Vertical Relational Contracts and Trade Credit

Marta Troya-Martinez[*]

University of Oxford

October 19, 2012

Abstract

This paper explores the consequences of using supplier trade credit within a vertical relational contract. The downstream firm operates in an environment where shocks may make it unable to repay. The shocks are unobservable to the supplier, which creates an asymmetric information problem.

Trade credit limits the supplier's possibilities to punish the downstream firm and termination is used in equilibrium. Surprisingly, we find that the supplier always sells too little despite having enough instruments to fix the double marginalization problem. The downward distortion in the quantity results from the need to make the contract self-enforced and/or to tackle the asymmetric information problem.

Furthermore, we show that the optimal contract resembles a simple debt contract: if the fixed repayment is met, the contract continues to the next period. Otherwise, the manufacturer asks for the highest possible repayment and terminates for a number of periods. The toughness of the termination policy decreases with the repayment.

JEL Classification: C73, D82, L14.

Keywords: Relational contracts, trade credit, imperfect monitoring.

1   Introduction

Trade credit plays an important role in market supply chains. See for example, Giannetti, Burkart and Ellingsen (2011). Vertically-related firms find it costly to conduct all their business-to-business transactions on pure cash-and-carry basis and the capacity to enter into contracts with delayed obligations is essential for a good business environment. Trade credit is rarely secured on collateral. Moreover, enforcing repayment through the courts can be problematic. The legal cost may be too high relative to the size of the transaction or the buyer may have been adversely affected by a shock leaving nothing to the supplier to foreclose on. This raises the issue of what determines when trade credit is paid back.

In this paper, we study supplier trade credit within a vertical relational contract in an environment where adverse shocks may make the downstream firm unable to honor the credit agreement. Since these shocks are not observable to the supplier, trade credit, by postponing the payment until the shock is realized, creates a source of asymmetric information. To induce repayment, the supplier ensures that it is worthwhile for the downstream firm to repay the credit rather than face retaliation (the simplest form being the refusal to transact further).

We identify a new mechanism that makes the supplier distort the quantity of the good downwards despite having enough instruments to set the final quantity (such as a quantity forcing contract). The reasons behind this distortion are twofold. First, suppose the shocks are observable. The surplus the supplier needs to leave to the downstream firm for repayment to be honored increases with the quantity supplied. Thus, the supplier does not internalize the positive externality that a larger quantity has on the downstream firm and the resulting quantity is too small. It is as if the supplier was facing an additional marginal cost as a result of the relational contract. Second, suppose the shocks are not observable but, even in a low revenue state, the downstream firm does not want to walk away from the contract. To tackle the asymmetric information problem, the supplier has to leave enough surplus in the high revenue states so that the true state is reported. We find that a tougher punishment policy is accompanied by a smaller quantity distortion.

This is because a tougher punishment can be used, instead of giving away surplus, to provide incentives to report the truth.

When the supplier needs to address the enforceability of the contract and the asymmetric information problems at the same time, a tension emerges between them. Increasing the repayment in low revenue states or toughening the punishment following small repayments decreases the incentives for the downstream firm to underreport revenues but at the same time make it less worthwhile continuing the relationship. As a result, the quantity distortion is even larger.

It has been extensively documented, both in developed and developing countries, that firms make deals with each other and get finance using ongoing relationships and trade credit. For example, trade credit accounts for about 15 percent of the assets of US manufacturing firms (Daripa and Nilsen (2011)). Similarly, the existence of relational contracts has been documented, for instance, by Bernstein (1992 and 1996) for the New York diamond trade and the US grain markets and by Uchida *et al.* (2006) for the Japanese small and mid-sized enterprises. In developing countries, where the absence of formal institutions forces firms to use informal substitutes instead, the evidence is even more prominent. For instance, McMillan and Woodruff (1999a and 1999b), Johnson, McMillan and Woodruff (2002) and Fafchamps (1997 and 2000) documents how firms in Vietnam, post-comunist and African countries use relational contracts to substitute for missing laws of contract and trade credit to make up for a lack of access to financial markets. Finally, relational contracts and trade credit are also expected to play an important role when the transactions are not entirely legal. For example, when firms operate in the shadow economy or in black markets, such as drug trade.

The model uses an agency setting where an upstream firm supplies a good and offers trade credit to a cashless downstream firm. For instance, the upstream firm ("she") can be a manufacturer and the downstream firm ("he") a retailer. The manufacturer's machinery is used as collateral making her less credit constrained than the retailer[1].

---

[1]High credit quality suppliers have comparative advantage in securing outside finance that they can pass on small, credit-constrained buyers (Boissay and Gropp (2007)).

3

The manufacturer proposes a quantity forcing contract of the form of a quantity and a repayment[2]. The retailer sells the good and pays back to the manufacturer. However, a shock may occur before the payment, making him unable to repay either part or the whole amount. The manufacturer cannot observe if the failure of payment is due to the shock or to the retailer stealing the money and she punishes him by terminating for a number of periods[3].

This paper belongs to the literature on inter-firm relational contracts with asymmetric information. Levin (2003) is the first paper to introduce moral hazard or adverse selection in a principal-agent relational contract to determine how self-enforceability reduces the provision of high-powered incentives. Trade credit bundled with limited liability imposes important restrictions on the contract that the literature initiated by Levin considers[4]. In Levin's model, the principal rewards the agent's costly action with an ex-ante fixed fee plus an ex-post discretionary positive or negative bonus. Loosely speaking, trade credit is an up-front payment to the agent. Unlike a fixed fee, it is equal to some uncertain revenues that depends on the quantity and hence the total surplus. The repayment, unlike a bonus, is always positive[5]. Most importantly, because of the limited liability, the manufacturer cannot punish the agent in monetary terms. Instead, incentives to repay are provided by the threat of termination (i.e. value burning). Since termination occurs in equilibrium, the outcome is bounded away from efficiency and the downward quantity distortion persists even as the firms become arbitrarily patient.

We consider a problem of moral hazard with hidden knowledge, where the retailer's actions are observable (whether he repaid or not) but not the information on which they

---

[2]Since the quantity is delivered before any private information becomes known to the retailer, it is not used for sorting purposes. As a result, using a two-part tariff or a more complex nonlinear scheme is equivalent to offering a quantity forcing contract. We choose this contract to eliminate the vertical externalities coming from the retailer's market power.

[3]The punishment we consider is similar to how credit reference agencies operate: they "simplify the information about each agent i with a credit report showing when the agent last 'cheated' (e.g., paid late or not at all). This information is kept on the agent's record for a set number of years T, after which time it is erased." Fafchamps (2010), page 57.

[4]See Malcomson (2010) for a recent survey.

[5]In terms of Levin (2003), the bonus is always negative - it is paid by the agent when the outcome is high (i.e. there is no shock) rather than when it is low.

are based (whether he received a shock and of which size).[6] This asymmetric information problem is reminiscent of the model of Green and Porter (1984) where an oligopoly is colluding in a market with noisy prices. When a low price is observed, firms do not know with certainty if this bad outcome is due to a market shock or a firm deviating to a larger quantity. The implications of this different information structure is that the quantity of the good is delivered to the retailer before the private information is learnt and hence it is not used as a sorting instrument.[7]

This paper is also closely related to the literature where an entrepreneur is wealth constrained and is financed by an investor, who cannot observe the investment's cashflows. Hence, the entrepreneur can potentially divert or steal them. Incentives to repay are given by liquidating the entrepreneur's assets (Hart and Moore (1998)), threatening to withhold the investment in the second (and last) period (Bolton and Scharfstein (1990))[8] or carrying on an audit as in the costly state verification models (Townsend (1979) and Gale and Hellwig (1985)).

As in this literature, we also find that the optimal contract is a debt contract[9]. The manufacturer asks for a fixed repayment that if met guarantees the continuation of the contract. Otherwise, the manufacturer asks for the highest possible repayment and punishes for a number of periods. To provide incentives to repay the right amount, a smaller repayment is associated with a larger termination period. A debt contract is optimal because it extract everything from the retailer in the default states allowing the manufacturer to soften the termination policy and to minimize the probability of default. The

---

[6]This problem is also known as post-contractual adverse selection, where the type of the agent (that is, the size of the shock received, which determines how much of the trade credit he is able to give back) becomes known to the downstream firm after having signed the contract. Furthermore, there is no sorting condition as repaying the credit is equally costly for all types.

[7]Buehler and Gartner (forthcoming) also explore the use of relational contracts within two vertically related firms, however their question is very different. They show, how in a vertical relationship where the manufacturer has private information about the manufacturing cost, recommending a retail price may be necessary in order to maximize total profits.

[8]Faure-Grimaud (2000) and Povel and Raith (2004a, 2004b) extend the basic Bolton and Scharfstein two-state model to the case where the cashflows are distributed continuously on a bounded interval.

[9]Innes (1990) finds that debt contracts are optimal in an environment with moral hazard with limited liability. A debt contract gives the best incentives as it makes the agent residual claimant in the good times and penalises him in the bad times by extracting all the surplus.

probability of liquidation (not refinancing or inspection) play the same role as the length of termination imposed by the manufacturer. First, it relaxes the incentive compatibility of the retailer. Second, like auditing or liquidating, imposing the termination policy is costly for the manufacturer, because the relation is always profitable (and hence socially valuable).

The more closely related paper to ours is Povel and Raith (2004a)[10]. The authors allow for the size of the investment to be secretly chosen by the entrepreneur as well as how much to repay and the probability of liquidating. Their focus is on showing that the optimal contract is still a debt contract even when the investment choice is not observed by the investor. They also find that the entrepreneur under-invests as compared to the first best as this decreases the fixed repayment and hence the likelihood of defaulting which may result in an inefficient liquidation[11]. The manufacturer of our model (the "investor") also chooses to sell less output than the first best to the retailer (i.e. the "entrepreneur"). However, because the manufacturer has the bargaining power, the choice is not only motivated to soften the (inefficient) termination policy but also to increase her share of the profits. As a result, the quantity distortion is even larger. The main fundamental difference is to introduce a relational contract in the analysis. Formally, we endogenize the future value that accrues to the retailer if he does not default, as it corresponds to the potential profits generated within the relationship. This change is important in this setup to capture the relational aspect of the relationship and it allows us to explore the impact of the future on the contract. We find that the quantity distortion persists even when there is symmetric information as a result of the enforceability problem.

The paper proceeds as follows. Section 2 introduces the model. Section 3 explores a simple example where an unlucky retailer may lose part of his revenues before paying

---

[10]Faure-Grimaud (2000) and Povel and Raith (2004b) focus on the effect that financial constraints have on the choice of output when the firm is competing a la Cournot with another (financially unconstrained) firm.

