega}_2) \leq \Phi(\omega)$.

If $R(\hat{\omega}_1) \geq R(\omega)$, then from $\Phi(\hat{\omega}_1) \leq \Phi(\omega)$, $T(\hat{\omega}_1) \leq T(\omega)$, which means a type-$\omega$ retailer can misreport $\hat{\omega}_1$ too. From (IC), $\Phi(\hat{\omega}_1) = \Phi(\omega) \geq \Phi(\hat{\omega}_2)$.

If $R(\hat{\omega}_1) < R(\omega)$, then $R(\hat{\omega}_1) < q - \hat{\omega}_2$, which means a type-$\hat{\omega}_2$ retailer can misreport $\hat{\omega}_1$. From (IC), $\Phi(\hat{\omega}_1) \geq \Phi(\hat{\omega}_2)$.

A.3 Proof of Lemma 2

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

Now we construct $(\hat{T}, \hat{R})$ in two steps.

Step 1. Consider an alternative contract $(\tilde{T}, \tilde{R})$, and the retailer's corresponding total payout $\tilde{\Phi}(\omega)$, such that:

(1) For $i = 1$, $\omega \leq \omega_1$, and any $i = 2, 3, \ldots, n$, $\omega \in (\omega_{i-1}, \omega_i]$,

$$\begin{cases} T(\omega) = p\omega, \quad R(\omega) = q - \omega, & \text{if } \omega \leq \frac{t_i - v_r q}{p - v_r}, \\ T(\omega) = p\omega, \quad R(\omega) = \frac{t_i - p\omega}{v_r}, & \text{if } \frac{t_i - v_r q}{p - v_r} < \omega \leq \frac{t_i - v_r(q - \omega_i)}{p}, \\ T(\omega) = t_i - v_r(q - \omega_i), \quad R(\omega) = q - \omega_i, & \text{if } \omega > \frac{t_i - v_r(q - \omega_i)}{p}; \end{cases}$$

For $\omega > q$, $T(\omega) = T(q)$, $R(\omega) = R(q)$.

(2) For $i = 1, 2, \ldots, n - 1$, t_i is determined by

$$\int_{\omega_{i-1}}^{\omega_i} \tilde{\Phi}(\omega) dF(\omega) = \int_{\omega_{i-1}}^{\omega_i} \Phi(\omega) dF(\omega); \tag{18}$$

t_n is determined by

$$\int_{\omega_{n-1}}^{\bar{\omega}} \tilde{\Phi}(\omega)dF(\omega) = \int_{\omega_{n-1}}^{\bar{\omega}} \Phi(\omega)dF(\omega).$$

Step 2. If, in the newly constructed contract $(\tilde{T}, \tilde{R})$, $t_{i-1} \geq t_i$ for some i, then we drop ω_{i-1} from $\{\omega_i\}_{i=0,1,\ldots,n}$, and repeat Step 1 based on $(\tilde{T}, \tilde{R})$.

After finitely many steps, we will get a contract $(\hat{T}, \hat{R})$ and the retailer's corresponding total payout $\hat{\Phi}(\omega)$, which is by definition a PBC. Observe that, compared with (T, R), the retailer's expected total payout conditional on $[0, q]$ is unchanged. So one immediately has (b) of the lemma.

To establish (a) of the lemma, we have to show that the manufacturer is not worse off in both steps. Given the equality in (b), this is equivalent to proving that the retailer pays more cash to the manufacturer in expectation. Mathematically,

$$\begin{array}{l} \int_0^q \tilde{\Phi}(\omega)dF(\omega) = \int_0^{\bar{\omega}} \Phi(\omega)dF(\omega) \\ \int_0^q \hat{T}(\omega)dF(\omega) \geq \int_0^{\bar{\omega}} T(\omega)dF(\omega) \end{array} \quad \Rightarrow \quad E_\omega u_m(\hat{T}, \hat{R}) \geq E_\omega u_m(T, R).$$

For this purpose, we will first prove that, for any $i = 1, 2, \ldots, n-1$,

$$\int_{\omega_{i-1}}^{\omega_i} \tilde{T}(\omega)dF(\omega) \geq \int_{\omega_{i-1}}^{\omega_i} T(\omega)dF(\omega). \tag{19}$$

In particular, there are two cases.

