

Strategic Capacity Investment under Hold-up Threats: The Role of Contract Length and Width*

Laure Durand-Viel[†] Bertrand Villeneuve[‡]

February 16, 2015

Abstract. We analyze the impact of the length of incomplete contracts on investment and surplus sharing. In the bilateral relationship explored, the seller controls the input and the buyer invests. With two-part tariffs, the length of the contract is irrelevant: the surplus is maximal and goes to the seller. In linear contracts, the seller prefers the shortest contract and the buyer the longest one. Further, the commitment period concentrates the incentives, whereas afterwards there is rent extraction. The socially efficient contract is as short as possible; yet, long contracts can be promoted because of the surplus they allocate to the buyer.

JEL Codes: D42, D45, D92, L95.

Keywords: Long-term Contracts, Incomplete Contracting, Investment.

1 Introduction

Some equipment is very long-lived, and yet its capacity is determined upfront and cannot be readjusted. When making this long-term decision, the investor needs to anticipate the possible evolution of the prices of equipment-related inputs and outputs. The investor often seeks to obtain guarantees about future prices, but the equipment might last longer than any reasonable contract's duration. This duration might be a concern when the equipment uses an input produced by a firm with market power. If the commitment is limited, then investors might fear a hold-up, that is, unilateral price increases that will extract rent from the installed capital (Williamson, 1971). The energy sector at large offers a number of instances in which trade-offs must be found via contracts.

We examine investments for which the installed capacity acts as a cap on subsequent purchases. Once the investment has been made, buyers are locked into the chosen technology or standard, but they keep the flexibility to purchase less than nominal capacity. The obvious examples are oil and gas pipelines and most large industrial fuel-consuming appliances. Other important sectors exhibit similar features. For instance, there are those firms that invest in specific production lines by using patented technologies. These firms

*This work was supported by FiME (laboratoire de Finance des Marchés de l'Énergie) and the "Chaire Finance et Développement Durable - Approches Quantitatives".

[†]Autorité de la Concurrence and Université Paris-Dauphine.

[‡]Université Paris-Dauphine and CREST. Contact: bertrand.villeneuve@dauphine.fr.

dimension their plants after evaluating the expected intensity of the license use and the cost of the corresponding fees. Ex post, elasticity in production is asymmetric: it is easier to scale down production than to exceed nominal capacity. Another example is the franchisor/franchisee relationship: when a franchisee chooses an outlet of a given size and location, this investment caps the future sales capacity. Investment is decided in view of the expected payment to the franchisor, for example, royalties calculated as a percentage of turnover.

The equilibrium strategies of a firm try to attract customers who, once locked in by their investment choice, are at the mercy of the firm's market power. The industrial organization literature on standard competition or aftermarket monopolization addresses this issue (Shapiro, 1995, Chen et al., 1998, Reitzes and Woroch, 2008). These models make rather extreme assumptions about the link between equipment and complementary goods. Strict complementarity means that consumers need one aftermarket good for each primary good (see Carlton and Waldman, 2010); independence implies that demand ex post can be adjusted freely (see, e.g., Borenstein et al., 2000, Morita and Waldman, 2004). In our model, complementarity is asymmetric: the buyer cannot consume more than the capacity allows, but the buyer is free to consume less. The capacity constraint is carefully chosen by the buyer: investment itself, as well as the threat of scaling down demand anytime, is a strategic tool against the seller's potential abuses. In effect, the seller must not only convince the buyer to engage in a specific relationship, but the seller must incite the buyer to make a sufficiently large investment in capacity and to use it. Thus, the seller must guarantee reasonable tariffs to avoid killing the goose that lays the golden egg.

In a perfect world, signing a complete contract that gives adequate incentives to make the optimal investment should be possible. Sometimes, well-designed incomplete contracts are also efficient. This is the case for the repeated short contracts in Crawford (1988) and Rey and Salanié (1990), or for continuously renegotiated finite term contracts in Guriev and Kvasov (2005). The contracts with limited width that we study in this paper are optimized with respect to the length of compensation. Conventional wisdom says that long contracts protect investors (their profit) as well as the investment (social surplus). Eventually, the expectation is that the investment level and social welfare rise as the commitment duration increases, and the extra surplus is shared between the parties depending on their bargaining powers. On the other hand, caution dictates that incomplete contracts should not last too long because they could become very inefficient. A time limit might be necessary to avoid applying obsolete clauses, especially if some random factor drives the economy away from the initial conditions on the basis of which the contract was calibrated (Crocker and Masten, 1988, for natural gas contracts; see also Creti and Villeneuve, 2004, for a survey on take-or-pay gas contracts). Still, the trade-off is not expected to lead to very short contracts as the benefits of commitment should still be dominant.

Our model shows a simple and realistic context in which the shortest contract is the best. After the contract expires, prices are designed for rent extraction. Contracts must therefore be structured so that the commitment period concentrates the incentives to invest, and the seller should concede temporary advantageous conditions. When the commitment period is long, the committed price inevitably mixes two objectives: incentives and rent extraction. But these objectives are antagonistic, and a longer contract is less efficient at inducing profitable investment. The negative impact on the social surplus of a longer contract is very unequally shared between the parties. One could think they both somehow lose. In fact, the seller is a net loser, whereas the buyer produces less surplus but is still a net winner.¹ A longer contract actually protects expropriable investors (their profit) rather than the investment itself (social surplus).