[11]In Green and Porter (1984), when quantities are chosen from a sufficiently fine grid of points, a similar result emerges. In particular, firms "produce quantities larger than the monopoly output to reduce the incentives to deviate from equilibrium play, which in turn allows equilibrium punishments to be less severe. Because punishments actually occur in equilibrium, this reduced severity is valuable". (Mailath and Samuelson (2006), p. 353).

to the manufacturer. Section 4 generalizes the previous example to the case where the retailer faces a continuum of demand states. Finally, Section 5 concludes.

2 The Setup

A manufacturer and a retailer have the opportunity to trade at dates $t = 0, 1, 2, ...$ In each period, the manufacturer produces a good at a marginal cost $c > 0$ and needs a retailer to market the product to the final consumer. The retailer can sell the good at no cost but he is completely credit constrained and needs to be fully financed by the manufacturer in order to sell. As a result, the manufacturer offers trade credit to the retailer, who will then pay the manufacturer back after selling the good but within the same period (so no interest rate is charged). To keep the analysis simple, we assume that the retailer is not able to save (i.e. any profits are consumed within the same period).

In order to remove distortions coming from the manufacturer not having enough instruments to determine the final quantity, we let the manufacturer offer[12] a quantity forcing contract. Since the retailer is credit constrained, if he does not repay, there are not many instruments available to the manufacturer to punish for his misbehavior. We let the manufacturer use the threat of a T period termination as a mean to provide incentives.[13] An alternative interpretation to the termination is to trade in less profitable terms (for instance by diminishing the quality of the good).[14]

We denote by $0 < \delta < 1$ the discount factor and we assume that in the periods of no trade, both firms get a constant outside option which is normalized to 0.

The timing is summarized in Figure 1. In each period, the manufacturer offers a contract to the retailer. The retailer rejects or accepts and if he accepts, he places an

---

[12]We assume that the manufacturer has all the bargaining power. We discuss in footnote 20 the implications of this assumption.

[13]McMillan and Woodruff (1999a and 1999b) find that such retaliation occurs in Vietnam, although it is not as forceful as one would expect in a standard repeated game framework. Fafchamps (2004) finds that only 48% of a sample of Sub-Saharan African manufacturers continue to trade following a (late or non) payment dispute.

[14]See Baker, Gibbons and Murphy (1994) for an example in the employer-employee relationship framework.
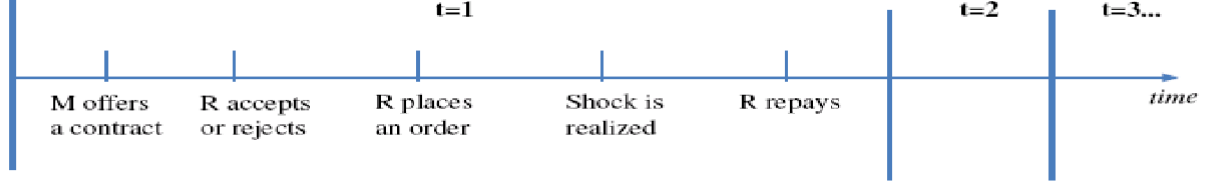
7

Figure 1: Timing of the game

order in the market. Then an *iid* shock is realized (and this information is only observed by the retailer) which determines the size of the revenues. Finally, the retailer decides how much to repay and the contract is terminated for a number of periods if it is specified in the contract.

The effect of the shock is to add randomness to the revenues of the retailer. We can interpret this shock as an uncertainty with the respect to the willingness to pay of final consumers. For instance, the retailer could be placing the quantity offered by the manufacturer in the market and some periods he is paid a high price for it and other periods a low price. Another interpretation is the one of an adverse shock whereby demand is certain but either the goods or the revenues are stolen now and then (for instance, by an organized crime group). Finally, the setup could also represent a situation where the uncertainty refers to how many units of a non-perishable (where it is not possible to give back the unsold units)[15] or perishable good are demanded every period in the market.

3   Example

As an example, we consider the case where the retailer receives revenues $R(q)$ from selling the good. However with probability s, there is a shock and the revenues are $lR(q)$ instead, where $0 \leq l < 1$.[16] The manufacturer offers a quantity forcing contract $\{q, D_H, D_L, T\}$

---

[15]If the good was non-perishable and the retailer could give back the unsold units, the problem would become trivial as the source of asymmetric information disappears.

[16]We use this multiplicative functional form for the revenues because of its simplicity, but the results are robust to using a general function. This functional form is, for instance, used in Green and Porter (1984).

where q is the quantity, D_H is the repayment if demand is high and D_L if it is low. In order to give incentives to repay the true amount, the manufacturer terminates the contract for T periods following D_L and forever following no payment[17].

In a benchmark situation with no asymmetric information and law enforcement, the parties maximize their joint expected one-period profits: $(1 - s(1 - l)) R(q) - cq$. The resulting first best quantity is then determined by:

$$q^{FB} : R'(q) = \frac{c}{1 - s(1 - l)} = \widetilde{c}$$

where $\widetilde{c}$ is the effective marginal cost, which accounts for the likelihood of the shock. This is the relevant marginal cost against which we make comparisons. The parties sell less than in the absence of the shock because with some probability they will not receive the entire revenues but will incur the production costs anyway. As the size of the shock, $1 - l$, decreases; the downward quantity distortion also decreases.

Let us now explore the consequences of asymmetric information as well as of the relational aspect of the contract. Let π_R denote the retailer's present discounted value of selling the good and repaying to the manufacturer from date t on:

$$\pi_R = (1 - s)[R(q) - D_H + \delta\pi_R] + s[lR(q) - D_L + \delta^{T+1}\pi_R]$$

The previous equation says that with probability $1 - s$ there is no shock. Therefore the retailer receives the entire revenues and pays back D_H to the manufacturer; in which case she renews the contract and hence, the game remains in this cooperative phase in the next period. However, with probability s, there is a shock that "destroys" part of the revenues. The retailer can only pay back D_L and the game moves to the termination phase in the next period. In this case, the retailer will earn again π_R only after the end

---

[17]When $l > 0$, the retailer can repay some amount following a shock. Not repaying anything implies stealing, in which case it is optimal to impose the maximum punishment because it helps to relax the dynamic enforcement constraint and it is not imposed in equilibrium. When $l = 0$, then $D_L = 0$ and the manufacturer punishes for T periods following no repayment.

of the punishment phase of T periods.

In a similar way, let π_M be the present discounted value for the manufacturer:

$$\pi_M = (1 - s)\left[D_H - cq + \delta\pi_M\right] + s\left[D_L - cq + \delta^{T+1}\pi_M\right]$$

With probability $1 - s$, the retailer repays D_H so the relationship move on to the next period and with the complementary probability, the retailer only repays D_L and the manufacturer terminates the contract for T periods.

Since the shocks are not observable to the manufacturer, she needs to ensure that the retailer has incentives to report the true demand. Due to the limited liability, only the retailer with a high demand can pretend to have encountered a low demand:

$$R(q) - D_H + \delta\pi_R \geq R(q) - D_L + \delta^{T+1}\pi_R \qquad (IC)$$

The incentive compatibility condition, IC, reflects the following retailer's trade-off: if he does not pay back the appropriate amount, he keeps $D_H - D_L$ but this will automatically trigger the termination phase, which yields valuation π_R only after T periods. For the retailer to pay back D_H, the manufacturer needs to ensure that tomorrow's gains from not being terminated are larger than the difference in payments today: $\delta\left(\pi_R - \delta^T\pi_R\right) \geq D_H - D_L$. The IC is not satisfied if the manufacturer never punishes. Similarly, the tougher the punishment, the easier it is for the manufacturer to satisfy the constraint.

Furthermore, the manufacturer need to ensure that the retailer wants to stay within the relational contract. The dynamic enforcement constraints ensure that a low (respectively, high) demand retailer does not want to walk away with all the revenues. These constraints are:

$$lR(q) - D_L + \delta^{T+1}\pi_R \geq lR(q) \qquad (DE)$$

and

$$R(q) - D_H + \delta\pi_R \geq R(q) \qquad (DE_H)$$

respectively. They require that the future expected value of being within the non-cooperative (cooperative) phase, rather than being terminated forever, are larger than the low (high) demand payment. Note that if IC and DE are satisfied, then DE_H is also satisfied. Since an $H - type$ retailer never wants to walk away with all his revenues, we can ignore DE_H.

In addition to the IC and DE constraints, the retailer is protected by limited liability. The retailer cannot be forced to make a repayment exceeding the revenues he reports in the current period. We call these constraints LLs.

Using the above information, the manufacturer's problem becomes:

$$\max_{q,D_H,D_L,T} \pi_M = \frac{(1-s)D_H + sD_L - cq}{1 - \delta(1-s) - s\delta^{T+1}}$$

s.t.

$$\delta\left(\pi_R - \delta^T \pi_R\right) \geq D_H - D_L \qquad\qquad (IC)$$

$$\delta^{T+1} \pi_R \geq D_L \qquad\qquad (DE)$$

$$R(q) \geq D_H \qquad\qquad (LL_H)$$

$$lR(q) \geq D_L \qquad\qquad (LL_L)$$

where $\pi_R = \frac{(1-s+sl)R(q)-[(1-s)D_H+sD_L]}{1-\delta(1-s)-s\delta^{T+1}}$.

By inspection of π_M, the manufacturer always wants to increase D_H and D_L as much as possible, therefore at least two constraints need to bind to bound D_H and D_L from above. We show in the Appendix that the IC is always binding and that the LL_H is never binding. The asymmetric information problem is always present and since the manufacturer need to give rents in at least one state to make the relationship valuable, she prefers to do so in the high state in order to relax the informational problem. Therefore, we are left with three possible combinations of binding constraints: $IC - DE$, $IC - DE - LL_L$ and $IC - LL_L$.

Note that there is a tension between addressing the asymmetric information problem (IC) and giving incentives to the retailer to keep trading (DE). Indeed, from IC, the

11

manufacturer would like to increase D_L[18] and T[19] as much as possible to give incentives to the high demand retailer to say the truth. At the same time, from DE we see that a larger D_L and T decreases the value of the relationship for a low demand retailer and hence increases his incentives to walk away.