(1) If $(\omega_{i-1}, \omega_i) \subseteq S_{(T,R)}$, then $\tilde{T}(\omega) \leq T(\omega)$ for any $\omega \in (\omega_{i-1}, \omega_i)$. We prove this by contradiction.

Suppose that $p\hat{\omega} \geq T(\hat{\omega}) > \tilde{T}(\hat{\omega})$ for some $\hat{\omega} \in (\omega_{i-1}, \omega_i)$. By the definition of $(\tilde{T}, \tilde{R})$,

$$\begin{aligned} p\hat{\omega} \geq T(\hat{\omega}) > \tilde{T}(\hat{\omega}) \quad &\Rightarrow \quad \tilde{T}(\hat{\omega}) = t_i < p\hat{\omega}, \\ &\Rightarrow \quad T(\hat{\omega}) > t_i. \end{aligned}$$

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

For any $\omega \in (\omega_{i-1}, \hat{\omega})$, Lemma 2 implies that $\Phi(\omega)$ is nonincreasing on (ω_{i-1}, ω_i), so we also have $\Phi(\omega) \geq \Phi(\hat{\omega}) > t_i$.

Thus $\Phi(\omega) > t_i$ for any $\omega \in (\omega_{i-1}, \omega_i)$, which is a violation of (18).

(2) When $(\omega_{i-1}, \omega_i) \nsubseteq S_{(T,R)}$, (ω_{i-1}, ω_i) is a subset of the interior of $[0, q]/S_{(T,R)}$. For any $\omega \in (\omega_{i-1}, \omega_i)$, (FC) must bind, otherwise $(\omega, q - R(\omega)) \subseteq S_{(T,R)}$, a contradiction.

Therefore $R(\omega) \geq \tilde{R}(\omega)$ for any $\omega \in (\omega_{i-1}, \hat{\omega})$. We immediately have (19).

Next we shall prove that

$$\int_{\omega_{n-1}}^{\bar{\omega}} \tilde{T}(\omega)dF(\omega) \geq \int_{\omega_{n-1}}^{\bar{\omega}} T(\omega)dF(\omega). \tag{20}$$

(1) If $(\omega_{n-1}, \omega_n) \subseteq S_{(T,R)}$, then $\Phi(\omega)$ is nonincreasing on (ω_{n-1}, ω_n). Moreover, the firm at state q should have no incentive to choose any s smaller than q. Thus, for any $\omega < q$,

$$\begin{aligned} p\omega + v_r(q-\omega) - \Phi(\omega) &\leq pq - \Phi(q), \\ \Rightarrow \quad \Phi(q) - \Phi(\omega) &\leq (p - v_r)(q-\omega). \end{aligned}$$

The right-hand side converges to zero as ω converges to q, which implies

$$\Phi(q) \leq \lim_{\omega \to q-} \Phi(\omega).$$

Hence, $\Phi(q) \leq \Phi(\omega)$ for any $\omega < q$. An argument similar to the previous one gives us (20).

(2) If $(\omega_{n-1}, \omega_n) \nsubseteq S_{(T,R)}$, then (FC) binds on $(\omega_{n-1}, \bar{\omega})$. We immediately have (20).

Finally, we show that

$$\int_0^q \hat{T}(\omega)dF(\omega) \geq \int_0^q \tilde{T}(\omega)dF(\omega). \tag{21}$$

Without loss of generality, assume that in $(\tilde{T}, \tilde{R})$, $t_{i-1} \geq t_i$ for some $i \in \{2, 3, \ldots, n\}$; moreover, after we drop ω_{i-1} and repeat step 1, $(\hat{T}, \hat{R})$ is finalized.