If we rewrite π_M in terms of π_R: $\pi_M = \frac{(1-s+sl)R(q)-cq}{1-\delta(1-s)-s\delta^{T+1}} - \pi_R$, we can see that the manufacturer cannot appropriate all the surplus. The ultimate amount of surplus left to the retailer depends on which of the constraints are binding. To the extent that π_R increases with q, it is as if the manufacturer had an extra marginal cost and, hence, we expect quantity distortions to emerge.[20]

In what follows, we characterize each of the three regimes relegating tedious computations to the Appendix. At the end of this section, we discuss more precisely how the occurrence of each regime depends on the size of the shock and illustrate it with a numerical example.

3.1 Small shock

When the revenues obtained in the low demand state are nonetheless large, LL_L does not bind. As a result, only IC and DE bind which determine the payments in each state as a function of q and T: $D_H = \delta (1 - s (1 - l)) R(q)$ and $D_L = \delta^{T+1} (1 - s (1 - l)) R(q)$. Because the retailer can walk away with the current revenues, the manufacturer can only ask for a repayment in the high state as large as what the retailer can steal tomorrow in expected terms if he keeps trading with her (i.e. tomorrow's discounted expected revenues) and in the low state what he can steal the period after the termination ends.

---

[18]An increase in D_L decreases the RHS of IC. It also decreases the LHS indirectly through a decrease in π_R but to a less extent so overall the first effect dominates.

[19]Note that $\frac{1-\delta^T}{1-\delta(1-s)-s\delta^{T+1}}$ is increasing in T.

[20]When the retailer has the bargaining power, he maximizes: $\pi_R = \frac{(1-s+sl)R(q)-cq}{1-\delta(1-s)-s\delta^{T+1}} - \pi_M$ subject to (IC), (DE), (LL_H), (LL_L) and the participation constraint of the manufacturer, $\pi_M = 0$. We conjecture the quantity distortion to be significantly smaller. The reason is twofold: first, the retailer does not need to share the surplus with the manufacturer (i.e. the second part of the objective function is zero), and second, since the retailer keeps all the profits, we expect (DE) not to be binding (i.e. the retailer does not have incentives to walk away from the contract).

Using this information, the manufacturer's problem is:

$$\max_{q,T} \pi_M = \frac{\left(\left(1-s\right)\delta + s\delta^{T+1}\right)\left(1 - s\left(1-l\right)\right)R(q) - cq}{1 - \delta\left(1-s\right) - s\delta^{T+1}}$$

Note that T decreases the objective function: not only it increases the inefficiency from terminating following a shock but it also reduces the repayment when the demand is low. As a result, the manufacturer never punishes following a payment of D_L, that is, $T = 0$. Hence, the repayment are the same in both states: $D_H = D_L = \delta\left(1 - s\left(1-l\right)\right)R(q)$. The optimal quantity is given by:

$$R'(q) = \frac{c}{\delta\left(1 - s\left(1-l\right)\right)} > \widetilde{c} \tag{1}$$

Note that a larger δ or a smaller s increases the quantity sold and the total repayment.

Proposition 1 *When the shock is small, the manufacturer only needs to address the dynamic enforcement problem. Despite the use of a quantity forcing contract, the quantity is distorted downwards.*

When the low demand is nonetheless large, the manufacturer resolves the tension between the asymmetric information and the contract enforcement by removing the asymmetric information (i.e. asking for the same repayment no matter the state). This case is strategically equivalent to the one where the shocks are observable and (LL_L does not bind).[21]

We are in this regime when the LL_L is not binding, which implies $\frac{\delta(1-s)}{1-\delta s} < l$. Hence, a smaller δ or a larger s makes it easier for this condition to be satisfied.[22]

---

[21] To see this, imagine that the shocks are observable. Then the manufacturer asks for D_H if demand is high, D_L if demand is low and terminate forever if the appropriate payment is not made. Therefore, the manufacturer needs to replace (IC) by (DE_H) and transform (DE) in the following way: $\delta\pi_R \geq D_L$. These two constraints imply: $D_H = D_L = \delta\left(1 - s\left(1-l\right)\right)R(q)$, and hence we obtain the same solution.

[22] Note that if LL_L is satisfied, LL_H is also satisfied because the high revenues are larger for any q.

3.2 Large shock

When the low demand is very small, DE stops binding because a retailer hit by a shock does not want to walk away with such small revenues. At the same time, LL_L binds because the manufacturer would like to recover as much revenues as possible from the low demand state: $D_L = lR(q)$. The IC also binds because a high demand retailer has high incentives to pretend he received low revenues and repay such a small amount. This constraint determines the high demand payment: $D_H = \frac{\delta(1-s)+(1-\delta(1-s))l-(1-s+ls)\delta^{T+1}}{1-\delta^{T+1}}R(q)$.

Using this information, the manufacturer maximizes:

$$\max_{q,T} \pi_M = \frac{\frac{l+\delta(1-s)^2(1-l)-((1-s)(1-s+ls)+sl)\delta^{T+1}}{1-\delta^{T+1}}R(q) - cq}{1-\delta(1-s)-s\delta^{T+1}}$$

The quantity is determined by:

$$R'(q) = \frac{c}{\frac{l+\delta(1-s)^2(1-l)-((1-s)(1-s+ls)+sl)\delta^{T+1}}{1-\delta^{T+1}}} > \widetilde{c} \tag{2}$$

T is the optimal punishment defined in the Appendix where we also show that there is downward distortion in the quantity.[23]

Proposition 2 *When the shock is large, the manufacturer only needs to address the asymmetric information problem. Despite the use of a quantity forcing contract, the quantity is distorted downwards. A shorter punishment is associated with a larger quantity distortion.*

When the low demand is really small, only the asymmetric information problem matters as the low demand retailer does not obtain enough revenues to walk away with them. To tackle the asymmetric information problem, the supplier has to leave surplus in the high revenue state so that this state is reported, which explains the quantity distortion. There is a trade-off between a tougher punishment and a smaller quantity, because a

---

[23]Note that LL_H is always satisfied as it requires: $l < 1$.

longer termination decreases the incentives of the retailer to steal, allowing the manufacturer to keep more surplus. With a tougher punishment, the manufacturer can ask for a larger D_H and offer a larger q. However, increasing the punishment comes at a cost: it increases the inefficiency following a shock as she terminates profitable trading with the unlucky retailer.

When the shock "destroys" all the revenues ($l = 0$), the manufacturer terminates with the retailer forever if $s < \frac{(1-\delta)^2}{\delta^2}$, where s is the probability of the shock and δ the discount factor. Otherwise, depending on the parameters, she may also terminate forever (for instance when $s \to 1$), or set a positive and finite length of the punishment (for instance when $\delta \to 1$).

This case occurs when the dynamic enforcement constraint does not bind, which implies: $l < \frac{\delta^{T+1}(1-s)}{1-s\delta^{T+1}}$.

3.3   Intermediate shock

Finally, we look at the case where IC, DE and LL_L bind. This happens when the difference between the high and low demand is large enough so the limited liability constraint binds. At the same time, the low demand is nonetheless worthwhile to steal and the dynamic enforcement problem matters. In particular, we are in this case when: $\frac{\delta^{T+1}(1-s)}{1-s\delta^{T+1}} < l \leq \frac{\delta(1-s)}{1-\delta s}$[24]. The LL_L determines the payment in the low state as a function of the quantity: $D_L = lR(q)$. Then IC and DE jointly determine the high demand payment and the punishment period also as a function of q: $D_H = \delta(1 - s(1 - l))R(q)$ and $\delta^{T+1} = \frac{l}{1-s(1-l)}$ where $0 < T < +\infty$[25]. The difference in repayment, $D_L < D_H$, requires the manufacturer to terminate the contract for a positive (but finite) number of periods after a low payment .

---

[24]These inequalities hold for any T.

[25]T is positive because $\frac{l}{1-s+sl} < 1$ and $\frac{l}{1-s+sl} < \delta$ as l is constrained by $l < \frac{\delta(1-s)}{1-\delta s}$. Moreover, it is never infinity because $\delta < 1$.

Using this information, the manufacturer solves:

$$\max_q \pi_M = \frac{\delta \left(1 - s\right) \left(1 - s \left(1 - l\right)\right) R(q) + slR\left(q\right) - cq}{1 - \delta \left(1 - s\right) - \frac{sl}{1 - s(1 - l)}}$$

The quantity is determined by:

$$R'(q) = \frac{c}{\delta \left(1 - s\right) \left(1 - s \left(1 - l\right)\right) + sl} > \widetilde{c} \tag{3}$$

Proposition 3 *When the shock is intermediate, the manufacturer needs to address both the asymmetric information and the dynamic enforcement problems. The quantity distortion is larger than in the absence of asymmetric information.*

Indeed, since $0 < l < 1$ the downward quantity distortion is larger than the one in condition (1). This is because the manufacturer, in addition to preventing an $L - type$ retailer from walking away with the revenues, she also needs to prevent an $H - type$ retailer from pretending to be an $L - type$.

Finally, we can see from (3) that q increases following an increase in δ. A larger quantity increases the revenues and hence the repayment amounts D_H and D_L. Since D_H increases proportionally more (because the future becomes more valuable) an increase in the length of the termination period is needed to enforce these larger repayments. Instead, when s increases, q decreases as well as both repayment amounts. Since D_H decreases proportionally more a smaller T is needed (see Appendix).[26]

3.4 The value of the future

From the previous analysis, it emerges that even though the manufacturer has enough instruments to set the quantity, she chooses to sell less than the efficient amount in each of the three regimes. However, as the parties become arbitrarily patient, the downward distortion of q only remains when the limited liability binds.

---

[26]Note that LL_H is always satisfied as it requires: $\frac{\delta sl}{1 - \delta(1 - s)} < 1$.

Proposition 4 *As the discount factor tends to 1, the outcome is bounded away from efficiency if and only if the limited liability constraint binds.*

Proof. See Appendix. ■

Intuitively, when LL_L does not bind, the manufacturer is better off asking for the same quantity regardless of the demand realization to avoid using a termination policy to separate the types, as this practice destroys surplus. When the retailer values more the future, the manufacturer can offer a larger quantity and ask for a larger repayment because the retailer has more incentives to repay and stick with the relationship in the future. However, as the size of the shock increases, the same repayment regardless of the revenues would imply a too low repayment. Hence, the manufacturer is pushed to ask for a larger payment in "good times" and use the termination policy to enforce it. Since the retailer is credit constrained and inter-temporal transfers of payoffs are impossible to use, termination occurs in equilibrium, which bounds q away from efficiency for no matter which δ.

It is worth highlighting that it is the impatience of the retailer the one that is creating the downward distortion in the quantity. Indeed, if the discount factor for the retailer were different from the one for the manufacturer, then it would be retailer's discount factor appearing in conditions (1), (2) and (3). In contrast, both discounts factors would affect the choice of the termination policy.