According to Step 2, on $(\omega_{i-2}, \omega_i]$, $(\hat{T}, \hat{R})$ is determined by

$$\begin{cases} \hat{T}(\omega) = p\omega, \quad \hat{R}(\omega) = q - \omega, & \text{if } \omega < \frac{\hat{t}_i - v_r q}{p - v_r}, \\ \hat{T}(\omega) = p\omega, \quad \hat{R}(\omega) = \frac{\hat{t}_i - p\omega}{v_r}, & \text{if } \frac{\hat{t}_i - v_r q}{p - v_r} \leq \omega < \frac{\hat{t}_i - v_r(q-\omega_i)}{p}, \\ \hat{T}(\omega) = \hat{t}_i - v(q - \omega_i), \quad \hat{R}(\omega) = q - \omega_i, & \text{if } \omega \geq \frac{\hat{t}_i - v_r(q-\omega_i)}{p}; \end{cases}$$

where $\hat{t}_i$ is determined by

$$\int_{\omega_{i-2}}^{\omega_i} \hat{\Phi}(\omega)dF(\omega) = \int_{\omega_{i-2}}^{\omega_i} \tilde{\Phi}(\omega)dF(\omega).$$

The left-hand side is the retailer's expected total payout conditional on $(\omega_{i-2}, \omega_i]$ given $(\tilde{T}, \tilde{R})$, which is nondecreasing in t_{i-1} and t_i. Therefore, $t_{i-1} \geq t_i$ implies $\hat{t}_i \geq t_i$. By construction we immediately have (21).

(19)–(21), along with the equality established in (b), imply (a). It can also be verified that the inequality in (a) is strict unless $(\hat{T}, \hat{R}) = (T, R)$. Hence our proof is complete.

A.4 Proof of Lemma 3

When $|J_{(T,R)}| = +\infty$, $\inf\{b_j - a_j : j \in J_{(T,R)}\} = 0$. For any $\delta > 0$ sufficiently small, the set $\{(a_j, b_j) : b_j - a_j \geq \delta, j \in J_{(T,R)}\}$ is nonempty, finite, and thus well-ordered. Following the similar methodology as what is applied in the proof of Lemma 2, we use $\{\omega_i : i = 0, 1, \ldots, n\}$ to represent $\{a_j, b_j : b_j - a_j \geq \delta, j \in J_{(T,R)}\} \cup \{0, q\}$, and assume that $0 = \omega_0 < \omega_1 < \cdots < \omega_n = q$. Moreover, let

$$\tilde{S} = \bigcup_{b_j - a_j \geq \delta, j \in J_{(T,R)}} (a_j, b_j).$$

The conjecture is also proceeded in two steps.

Step 1. Consider an alternative contract $(\tilde{T}, \tilde{R})$, and the retailer's corresponding total payout $\tilde{\Phi}(\omega)$, such that:

(1) For $i = 1$, $\omega \leq \omega_1$, and any $i = 2, 3, \ldots, n$, $\omega \in (\omega_{i-1}, \omega_i]$,

$$\begin{cases} T(\omega) = p\omega, \quad R(\omega) = q - \omega, & \text{if } \omega \leq \frac{t_i - v_r q}{p - v_r}, \\ T(\omega) = p\omega, \quad R(\omega) = \frac{t_i - p\omega}{v_r}, & \text{if } \frac{t_i - v_r q}{p - v_r} < \omega \leq \frac{t_i - v_r(q - \omega_i)}{p}, \\ T(\omega) = t_i - v_r(q - \omega_i), \quad R(\omega) = q - \omega_i, & \text{if } \omega > \frac{t_i - v_r(q - \omega_i)}{p}; \end{cases}$$

For $\omega > q$, $T(\omega) = T(q)$, $R(\omega) = R(q)$.

(2) For $i = 1, 2, \ldots, n-1$, t_i is determined by

$$\begin{cases} \int_{\omega_{i-1}}^{\omega_i} \tilde{\Phi}(\omega) dF(\omega) = \int_{\omega_{i-1}}^{\omega_i} \Phi(\omega) dF(\omega) & \text{if } (\omega_{i-1}, \omega_i) \subseteq \tilde{S}, \\ \int_{\omega_{i-1}}^{\omega_i} \tilde{\Phi}(\omega) dF(\omega) = \int_{\omega_{i-1}}^{\omega_i} [T(\omega) + v_r(q - \omega)] dF(\omega) & \text{if } (\omega_{i-1}, \omega_i) \not\subseteq \tilde{S}; \end{cases}$$

t_n is determined by

$$\begin{cases} \int_{\omega_{n-1}}^{\bar{\omega}} \tilde{\Phi}(\omega) dF(\omega) = \int_{\omega_{n-1}}^{\bar{\omega}} \Phi(\omega) dF(\omega) & \text{if } (\omega_{n-1}, \omega_n) \subseteq \tilde{S}, \\ \int_{\omega_{n-1}}^{\bar{\omega}} \tilde{\Phi}(\omega) dF(\omega) = \int_{\omega_{n-1}}^{\bar{\omega}} [T(\omega) + v_r(q - \omega)] dF(\omega) & \text{if } (\omega_{n-1}, \omega_n) \not\subseteq \tilde{S}. \end{cases}$$

Step 2. If in the newly constructed contract $(\tilde{T}, \tilde{R})$, $t_{i-1} \geq t_i$ for some i, then we drop ω_{i-1} from $\{\omega_i : i = 0, 1, \ldots, n\}$, and repeat Step 1 based on $(\tilde{T}, \tilde{R})$.

After finitely many steps, we will get a contract $(\hat{T}, \hat{R})$ and the retailer's corresponding total payout $\hat{\Phi}(\omega)$, which is by definition a PBC. Recall that by our construction $\tilde{S} \subset S_{(T,R)}$, so $[0, q]/\tilde{S}$ contains some small intervals that actually belongs to $S_{(T,R)}$. These intervals are treated as if they have a binding (FC) in the conjecture of $(\tilde{T}, \tilde{R})$. Such a transformation serves to ensure that we can apply the same argument in the proof of Lemma 2 to establish (a), i.e., the manufacturer is better off in Step 1. Actually, since the retailer's total payout is increased in $(\tilde{T}, \tilde{R})$, the manufacturer is strictly better off. Similarly, one can prove that Step 2 will not make the manufacturer worse off, so

$$E_\omega u_m(\hat{T}, \hat{R}) \geq E_\omega u_m(T, R). \tag{22}$$

Although the retailer is worse off in $(\tilde{T}, \tilde{R})$, its utility loss is bounded above. For any $\omega \in (a_j, b_j)$ such that $b_j - a_j < \delta$, we have

$$R(\omega) \geq q - b_j \quad \Rightarrow \quad q - \omega - R(\omega) < b_j - a_j < \delta.$$

Therefore

$$\int_0^{\bar{\omega}} \hat{\Phi}(\omega) dF(\omega) - \int_0^{\bar{\omega}} \Phi(\omega) dF(\omega) \leq \int_{[0,q]/\tilde{S}} v_r[q - \omega - R(\omega)] dF(\omega) < v_r \delta,$$

which is equivalent to

$$E_\omega u_r(\hat{T}, \hat{R}) > E_\omega u_r(T, R) - v_r \delta. \tag{23}$$

As a final step, we reduce t_1 in $(\hat{T}, \hat{R})$ by a small amount to compensate the retailer's utility loss. Afterwards (22) and (23) become

$$E_\omega u_m(\hat{T}, \hat{R}) > E_\omega u_m(T, R) - v_r \delta,$$
$$E_\omega u_r(\hat{T}, \hat{R}) = E_\omega u_r(T, R).$$

Note that δ can be arbitrarily small, so for any $\varepsilon > 0$, choosing $\delta = \varepsilon / v_r$ will give us a contract that satisfies both (a) and (b).