Finally, to illustrate how the interaction between the different parameters determines the regimes, let us assume that the demand is linear: $R(q) = (a - bq)q$. Figures 2 and 3 depict which regime yields the largest profits for the manufacturer in the space δ and l. The difference between figures is the probability of the shock s.

Figure 2 depicts the case where the likelihood of the shock is low and equal to $s = 0.1$. The line that separates the green and the white area is $l = \frac{\delta(1-s)}{1-\delta s}$ and the line that separates the white and the red area is $l = \frac{\delta^{T+1}(1-s)}{1-s\delta^{T+1}}$ where T is the optimal termination policy. When the value of the future is small (i.e. δ is small) no contract can be implemented
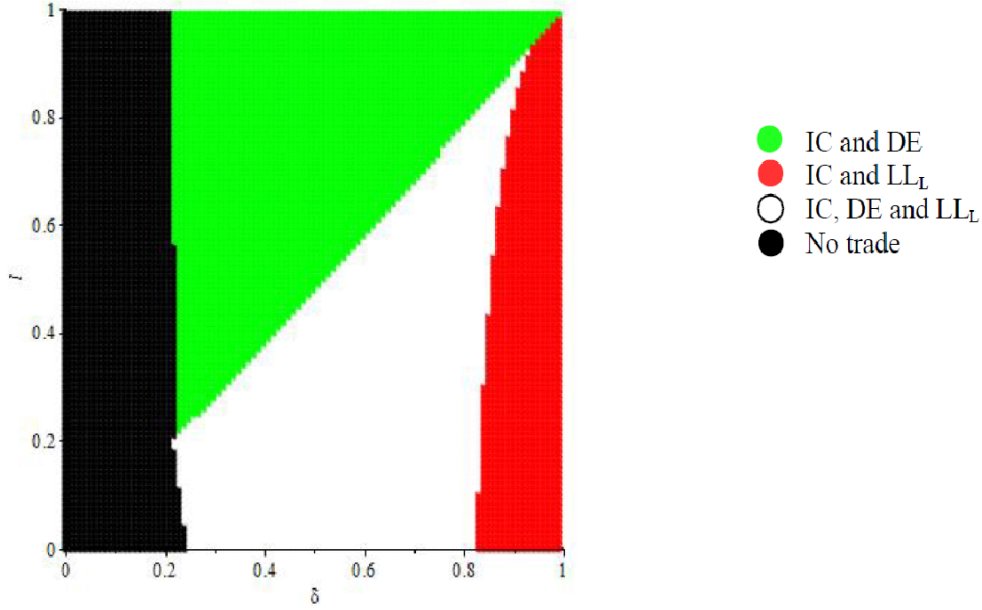
Figure 2: Optimal regime when $s = 0.1, a = 10, b = 1$ and $c = 2$

(and hence no trade occurs).[27] If the demands in both states are similar (i.e. l large and above $\frac{\delta(1-s)}{1-\delta s}$), then the manufacturer asks for the same payment and hence does not need to use any termination policy. As the low demand becomes smaller (l decreases below $\frac{\delta(1-s)}{1-\delta s}$), then LL_L starts binding and the manufacturer asks for a lower payment in the low state than in the high state ($D_L < D_H$) and punishes the retailer following the low payment for a finite number of periods. Finally, as l decreases (when $l < \frac{\delta^{T+1}(1-s)}{1-s\delta^{T+1}}$) DE stops binding, that is, the $L - type$ retailer does not have incentives to walk away with the revenues. This case also occurs as δ increases because the value of the relationship increases.

Figure 3 depicts a shock of likelihood $s = 0.9$. Note that as s increases, less contracts can be self-enforced (for low l and δ). Since the shock is more likely and the manufacturer does not want to punish an unlucky retailer, the regime IC and DE, where the payment is the same for both states and there is no termination, becomes more common.

---

[27] Although this is out of the scope of this model, the no trade situation could also correspond to a manufacturer that vertically integrates downwards and serves the consumers.
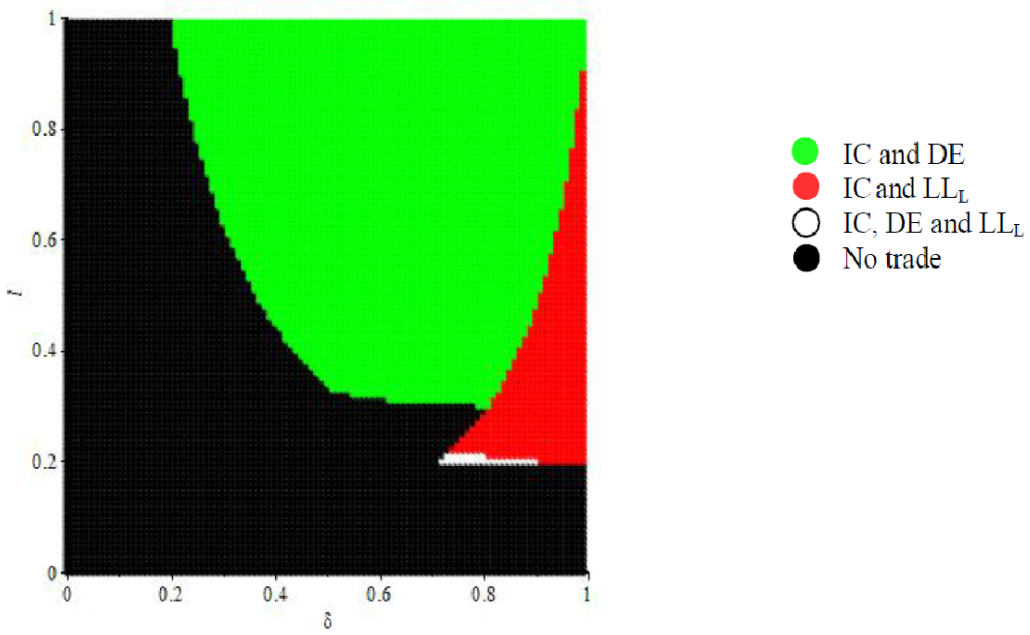
Figure 3: Optimal regime when $s = 0.9, a = 10, b = 1$ and $c = 2$

4 The Model

In this Section, we generalize the example by exploring the scenario where the retailer faces a continuum of demand states. The shock s is an *iid* random variable distributed on the interval $[0, \overline{s}]$ with $h(s)$ and $H(s)$ as the density and cumulative distribution functions, respectively. Denote the revenues for a given state s as $R(q; s)$ and the expected revenues as $R_E(q) = \int_0^{\overline{s}} R(q; s)h(s)ds$. Two observations are in order: first, and in contrast with the Example, a larger s means a better state of demand for a given q (that is, $\frac{\partial R(q;s)}{\partial s} > 0$); and second, the largest shock "destroys" all the revenues so the regime of a small shock in Section 3.1 is ruled out. We do this to focus on the effect of asymmetric information on the optimal contract. As in the example, the manufacturer offers a quantity q, a repayment $D(\widetilde{s})$ and a termination policy $T(\widetilde{s})$, for each particular shock reported $\widetilde{s}$.

19

4.1 Benchmark

When the shocks are observable and contracts can be enforced, the firms maximize their one-period joint profits:

$$\max_q R_E(q) - cq$$

The first best quantity is determined by:

$$q^{FB} : \int_0^{\bar{s}} \frac{\partial R(q;s)}{\partial q} h(s)ds = c$$

If the demand has the multiplicative functional form: $R(q;s) = sR(q)$, the optimal quantity is then given by:

$$R'(q) = \frac{c}{E(s)}$$

where $E(s) = \int_0^{\bar{s}} sh(s)ds$. If on average, the shocks are detrimental, there is reduction in quantity as in the Example. The effective marginal cost is $\widetilde{c} = \frac{c}{E(s)}$.

4.2 Relational contract

In this Section, we introduce the unobservability of the shocks and the need of relational contracts. We first find the conditions under which the contract is incentive compatible and/or dynamically enforceable and then we proceed to characterize the optimal contract. We conclude the Section with an example.

4.2.1 Incentive compatibility and dynamic enforceability

First, we establish the structure of the contract and then we lay down the conditions under which the retailer reports the true state of demand and does not have incentives to walk away from the contract.

Since the quantity is chosen before the state of the demand is realized, it is not used for sorting purposes. For a given state of the demand s and a given quantity q, the

retailer chooses a report $\widetilde{s}$ to maximize his profits:

$$\pi_R(\widetilde{s}; s) = \underbrace{R(q; s) - D(\widetilde{s})}_{\text{Today's payoff of reporting } \widetilde{s}} + \underbrace{\delta^{T(\widetilde{s})+1} \pi_R}_{\text{Expected payoff of reporting } \widetilde{s}}$$

where π_R is the expected discounted profits from staying in the relationship. Note that if the retailer were not credit constrained, the choice about which $\widetilde{s}$ to report would not depend on the true s as it is equally costly for any type of the retailer to report any $\widetilde{s}$. In other words, there is no sorting condition in this model, $\frac{\partial^2 \pi_R(\widetilde{s}; s)}{\partial \widetilde{s} \partial s} = 0$, as in Levin (2003). Because the retailer is credit constrained, however, he cannot repay a larger $D(\widetilde{s})$ than his actual revenues $R(q; s)$:

$$D(\widetilde{s}) \leq R(q; s) \quad \forall s, \widetilde{s} \tag{4}$$

The limited liability condition (4) links the choice of the report with the true state.

For a given q, the manufacturer has two instruments to deter the retailer from under-reporting the state of the demand: following a low $\widetilde{s}$, she can either increase the payment today $D(\widetilde{s})$ or increase the length of the termination period $T(\widetilde{s})$, which decreases the retailer's continuation value. Given that increasing $T(\widetilde{s})$ is also costly for the manufacturer (because she loses future trade), whenever it is possible, the manufacturer asks for the largest repayment: $D(\widetilde{s}) = R(\widetilde{s}, q)$.

In equilibrium, however, the manufacturer cannot extract all the revenues from the retailer in all the states, as this makes the relationship worthless to the retailer (i.e., $\pi_R = 0$). There must be a report s^* for which the manufacturer does not ask for all the revenues and, consequently, does not terminate, $T(s^*) = 0$[28]:

$$\pi_R(s^*; s) = R(q; s) - R(q; s^*) + \delta \pi_R \tag{5}$$

Hence for $\widetilde{s} \geq s^*$, the manufacturer can only offer the contract: $T(\widetilde{s}) = 0$ and $D(\widetilde{s}) = R(q; s^*)$, because she can no longer decrease the punishment period to compensate the

---

[28] If M were to terminate for some periods following the report s^*, she would strictly prefer to increase the repayment to decrease the termination length (and hence to reduce the inefficiency generated by the no-trade situation).

retailer for a larger repayment. Since the largest shock leads to zero revenues, the limited liability constraint (4) is always binding for some states and hence $s^* > 0$. To summarize, the manufacturer offers the following repayment schedule:

$$D(\widetilde{s}) = \begin{cases} R(q; \widetilde{s}) \text{ if } \widetilde{s} < s^* \\ R(q; s^*) \text{ if } \widetilde{s} \geq s^* \end{cases}$$

and termination policy:

$$T(\widetilde{s}) = \begin{cases} T(\widetilde{s}) \text{ if } \widetilde{s} < s^* \\ 0 \ \text{ if } \widetilde{s} \geq s^* \end{cases}$$

where $T(\widetilde{s})$ is to be defined[29]. The contract offered by the manufacturer resembles to what it is known in the corporate finance literature as a debt contract. Debt contracts leave no money to the borrower in the bad states while making the borrower the residual claimant in the good states.

Lemma 1 *The optimal contract is a debt contract.*

A debt contract is optimal in this model because it minimizes the inefficiency associated with the termination (by trading-off larger repayments for lower termination periods in the default states) while inducing the retailer to report the true.

Let us proceed to lay down the conditions under which this contract induces truth-telling. Since $\frac{\partial^2 \pi_R(\widetilde{s};s)}{\partial \widetilde{s} \partial s} = 0$, let us denote by $u(\widetilde{s})$ the part of the retailer's payoff that does not depend on his type s:

$$u(\widetilde{s}) = -D(\widetilde{s}) + \delta^{T(\widetilde{s})+1} \pi_R \tag{6}$$

The independence between the incentives to report a demand state and the actual demand state makes the task of inducing truth-telling quite simple. Intuitively, if it were feasible, the retailer would always report the demand state $\widetilde{s}$ with the highest $u(\widetilde{s})$, regardless of the true s. The retailer reports the truth if the combination of the repayment and

---

[29] In equation (7).

continuation value is not dependent on s: $u'(\widetilde{s})\mid_{\widetilde{s}=s} = 0 \ \forall s$. Furthermore, the retailer does not want to walk away from the contract if this combination is non-negative: $u(\widetilde{s})\mid_{\widetilde{s}=s} \geq 0$ $\forall s$. In terms of the Example in Section 3, the first constraint corresponds to the incentive compatibility constraint, IC, and the second to the dynamic enforcement constraint, DE.

Proposition 5 *The agent's payoff consist of two parts: first, the current period revenues and second, a combination of the repayment and continuation value. The second part is non-negative and state-independent.*

Proof. See Appendix. ∎

Since the rent never depends on the report, let us redefine $u(\widetilde{s}) = u$. Equation (5) becomes $u = -R(q;s^*) + \delta\pi_R$ when $u > 0$. We can combine it with (6) to define the termination policy:

$$\delta^{T(s)+1} = \frac{\delta\pi_R - R(q;s^*)}{\pi_R} + \frac{R(q;s)}{\pi_R} \tag{7}$$

Equation (7) defines $T(\widetilde{s})$. Note that when the retailer reports $s = 0$, if $u > 0$, then the the manufacturer terminates for a finite number of periods, while if $u = 0$, she terminates the contract forever. In the last case, equation (7) becomes $\delta^{T(s)+1} = \frac{R(q;s)}{\pi_R}$. In both cases, a larger report is coupled with a shorter punishment period.

Before solving the problem of the manufacturer, let us find π_R using Lemma 1:

$$\pi_R = \int_0^{s^*} \delta^{T(s)+1}\pi_R h(s)ds + \int_{s^*}^{\overline{s}} \left[R(q;s) - R(q;s^*) + \delta\pi_R\right] h(s)ds$$

The retailer gives back all his revenues if the shock is smaller than s^* and has the contract terminated for $T(s)$ periods. Otherwise, he repays the constant amount $R(q;s^*)$ and keeps trading with the manufacturer in the next period. Using condition (7), π_R can be simplified to:[30]

$$\pi_R = R_E(q) - R(q;s^*) + \delta\pi_R$$

Consider a first best world where the retailer is not credit constrained. Then an optimal contract could be a quantity (chosen so that the expected revenues minus the production

---

[30]If $u = 0$, then $\pi_R = R_E(q)$.

costs are maximized) and a fixed up-front payment from the retailer. If the manufacturer has all the bargaining power, then this payment would be equal to the expected revenues (and hence the manufacturer would extract all the rents, $\pi_R = 0$) and the retailer would never have his contract terminated.[31] We can interpret the above equation in these terms, where the retailer keeps the expected profits minus a fixed payment, $R(q; s^*)$, and he never has his contract terminated. The difference from the first best world comes in terms of a different quantity and a smaller repayment amount in order to leave some rents to the retailer ($\pi_R > 0$) so it is worth for him to stay in relationship. Solving for π_R yields:

$$\pi_R = \frac{R_E(q) - R(q; s^*)}{1 - \delta} \tag{8}$$

Hence, for the expected discounted profits to be positive, it needs to be the case that the revenues at s^* are smaller than the expected revenues.

Finally, using (8), the dynamic enforcement condition of Proposition 5 becomes:

$$u = \frac{\delta R_E(q) - R(q; s^*)}{1 - \delta} \geq 0 \tag{9}$$

In order to ensure the repayment, the manufacturer needs to guarantee that the retailer obtains at least the difference between what he can steal tomorrow if he stays in the relationship and the maximum repayment that he can potentially face today.

4.2.2   Optimal contract

Using Lemma 1, the profits of the manufacturer are:

$$\pi_M = \int_0^{s^*} \left[R(q; s) + \delta^{T(s)+1} \pi_M \right] h(s) ds + (1 - H(s^*)) \left[R(q; s^*) + \delta \pi_M \right] - cq$$

that is, the manufacturer receives all the revenues in the bad states ($s < s^*$) and terminates the contract with the retailer for $T(s)$ periods. Otherwise, the manufacturer charges

---

[31] Note that an ex-ante equivalent possibility is for the manufacturer to ask for the obtained revenues in all the states.

a fixed repayment, $R(q; s^*)$, and never terminates the contract. Finally, the manufacturer incurs the production costs.

When choosing s^*, the manufacturer faces the following trade-off: a larger s^* allows the manufacturer to extract a larger expected repayment from the retailer today (and hence π_R decreases); however, by (7), and in order to keep the incentives unchanged for all other types, this comes at a cost of a longer termination period.[32] On the other hand, an increase in the quantity q leads to an increase in the total payment as well as an increase in the production costs. Since it also leads to an increase of π_R, the effect on the termination policy is less clear and depends on the particular form of the demand function.[33]

Rewriting π_M in terms of π_R, illustrates the fact that the manufacturer cannot appropriate all the surplus:

$$\pi_M = \frac{R_E(q) - cq}{1 - \delta \int_{s^*}^{\bar{s}} h(s)ds - \int_0^{s^*} \delta^{T(s)+1} h(s)ds} - \pi_R$$

As in the Example of Section 3, since π_R increases with q, it is as if the manufacturer had an extra marginal cost associated with the asymmetric information and enforceability problem. As a result, we expect quantity distortions to emerge.

The problem of the manufacturer is then to choose s^* and q to maximize π_M subject

---

[32] After substituting for $\pi_R(s^*, q)$, the sign of the derivative is negative:

$$\frac{\partial \delta^{T(s,q)}}{\partial s^*} = -\frac{1 - \delta}{\delta} \frac{\partial R(s^*, q)}{\partial s^*} \frac{R_E(q) - R(s, q)}{(R_E(q) - R(s^*, q))^2} < 0 \ \forall s \leq s^*$$

where the inequality follows from the fact that in order to have $\pi_R(s^*, q) > 0$, it needs to be the case that $R_E(q) > R(s^*, q)$.

[33] In the multiplicative example that we present below, $T(s)$ does not depend on q.

to (9). After replacing π_R, the problem of the manufacturer becomes[34]:

$$\max_{q,s^*} \pi_M = \frac{\int_0^{s^*} R(q;s)h(s)ds + (1 - H(s^*)) R(q;s^*) - cq}{(1 - \delta) \frac{\int_{s^*}^{\bar{s}} (R(q;s) - R(q;s^*))h(s)ds}{R_E(q) - R(q;s^*)}} \tag{10}$$

$$\text{s.t. } \frac{\delta R_E(q) - R(q;s^*)}{1 - \delta} \geq 0$$

Note that as $\delta \to 1$, the dynamic enforcement constraint (9) does not bind. When this is the case, the optimal s^* and q do not depend on δ. Conversely, if (9) binds, both choices depend on δ. In particular, and in line with the Example, the fixed repayment that the manufacturer asks for the types above s^* is equal to what the retailer expects to obtain tomorrow if he does not repay: $R(q;s^*) = \delta R_E(q)$. Therefore, a larger discount factor allows the manufacturer to impose a larger s^*. Also, when (9) binds, an increase in s^* does not unambiguously increase the expected repayment because the fixed repayment $R(q;s^*)$ is constrained by $\delta R_E(q)$ and hence remains unchanged. Since the expected repayment remains unchanged, the termination policy does not change either. Thus, the only instrument available for the manufacturer is the choice of q. Note as well that the manufacturer will terminate with the retailer forever if there is no repayment at all.

The following Proposition summarizes this discussion.

Proposition 6 *The optimal s^* and q are independent of the discount factor, δ, if and only if δ is large enough: $\delta > \frac{R(q;s^*)}{R_E(q)}$, where $R(q;s^*)$ is the fixed repayment and $R_E(q)$ the expected revenues.[35]*

Proof. See Appendix. ∎

Since the solution depends on the particular demand function, in what follows, we illustrate these results with an example. For simplicity and comparability with the previous example we use the multiplicative demand function.

---

[34]Note that the retailer's participation constraint is never binding as he always has the option of walking away with the current revenues $R(q;s)$ by reporting $\widetilde{s} = 0$ and hence not repaying anything.

[35]The full version of this Proposition in the Appendix contains the first-order conditions from which s^* and q are found both for this and lower δ.