A.5 Proof of Proposition 3

We prove the result by contradiction. Suppose that (T, R) maximizes $E_\omega u_m(T, R)$ subject to (IR_r) among all the PBCs, but is not q-optimal. By definition, there exists an admissible contract $(\tilde{T}, \tilde{R})$, which satisfies (IR_r) and

$$E_\omega u_m(\tilde{T}, \tilde{R}) > E_\omega u_m(T, R).$$

Let

$$\varepsilon = \frac{1}{2}[E_\omega u_m(\tilde{T}, \tilde{R}) - E_\omega u_m(T, R)] > 0.$$

According to Proposition 1, there exists a PBC, $(\hat{T}, \hat{R})$, such that

$$E_\omega u_m(\hat{T}, \hat{R}) > E_\omega u_m(\tilde{T}, \tilde{R}) - \varepsilon > E_\omega u_m(T, R),$$
$$E_\omega u_r(\hat{T}, \hat{R}) = E_\omega u_r(\tilde{T}, \tilde{R}),$$

which means $(\hat{T}, \hat{R})$ outperforms (T, R) subject to (IR_r), a contradiction. Therefore (T, R) must be q-optimal.

A.6 Proof of Proposition 4

We start from discussing the range of λ. Given a PBC, (T, R), both contracting parties' expected payoffs on (ω_{i-1}, ω_i) is determined by ω_{i-1}, ω_i, and t_i. In what follows, we restrict our attention to cases with $i = 1, 2, \ldots, n-1$. For $i = n$, all the discussion is the same except that $F(\omega_i)$ should be replaced by 1. Specifically, there are three cases.

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

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

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

$$\begin{aligned}
\omega < \omega_i^1 \quad &\Rightarrow \quad T(\omega) = p\omega, \quad R(\omega) = \frac{t_i - p\omega}{v_r}, \\
\omega \geq \omega_i^1 \quad &\Rightarrow \quad T(\omega) = t_i - v_r(q - \omega_i), \quad R(\omega) = q - \omega_i.
\end{aligned}$$

Therefore,

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

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

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

Therefore,

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

It should also be noted that $\hat{L}$ is continuous at the two nondifferentiable points: $t_i = p\omega_{i-1} + v_r(q - \omega_i)$ and $t_i = p\omega_{i-1} + v_r(q - \omega_{i-1})$. Lemma A.1 states the range of λ, which directly comes from the first-order necessary conditions. We can immediately infer from Lemma A.1 that (IR_r) binds in any optimal PBC.

Lemma A.1. *$\lambda \in (v_m/v_r, 1)$, and for any $i = 1, 2, \ldots, n$, $t_i \geq p\omega_{i-1} + v_r(q - \omega_i)$.*

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

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

If $\lambda = 1$, $\hat{L}$ is negatively correlated with the retailer's expected return of inventory. The contract that maximizes $\hat{L}$ must have $R(\omega) = 0$, $T(\omega) \leq 0$ for any ω, which is a violation of (IR_m).

If $\lambda > 1$, $\partial \hat{L}/\partial t_i < 0$ for any t_i, a contradiction since t_1 has no lower bound. □

Now we are ready to solve for the q-optimal contract based on Lemma A.1. For any $i = 1, 2, \ldots, n-1$, let

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

denote the retailer's Lagrangian conditional on $\omega \in (\omega_{i-1}, \omega_i)$. Then we have

$$\frac{\partial \hat{L}}{\partial \omega_i} = \frac{\partial L_i}{\partial \omega_i} + \frac{\partial L_{i+1}}{\partial \omega_i}.$$

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

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

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

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

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

Therefore,

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

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

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

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

Thus $t_i \leq p\omega_i + v_r(q-\omega_i)$ binds with equality, which essentially means $\omega_i = \omega_{i-1}$.

In sum, (T, R) is an optimal PBC only if $n = 1$. A binding (IR_r) is immediate from Lemma A.1. By Proposition 1, any PBC with $n = 1$ is admissible, and thus is q-optimal according to Proposition 3. Such PBC is unique because first-order conditions are all necessary.