4.2.3 Example

Let us consider the following revenue function $R(q; s) = sR(q)$. Then the retailer's profits are: $\pi_R = \frac{(E(s)-s^*)R(q)}{1-\delta}$, and thus, in order for $\pi_R > 0$, the threshold s^* needs to be smaller than the expected shock. The problem of the manufacturer (10) becomes:

$$\max_{q,s^*} \pi_M = \frac{\widehat{E}(s,s^*)R(q) - cq}{(1-\delta)\frac{E(s)-\widehat{E}(s,s^*)}{E(s)-s^*}}$$

$$\text{s.t.} \quad \frac{\delta E(s) - s^*}{1-\delta} \geq 0 \tag{11}$$

where $\widehat{E}(s,s^*) = H(s^*)E(s \mid s \leq s^*) + (1 - H(s^*))s^*$ is the expected shock that matters to the manufacturer in terms of the repayment. Indeed, for $s \leq s^*$ the manufacturer gets all the revenues repaid while for $s > s^*$ she gets the revenues of the state s^*. Clearly, for any $s^* < \bar{s}$, $\widehat{E}(s,s^*) < E(s)$, which leads to the downward distortion of q compared to the benchmark. The dynamic enforcement constraint (9) becomes (11). It does not bind if s^* is smaller than the discounted value of the expected shock tomorrow, which happens for a large δ.

Figure 4 depicts the current and future profits of the manufacturer as a function of the realization of the shock when it is distributed like a Uniform on $[0, 1]$.

The payment today is the current revenues, $R(q; s)$, up to s^*, and a fixed payment, $R(q; s^*)$, from s^* onwards. The future discounted benefits from maintaining the relationship are constant and equal to $\delta\pi_M$ for s larger than s^*, and $\delta^{T(s)+1}\pi_M$ for s smaller than s^*.

Before we state the first order conditions, let us understand the trade-offs that the manufacturer is facing when choosing s^* and q. If the dynamic enforcement constraint (11) binds, then s^* is determined by it so the manufacturer can only choose q. In particular, the manufacturer sells the quantity that maximizes her profits taking into account the repayment that she is expected to obtain and that it is determined by s^*.

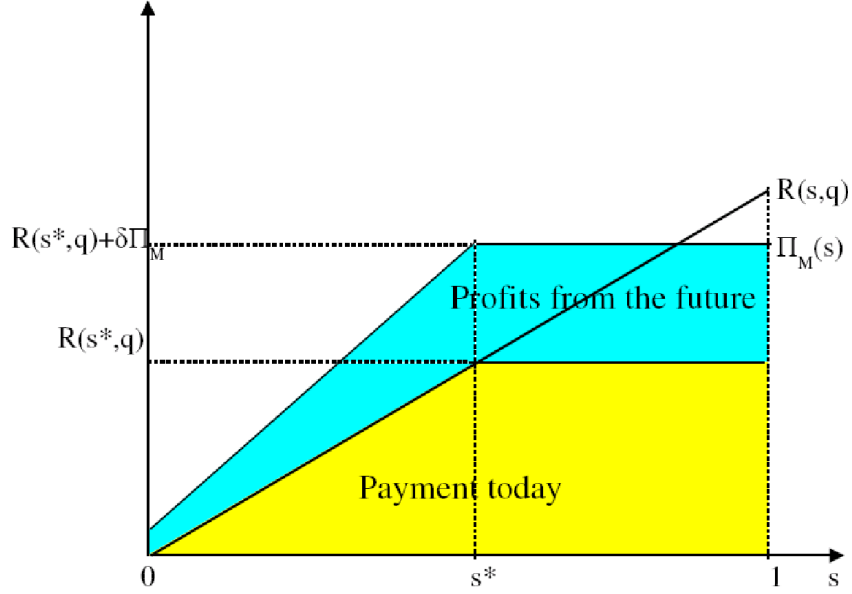When the constraint (11) does not bind, the manufacturer can also choose s^*. A larger

Figure 4: $p(s, q) = sp(q)$, $s \rightsquigarrow U[0, 1]$

s^*, on one hand, increases the expected repayment[36] and hence π_M. On the other, because it also decreases π_R and the retailer's incentives to repay, it increases the inefficiency due to tougher punishments[37].

Corollary 2 *If $R(q; s) = sR(q)$, then the optimal s^* and q are determined by the following first order conditions when $\delta E(s) > s^*$:*

$$q : R'(q) = \frac{c}{\widehat{E}(s, s^*)} \tag{12}$$

$$s^* : \frac{p(q)}{c} = \frac{H(s^*)}{1 - H(s^*)} \frac{E(s) - E(s \mid s \leq s^*)}{\widehat{E}(s, s^*)(E(s \mid s \geq s^*) - s^*) - E(s)(E(s) - s^*)} \tag{13}$$

Otherwise, $s^* = \delta E(s)$ and q is determined by (12). The optimal s^* and q are strategic complements.

Note that (12) and (13) do not contain δ. Thus if the value of the future is large, the constraint (11) does not bind, and the choice of q and s^* is not affected by δ. In this

---

[36] Note that $\frac{\partial \widehat{E}(s, s^*)}{\partial s^*} > 0$.

[37] This is reflected in the denominator of π_M, which is increasing in s^* because $\frac{\partial \widehat{E}(s, s^*)}{\partial s^*} < 1$.

case, the manufacturer would set the first best quantity only if $s^* = \bar{s}$[38], which is not possible as this makes the retailer's profits, π_R, zero. Therefore, Proposition 4 applies and the outcome is bounded away from efficiency even for δ arbitrarily close to 1. This is because in order to make the relationship profitable for the manufacturer, the limited liability constraint (4) binds for some states (i.e. $0 < s^*$[39]). This in turn implies that the termination is used in equilibrium and thus the surplus is bounded away from efficiency.

When the value of the future is low enough, $\delta < \frac{s^*}{E(s)}$, then the dynamic enforcement constraint (11) binds, in which case s^* and q do depend on δ. Then an efficient quantity will require $\delta > 1$[40], which is not possible either. Finally, note that, since $\frac{\partial \widehat{E}(s,s^*)}{\partial s^*} > 0$, the downward distortion in the quantity decreases with s^*, regardless of whether the constraint binds. Therefore, q and s^* are strategic complements, that is, a tougher punishment policy is accompanied by a smaller quantity distortion. This is because a tougher punishment can be used, instead of giving away surplus, to provide incentives to report the truth. The intuition behind the quantity distortion is the following. The manufacturer always needs to address the asymmetric information problem. As a result, she has to leave enough surplus in the high revenue states so that the true state is reported. When the dynamic enforcement constraint binds, the manufacturer also needs to leave surplus so the retailer does not walk away from the contract. The manufacturer does not internalize the positive externality that a larger quantity has on the retailer and the resulting quantity is too small. It is as if the manufacturer was facing an additional marginal cost as a result of the relational contract and the asymmetric information.

Finally, in order to compare the general model with the numerical example in Section 3, we assume the linear demand function $R(q; s) = s(a - bq)q$ and that s is distributed

---

[38]Because only then: $E(s) = \widehat{E}(s, s^*)$.

[39]Otherwise, M would produce no quantity because $\widehat{E}(s, s^*) = 0$.

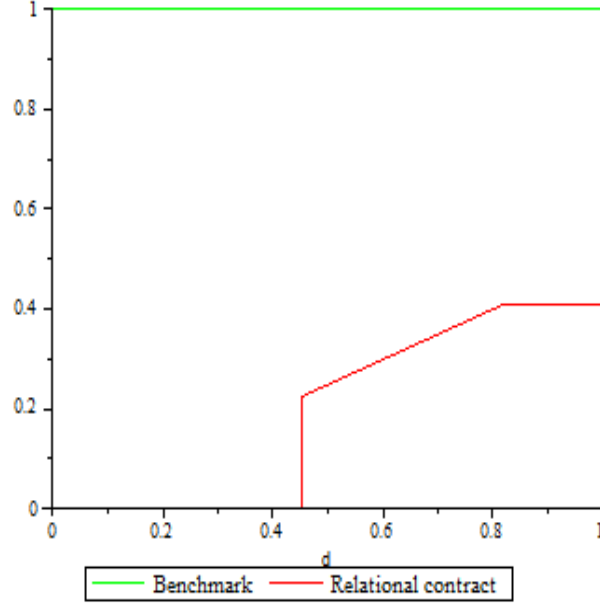[40]Indeed, $\delta = \frac{\bar{s}}{E(s)} > 1$.

Figure 5: s^* when $a = 10$, $b = 1$ and $c = 2$

uniformly on the interval $[0, 1]$[41]. For $a = 10$, $b = 1$ and $c = 2$, the optimal $s^* = 0.41$ and $q = 1.92$ if δ is at least 0.82. Otherwise, s^* and q are increasing in the discount factor. If $\delta < 0.45$, then there is no trade. Figures 5 and 6 depict s^* and q as a function of δ for this particular example.

5 Conclusions

The goal of this paper has not been to explain why and how much trade credit is offered by a supplier[42]. We take this decision as given, and rather we explore how trade credit

---

[41] The first order conditions, (12) and (13) become:

$$q \quad : \quad p'(q)q + p(q) = \frac{c}{s^* \left(1 - \frac{s^*}{2}\right)}$$

$$s^* \quad : \quad \frac{p(q)}{c} = \frac{s^*}{s^* \left(1 - \frac{s^*}{2}\right)(1 - s^*) - \left(\frac{1}{2} - s^*\right)}$$

[42] Answers to these questions can, for instance, be found in Burkart and Ellingsten (2004) that show how trade credit and bank lending are complements because goods are less divertable to private benefits than money. In the same way, Cunat (2007) shows how suppliers of services and differentiated goods are more willing to sell on credit than suppliers of standarized goods because they may be harder to replace and hence the downstream firm is more reluctant to default. Also, Smith (1987) finds that trade credit
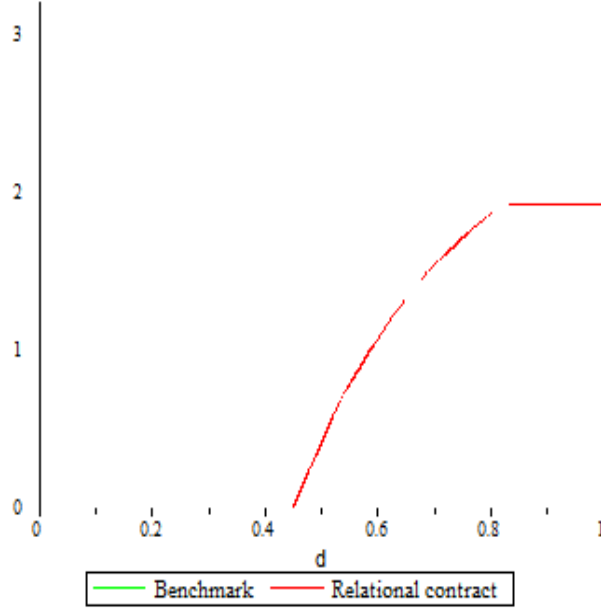
Figure 6: q when $a = 10$, $b = 1$ and $c = 2$

affects the different contract characteristics, such as late payment penalty, quantity and price of the good sold. The main prediction of this analysis is that the use of trade credit does have an important impact on the market outcome. In particular, the quantity sold in the market is expected to be lower than the efficient one. The supplier shares the surplus with the downstream firm to give him incentives to reveal the true demand and/or stay within the contract. As a result, the upstream firm does not internalize the effect of a larger quantity on the downstream firm.

We also show that the optimal contract resembles a debt contract. Debt contracts are successful in keeping the termination policy to the minimum while still providing the downstream firm incentives to repay the appropriate amount.

In our analysis, we have assumed that the quantity offered by the downstream firm does not change depending on the past repayment history. This framework would be suitable for industries where it is very costly to adjust the production quantity from

___

may be a consequence of an agency problem. Indeed, if there is quality variation in the good supplied, the downstream firm may be more reluctant to pay before having had the time to inspect the good. Another instance is Daripa and Nilsen (2011) who point out that trade credit mitigates the negative externality on the manufacturer from the retailer's trade off between loss sales and inventory costs.

one period to the other. In the future, it would be interesting to explore the form of non-stationary contracts and determine in which particular way the upstream firm will increase or decrease the quantity in each period as a function of the previous period repayment.

Finally, since the use of trade credit does not allow the firms to share the joint profits in an arbitrary way (i.e. using fixed transfers), the contract is expected to change depending on whether it is the upstream or downstream firm making the offer. It is left for future work to explore in which particular way the contract would change if it were offered by the downstream firm.

6 Appendix

Computations for the Example. When IC and LL_L bind, simple algebra shows that the RHS of condition (2) is larger than $\widetilde{c}$ when $1 - \delta (1 - s) - s\delta^{T+1} > 0$, which is always true, and that it is decreasing in T:

$$\frac{\partial \text{RHS of (2)}}{\partial T} = \frac{c \ln \delta (1-s)^2 (1-\delta)(1-l)\delta^{T+1}}{\left(\delta (1-s)^2 + \left(1 - \delta (1-s)^2\right) l - ((1-s)(1-s+ls) + sl)\delta^{T+1}\right)^2} < 0$$

The first order condition for T is:

$$\frac{\partial \pi_M}{\partial T} = \frac{-\delta^{T+1} \ln \delta \left[\frac{\left[(1-\delta)^2 - s\delta^2\left(1-\delta^T\right)^2\right](1-s)(1-s+sl) - \left(1-\delta(1-s)-s\delta^{T+1}\right)^2 l}{\left(1-\delta^{T+1}\right)^2} R(q) + scq\right]}{\left(1 - \delta (1-s) - s\delta^{T+1}\right)^2}$$

It is possible to give a more precise answer when $l = 0$ (i.e. the shock "destroys" all the revenues), which implies $D_L = 0$. In this case, $\frac{\partial \pi_M}{\partial T} > 0$ when $T = 0$ and it tends to 0^+ as $T \to +\infty$ if $s < \frac{(1-\delta)^2}{\delta^2}$. Thus, in the white region of Figure 7 M terminates forever. If $s > \frac{(1-\delta)^2}{\delta^2}$, M still terminates forever if:

$$\frac{cq}{\delta (1-s)^2} > \left[\frac{\delta^2 s - (1-\delta)^2}{\delta s}\right] R(q)$$
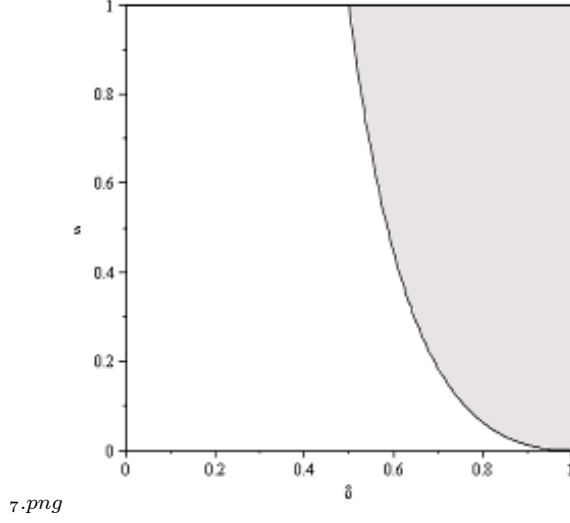
in Ex1

7.png

Figure 7: In the white region the optimal T is termination forever

Since by (2), $\frac{cq}{\delta(1-s)^2} < R(q)$, and because $\frac{\delta^2 s - (1-\delta)^2}{\delta s} \in [0,1]$, for some parameter values, this inequality may be true (for instance when $s \to 1$). When the inequality does not hold (for instance when $\delta \to 1$), the optimal termination period is positive and finite. Since the second derivative (evaluated at $\frac{\partial \pi_M}{\partial T} = 0$) is negative, the solution is a maximum.

$$\frac{\partial^2 \pi_M}{\partial T^2} = -\frac{\delta^{2(T+1)} (\ln \delta)^2 (1-s)^2 \left((1-\delta)^2 + \delta s \left(1-\delta^T\right)\left(1-\delta+1-\delta^{T+1}\right)\right)}{\left(1-\delta+s\delta\left(1-\delta^T\right)\right)^2 \left(1-\delta^{T+1}\right)^4} R(q)$$

For the case where IC, DE and LL_L bind, the comparative statics of s are:

$$\frac{\partial^2 \pi_M}{\partial q \partial s} \propto -\frac{c\left(\delta\left[1-s\left(1-l\right)+\left(1-s\right)\left(1-l\right)\right]-l\right)}{\left(\delta\left(1-s\right)\left(1-s\left(1-l\right)\right)+sl\right)^2}$$

therefore, q decreases following an increase in s if:

$$\delta > \frac{l}{1-s+(1-s)(1-l)+sl}$$

which is always true because the RHS is increasing in l, and the inequality holds at the upper bound of l, $\frac{\delta(1-s)}{1-\delta s}$. By inspection, it is easy to see that $\frac{\partial D_H}{\partial s} < 0$, $\frac{\partial D_L}{\partial s} < 0$ and

33

$\frac{\partial T}{\partial s} < 0$.

Finally, let us show that there are no other regimes. Suppose that LL_H binds, then $D_H = R(q)$ and either DE or LL_L need to bind to bound D_L. LL_L binding is not possible as this would make $\pi_R = 0$. DE binding makes the low repayment equal $D_L = \frac{s\delta^{T+1}lR(q)}{1-\delta(1-s)}$. With these two repayments, IC is satisfied only if $l > \frac{1-\delta(1-s)}{s\delta}$, which leads to a contradiction as $\frac{1-\delta(1-s)}{s\delta} > 1$. The IC needs to be binding because the case where only DE and LL_L bind is not possible. This is because an increase in the punishment period not only decreases π_M but also makes DE more binding. Therefore, the manufacturer chooses no punishment ($T = 0$) and as a result IC is violated. ∎

Proof of Proposition 4. By inspection, the RHS of (1) tends to $\widetilde{c}$ as $\delta \to 1$. Similarly, the RHS of (3) tends to $\frac{c}{(1-s)(1-s+sl)+sl}$, which is larger than $\widetilde{c}$. Using l'Hopital's rule, equation (2) becomes:

$$\lim_{\delta \to 1} R'(q) = \frac{c}{\frac{l+[(1-s)(1-s+sl)+sl]T}{T+1}} \tag{14}$$

Note that at $T = 0$, the RHS of (14) is larger than $\widetilde{c}$ and that it is decreasing in T. Noting that as $T \to +\infty$, the RHS of (14) tends to $\frac{c}{(1-s)(1-s+sl)+sl}$ completes the proof. ∎

Proof of Proposition 5. Because of condition (4), a retailer can lie only downwards. To avoid the retailer underreporting his type, the manufacturer needs to ensure that $u'(\widetilde{s}) \geq 0$, that is, the payoff $u(\widetilde{s})$ increases with reported $\widetilde{s}$. However, the manufacturer cannot design a contract that satisfies $u'(\widetilde{s}) > 0$ for $s \in [s^*, \overline{s}]$ because she has no means to increase $u(\widetilde{s})$ (i.e., she is already not terminating with the retailer and asking for a smaller repayment $R(q; s\prime) < R(q; s^*)$ would make types $s \in [s', s^*]$ overreport their types). Furthermore, for $s \in [0, s^*)$ the contract cannot satisfy $u'(\widetilde{s}) > 0$ either, because the manufacturer is already extracting the maximum payment and increasing the payoff with $\widetilde{s}$ would imply a costly (for the manufacturer) increase in the termination period for smaller types. Indeed, the manufacturer wants to set $T(s)$ as low as possible as long as there is truth-telling. Finally, to prevent the retailer walking away with the earnings, the manufacturer need to ensure that $u(\widetilde{s})$ is non-negative, which completes the proof. ∎

Proposition 6. The optimal s^* and q are determined by

$$\left(1 - H(s^*)\right)\pi_R - H(s^*)\pi_M \frac{E(s) - E\left(s \mid s \leq s^*\right)}{E(s) - s^*} = \widehat{\lambda} \tag{15a}$$

$$\pi_R \frac{\frac{\partial R(q;s^*)}{\partial q} - c}{\frac{\partial R(q;s^*)}{\partial q} - \delta \frac{\partial R_E(q)}{\partial q}} + \frac{\int_0^{s^*} \left(\frac{\partial R(q;s)}{\partial q} - \frac{\partial R(q;s^*)}{\partial q}\right) h(s) ds}{\frac{\partial R(q;s^*)}{\partial q} - \delta \frac{\partial R_E(q)}{\partial q}} \left(\pi_R + \pi_M\right) \tag{15b}$$

$$+\pi_M \left(\frac{\int_0^{s^*} \left(R(q;s^*) - R(q;s)\right) h(s) ds}{R_E(q) - R(q;s^*)} \frac{\frac{\partial R_E(q)}{\partial q} - \frac{\partial R(q;s^*)}{\partial q}}{\frac{\partial R(q;s^*)}{\partial q} - \delta \frac{\partial R_E(q)}{\partial q}}\right) = \widehat{\lambda}$$

where λ is the shadow cost of the dynamic enforcement constraint and $\widehat{\lambda}$ is adjusted by the second best loss of revenues, $\widehat{\lambda} = \lambda \left(R_E(q) - \int_0^{s^*} R(q;s)h(s)ds - \left(1 - H\left(s^*\right)\right)R(q;s^*)\right)$. If and only if δ is large enough, $\delta > \frac{R(q;s^*)}{R_E(q)}$, then $\widehat{\lambda} = 0$ and the optimal s^* and q are independent of δ. ∎