A.7 Proof of Proposition 5

Suppose that (T, R) is q-optimal. As in (1), we denote the union of all the open intervals $(\omega, q - R(\omega))$ by $S_{(T,R)}$, and the partition of $S_{(T,R)}$ by $J_{(T,R)}$.

If $|J_{(T,R)}| < +\infty$, then from Lemma 2, (T, R) must be a PBC, otherwise it will be strictly dominated by a PBC. Again from Proposition 4, (T, R) must be equivalent to (T^*, R^*).

If $|J_{(T,R)}| = +\infty$, then we pick one of the endpoints of the open intervals in the partition of $S_{(T,R)}$, which is denoted by ω_i. Assume $\omega_i < q$. Now we treat $[0, \omega_i]$ as if it is the type set, and apply exactly the same method as in the proof of Lemma 2, Lemma 3, and Proposition 4. As a result, we can conclude that on the "truncated" type set $[0, \omega_i]$, (T, R) is weakly dominated by a PBC with only one piece, which is also defined on

$[0, \omega_i]$. Similarly, on the other half of the type set $(\omega_i, \bar{\omega}]$, (T, R) is, for the same reason, weakly dominated by a PBC with only one piece defined on $(\omega_i, \bar{\omega}]$. Therefore, on the complete type set $[0, \bar{\omega}]$, (T, R) is weakly outperformed by a PBC with $n = 2$. However, by Proposition 4, a PBC with $n = 2$ can never be q-optimal, a contradiction.

A.8 Proof of Proposition 7

We first show that, (p_1, p_2) constitutes an equilibrium price vector only if $p_1 = p_2$. Suppose that $p_1 < p_2$. Then, Retailer 1 can raise his price as long as it does not exceed p_2 so that customers will still buy from Retailer 1 first. Retailer 1 thus may increase his revenue without changing the quantity sold. Hence $p_1 < p_2$ cannot happen in an equilibrium.

Therefore, in any equilibrium the two retailers share the market demand. Anticipating this, the manufacturer will offer identical buy-back contracts to both retailers as they are optimal for any given order quantity. Moreover, The contracts will specify cash payments that can be satisfied only if both retailers set the highest possible prices, i.e., $p_1 = p_2 = v$. By symmetry, $q_1 = q_2$.

Let $\underline{q}_1$ be the solution for

$$\int_0^{2q} [\frac{1}{2} v\omega + v_r(q - \frac{1}{2}\omega)] dF(\omega) + \int_{2q}^{\bar{\omega}} vq dF(\omega) = \underline{u}_r; \tag{24}$$

and $\overline{q}_1$ be the solution for

$$\int_0^{2q} [\frac{1}{2} v\omega + v_r(q - \frac{1}{2}\omega)] dF(\omega) + \int_{2q}^{\bar{\omega}} vq dF(\omega) - v_r q = \underline{u}_r. \tag{25}$$

Then, $t_1^* > 0$ if and only if $q_1 > \underline{q}_1$. When $\underline{q}_1 < q_1 < \overline{q}_1$, $t_1^* < v_r q_1$; when $q_1 \geq \overline{q}_1$, $t_1^* \geq v_r q_1$. However, contracts with $t_1^* < v_r q_1$ cannot be sustained in the equilibrium, as retailers may secretly undercut their prices to compete with each other. Therefore, we have to find conditions that can ensure $q_1 \geq \overline{q}_1$.