Proof of Proposition 6. The Lagrangian function is:

$$\mathcal{L} = \frac{\int_0^{s^*} R(q;s)h(s)ds + \left(1 - H\left(s^*\right)\right)R(q;s^*) - cq}{(1 - \delta)\left(1 + \frac{H(s^*)R(q;s^*) - \int_0^{s^*} R(q;s)h(s)ds}{R_E(q) - R(q;s^*)}\right)} - \lambda \left(R(q;s^*) - \delta R_E(q)\right)$$

where λ is the shadow cost associated with the constraint. The first order conditions are:

$$\frac{\partial \mathcal{L}}{\partial s^*} = \pi_R \frac{\partial R(q;s^*)}{\partial s^*} \frac{1 - H\left(s^*\right)}{R_E(q) - \int_0^{s^*} R(q;s)h(s)ds - \left(1 - H\left(s^*\right)\right)R(q;s^*)}$$

$$-\pi_M \frac{\partial R(q;s^*)}{\partial s^*} \frac{\frac{H(s^*)R_E(q) - \int_0^{s^*} R(q;s)h(s)ds}{R_E(q) - R(q;s^*)}}{R_E(q) - \int_0^{s^*} R(q;s)h(s)ds - \left(1 - H\left(s^*\right)\right)R(q;s^*)}$$

$$-\lambda \frac{\partial R(q;s^*)}{\partial s^*} = 0$$

$$\frac{\partial \mathcal{L}}{\partial q} = \pi_R \frac{\int_0^{s^*} \frac{\partial R(q;s)}{\partial q} h(s)ds + \left(1 - H\left(s^*\right)\right)\frac{\partial R(q;s^*)}{\partial q} - c}{R_E(q) - \int_0^{s^*} R(q;s)h(s)ds - \left(1 - H\left(s^*\right)\right)R(q;s^*)}$$

$$-\pi_M \frac{H\left(s^*\right)\frac{\partial R(q;s^*)}{\partial q} - \int_0^{s^*} \frac{\partial R(q;s)}{\partial q} h(s)ds - \frac{\left(H(s^*)R(q;s^*) - \int_0^{s^*} R(q;s)h(s)ds\right)\left(\frac{\partial R_E(q)}{\partial q} - \frac{\partial R(q;s^*)}{\partial q}\right)}{R_E(q) - R(q;s^*)}}{R_E(q) - \int_0^{s^*} R(q;s)h(s)ds - \left(1 - H\left(s^*\right)\right)R(q;s^*)}$$

$$-\lambda \left(\frac{\partial R(q;s^*)}{\partial q} - \delta \frac{\partial R_E(q)}{\partial q}\right)$$

and

$$\lambda \left(R(q; s^*) - \delta R_E(q) \right) = 0$$

Setting the first two conditions equal to zero, give us equation (15a) and (15b), respectively.

We assume that the second order conditions hold. Note that the optimal s^* cannot be in a corner as this would give zero profits to either the manufacturer (if $s^* = 0$) or to the retailer (if $s^* = 1$). We assume that the demand and the cost function are such that production takes place, i.e., $0 < q < +\infty$. ∎

Proof of Corollary 2. Suppose that the constraint does not bind, the first order conditions are:

$$\frac{\partial \pi_M}{\partial q} = \frac{\widehat{E}(s, s^*) \left(p'(q)q + p(q) \right) - c}{(1 - \delta) \frac{E(s) - \widehat{E}(s, s^*)}{E(s) - s^*}} = 0$$

$$\begin{aligned}
\frac{\partial \pi_M}{\partial s^*} &= (1 - H(s^*)) \frac{E(s) - s^*}{E(s) - \widehat{E}(s, s^*)} p(q)q \\
&+ H(s^*) \frac{E(s \mid s \leq s^*) - E(s)}{\left(E(s) - \widehat{E}(s, s^*) \right)^2} \left(\widehat{E}(s, s^*) p(q)q - cq \right) = 0
\end{aligned}$$

The cross derivative is:

$$\begin{aligned}
\frac{\partial^2 \pi_M}{\partial s^* \partial q} &= (1 - H(s^*)) \frac{E(s) - s^*}{E(s) - \widehat{E}(s, s^*)} \left(p'(q)q + p(q) \right) \\
&+ H(s^*) \frac{E(s \mid s \leq s^*) - E(s)}{\left(E(s) - \widehat{E}(s, s^*) \right)^2} \left(\widehat{E}(s, s^*) \left(p'(q)q + p(q) \right) - c \right) = 0
\end{aligned}$$

Using $\frac{\partial \pi_M}{\partial q} = 0$, it is easy to see that $\frac{\partial^2 \pi_M}{\partial s^* \partial q} > 0$. ∎

References

[1] BAKER, G., R. GIBBONS, and MURPHY, K. J., (1994) "Subjective Performance Measures in Optimal Incentive Contracts," *The Quarterly Journal of Economics,*

Vol.109(4), pp. 1125-1156.

[2] BERNSTEIN, L., (1992). "Opting Out of the Legal System: Extralegal Contractual Relations in the Diamond Industry.", *Journal of Legal Studies*, Vol. XXI, pp. 115–157.

[3] BERNSTEIN, L., (1996). "Merchant Law in a Merchant Court: Rethinking the Code's Search for Immanent Business Norms.", *University of Pennsylvania Law Review*, Vol. 144(5), pp. 1765–1821.

[4] BOISSAY, F., and GROPP, R., (2007), "Trade Credit Defaults and Liquidity Provision by Firms", European Central Bank Working Paper No. 753.

[5] BOLTON, P. and SCHARFSTEIN, D.S., (1990), "A Theory of Predation Based on Agency Problems in Financial Contracting." *American Economic Review*, Vol. 80, pp. 93-106.

[6] BUEHLER, S., and GARTNER D.L., (forthcoming), "Making Sense of Non-Binding Retail-Price Recommendations", *American Economic Review.*

[7] BURKART, M., and ELLINGSEN, T., (2004), "In-Kind Finance: A Theory of Trade Credit", *American Economic Review*, Vol. 94, pp. 569-590.

[8] CUNAT, V. M., (2007), "Trade Credit: Suppliers as Debt Collectors and Insurance Providers", *Review of Financial Studies*, Vol. 20, pp. 491-527.

[9] DARIPA, A., and NILSEN, J., (2011), "Ensuring Sales: A Theory of Inter-Firm Credit", *American Economic Journal: Microeconomics,* Vol. 3, pp. 245–279.

[10] FAFCHAMPS, M., (1997), "Trade credit in Zimbabwean manufacturing", *World Development*, Vol. 24 No. 3, pp. 795-815.

[11] FAFCHAMPS, M., (2000), "Ethnicity and credit in African manufacturing", *Journal of Development Economics*, Vol. 61, pp. 205-235.

[12] FAFCHAMPS, M., (2004), *Market Institutions in Sub-Saharan Africa: Theory and Evidence*, MIT Press, Cambridge Mass.

[13] FAFCHAMPS, M., (2010), "Spontaneous Markets, Networks and Social Capital: Lessons from Africa" Chapter in *Institutional Microeconomics of Development*, (edited by Timothy Besley and Rajshri Jayaraman), MIT Press, Cambridge Mass.

[14] FAURE-GRIMAUD, A., (2000), "Product Market Competition and Optimal Debt Contracts: The Limited Liability Effect Revisited." *European Economic Review*, Vol. 44, pp.1823-1840.

[15] GALE, D. and HELLWIG, M., (1985), "Incentive-Compatible Debt Contracts: The One-Period Problem." *Review of Economic Studies*, Vol. 52, pp. 647-663.

[16] GIANNETTI, M, BURKART, M and ELLINGSEN, T., (2011), "What You Sell Is What You Lend? Explaining Trade Credit Contracts", *The Review of Financial Studies,* Vol. 24, pp. 1261-1298.

[17] GREEN, E.J., and PORTER R.H., (1982), "Noncooperative collusion under imperfect price information", *Econometrica*, Vol. 52, No. 1 (Jan., 1984), pp. 87-100.

[18] HART, O. and MOORE, J., (1998), "Default and Renegotiation: A Dynamic Model of Debt." *Quarterly Journal of Economics*, Vol. 113, pp.1-41.

[19] INNES, R.D., (1990), "Limited Liability and Incentive Contracting with Ex-ante Action Choices." *Journal of Economic Theory*, Vol. 52, pp. 45-67.

[20] JOHNSON, S., MCMILLAN, J., and WOODRUFF, C., (2002), "Courts and Relational Contracts", Journal of Law, Economics and Organization, Vol. 18, pp. 221-77.

[21] LEVIN, J., (2003), "Relational Incentive Contracts", *American Economic Review*, Vol. 93, No. 3, pp. 835-857(23).

[22] MAILATH, G., and SAMUELSON, L., (2006), *Repeated Games and Reputations: Long-Run Relationships.* New York: Oxford University Press.

[23] MALCOMSON, J., (2010), "Relational incentive contracts", Chapter for *Handbook of Organizational Economics* (edited by Robert Gibbons and John Roberts), Princeton University Press.

[24] MCMILLAN, J. and WOODRUFF, C., (1999a), "Interfirm Relationships and Informal Credit in Vietnam", *Quarterly Journal of Economics*, Vol.114, pp. 1285-1320.

[25] MCMILLAN, J. and WOODRUFF, C., (1999b), "Dispute Prevention Without Courts in Vietnam," *Journal of Law, Economics, and Organization*, No.15, pp. 637-58.

[26] POVEL, P. and RAITH, M., (2004a), "Optimal Debt with Unobservable Investments." *RAND Journal of Economics*, Vol. 35, pp. 599-616.

[27] POVEL, P. and RAITH, M., (2004b), "Financial Constraints and Product Market Competition: Ex Ante vs. Ex Post Incentives." *International Journal of Industrial Organization*, Vol. 22, pp. 917-949.

[28] SMITH, J., (1987), "Trade Credit and Informational Asymmetry", *Journal of Finance*, Vol. 42, pp. 863-69.

[29] TOWNSEND, R.M., (1979), "Optimal Contracts and Competitive Markets with Costly State Verification." *Journal of Economic Theory*, Vol. 21, pp. 265-293.

[30] UCHIDA, H., UDELL, G.F., and WATANABE, W., (2006), "Are Trade Creditors Relationship Lenders?" http://www.rieti.go.jp/jp/publications/dp/06e026.pdf.