When $t_1^* < v_r q_1$ the manufacturer's expected utility is

$$E_\omega u_m(T_1^*, R_1^*) = 2 \int_0^{\frac{2t_1^*}{v}} [\frac{1}{2} v\omega + \frac{v_m}{v_r}(t_1^* - \frac{1}{2} v\omega)] dF(\omega) + 2 \int_{\frac{2t_1^*}{v}}^{\bar{\omega}} t_1^* dF(\omega) - 2cq_1,$$

where t_1^* is determined by

$$\int_0^{2q_1} [\frac{1}{2} v\omega + v_r(q_1 - \frac{1}{2}\omega)] dF(\omega) + \int_{2q_1}^{\bar{\omega}} vq_1 dF(\omega) - t_1^* = \underline{u}_r.$$

The first-order condition is

$$[1-(1-\frac{v_m}{v_r})F(\frac{2t_1^*}{v})][v-(v-v_r)F(2q_1)]=c. \tag{26}$$

Note that the left-hand side of (26) decreases with q_1. Then, when $q_1 \to \underline{q}_1$, it converges to $v-(v-v_r)F(2\underline{q}_1)$. By (25), when $q_1=\underline{q}_1$, $t_1^*=0$, so we must have $\underline{q}_1 \leq q_1^{FB}$, which implies

$$v-(v-v_r)F(2\underline{q}_1) \geq v-(v-v_r)F(2q_1^{FB})=c.$$

Therefore, (26) has a solution on $[\underline{q}_1, \overline{q}_1]$ if and only if

$$[1-(1-\frac{v_m}{v_r})F(\frac{2v_r\overline{q}_1}{v})][v-(v-v_r)F(2\overline{q}_1)] \leq c. \tag{27}$$

When $t_1^* > v_r q_1$, the manufacturer's expected utility is

$$\begin{aligned}E_\omega u_m(T_1^*, R_1^*) = {} & 2\int_0^{\omega_1''}[\frac{1}{2}v\omega+v_m(q_1-\frac{1}{2}\omega)]dF(\omega)+2\int_{\omega_1''}^{\frac{2t_1^*}{v}}[\frac{1}{2}v\omega+\frac{v_m}{v_r}(t_1^*-\frac{1}{2}v\omega)]dF(\omega)\\ & +2\int_{\frac{2t_1^*}{v}}^{\bar{\omega}} t_1^* dF(\omega)-2cq_1,\end{aligned}$$

where $\omega_1''=\frac{2(t_1^*-v_r q_1)}{v-v_r}$, and t_1^* is determined by

$$\int_{\omega_1''}^{2q_1}[\frac{1}{2}v\omega+v_r(q_1-\frac{1}{2}\omega)]dF(\omega)+\int_{2q_1}^{\bar{\omega}} vq_1 dF(\omega)-t_1^*[1-F(\omega_1'')]=\underline{u}_r. \tag{28}$$

The first-order condition is

$$\{1-F(\omega_1'')-(1-\frac{v_m}{v_r})[F(\frac{2t_1^*}{v})-F(\omega_1'')]\}\{v-(v-v_r)\frac{F(2q_1)-F(\omega_1'')}{1-F(\omega_1'')}\}+v_mF(\omega_1'')=c. \tag{29}$$

Similarly, the left-hand side of (29) decreases with q_1, and when $q_1 \to \overline{q}_1$, it converges to the left-hand side of (26); when $q_1 \to \bar{\omega}$ and $\underline{u}_r \to 0$, it converges to v_m. Therefore, the first-order derivative of $E_\omega u_m(T_1^*, R_1^*)$ is continuous at the cutoff $\overline{q}_1$, and when $\underline{u}_r$ is small enough, first-order conditions (26) and (29) have a unique solution q_1^*. Moreover, since the left-hand side of (27) decreases with $\overline{q}_1$, we shall have $q_1^* \geq \overline{q}_1$ if $\underline{u}_r$ is sufficiently small. It is then straightforward to show that $q_1^* < q_1^{FB}$.

We will then compare q_1^* with q^*. Let $p=v$ in (12) and (13). Suppose that $2q_1^* \leq q^*$.

Then by comparing (12) with (28), we know that $2t_1^* < t^*$, and $\omega_1'' < \omega_1'$. Since

$$v - (v - v_r)\frac{F(2q_1^*) - F(\omega_1'')}{1 - F(\omega_1'')} > v_r,$$

the left-hand side of (29) is larger than the left-hand side of (13), a contradiction. Therefore, $2q_1^* > q^*$.