We also argue that simpler contracts that are less efficient in general, such as contracts with purely linear prices, are robust and, more importantly, guarantee buyers a share of the surplus. Linear pricing as well as the duration of the contract could be seen as general rules that parties can agree on beforehand to prevent contract failure, or that can be imposed by regulators in the interest of buyers, while letting the parties negotiate about contractual details. In our view, contract length and width (price structure) can be viewed as key features of the “‘constitution’ governing the ongoing relationship” (Goldberg, 1976, p. 428). This view motivates our comparative statics analysis of the contract length in relation to the restrictions on the price structure.

The paper is organized as follows. Section 2 sets up the model. We analyze various scenarios with two-part tariffs in section 3. Linear contracts are justified and characterized in section 4. The proofs are relegated to the [Appendix](#).

2 General setting

2.1 Trade and payoffs

The game involves two players: the seller and the buyer. The model is in continuous time with infinite horizon. At instant t , the seller sells to the buyer a quantity q_t of a commodity produced at a constant unit cost c ; he receives in exchange a payment τ_t , expressed in units of the numéraire good.

The instantaneous utility of the buyer is quasi-linear. It can be written (up to some

¹Castaneda (2006) also finds a negative effect of longer contracts on investment, but in his model both parties are worse off. The contract in his analysis is such that the buyer pays a lump sum upfront and that the seller invests after acceptance. The buyer thus is reluctant to accept long contracts that could establish extortion for a long time, which in turn means that the seller’s return on them is poor that depresses investment.

irrelevant constant) $u(q_t) - \tau_t$. We will use the iso-elastic utility function

$$(1) \quad u(q) = \frac{\varepsilon d^{\frac{1}{\varepsilon}}}{\varepsilon - 1} q^{\frac{\varepsilon - 1}{\varepsilon}},$$

where d is a positive scale parameter and where ε is the price elasticity of demand. We assume that $\varepsilon > 1$.

The seller sets the tariffs $\{\tau_t(\cdot), t \geq 0\}$ where the argument of $\tau_t(\cdot)$ is q_t . Only linear and two-part tariffs are used. The general shape of the total payment at date t to the seller for quantity q_t is therefore $\tau_t(q_t) = m_t + p_t q_t$, where m_t is a per-time-unit fixed fee and p_t is the marginal price.

Thus, the instantaneous buyer demand $Q(\cdot)$ exhibits constant price elasticity:

$$(2) \quad Q(p_t) = d p_t^{-\varepsilon}.$$

The inverse demand function Q^{-1} is denoted P .

The consumption of the commodity requires an equipment with an indefinite service life, installed at date $t = 0$, whose size A determines maximum consumption once and for all. For example if the equipment is a pipeline, imports cannot exceed the pipeline's capacity:

$$(3) \quad q_t = \min\{Q(p_t), A\}.$$

Unless otherwise specified, the investment cost (k per unit) is assumed to be borne by the buyer, who chooses non-cooperatively the investment size A (the case where the seller is the investor is also discussed further on). The future is discounted at rate r .

The intertemporal surplus of the buyer is

$$(4) \quad S = \int_0^{+\infty} (u(q_t) - \tau_t(q_t)) e^{-rt} dt - kA.$$

The intertemporal profit of the seller is

$$(5) \quad \Pi = \int_0^{+\infty} (\tau_t(q_t) - c q_t) e^{-rt} dt.$$

The first best. The social optimum requires $u'(q_t) = c + rk$, which means constant consumption: $\forall t, q_t = Q(c + rk)$. The term rk is the unit amortization cost of the durable investment. The efficient investment is thus

$$(6) \quad A^* = Q(c + rk).$$

2.2 Market power and limited commitment

Trade takes place only after the equipment has been built, however trading terms can be determined beforehand by means of an agreed tariff. T is the commitment duration.

1. At date 0, the seller makes a take-it-or-leave-it offer to the buyer consisting of tariffs $\{\tau_t(\cdot), t \in [0, T)\}$, valid until T . The buyer accepts this offer if it leaves him with a non-negative surplus.
2. The buyer invests A .
3. From $t = 0$ to T , the buyer purchases q_t at each date t .
4. At date T , the contract expires. At each date $t \geq T$, the seller makes a take-it-or-leave-it offer to the buyer consisting of tariff $\tau_t(\cdot)$. The buyer accepts this offer if it leaves him with a non-negative instantaneous surplus, and consumes the corresponding q_t .

Without loss of generality, the seller's pricing strategies can be restricted to constant marginal prices over each period: p_0 during the first period ($0 \leq t < T$) and p_T during the second period (for $t \geq T$). The way the fixed payments are staggered over time within the first period does not matter in the sense that paying m_t at all dates $t \in [0, T)$ is equivalent to an upfront payment²

$$(7) \quad M_0 = \int_0^T m_t e^{-rt} dt.$$

The contract can be summarized by (M_0, p_0) .

In the 4th step, the buyer cannot change his investment, thus whether the seller offers a durable contract or (as we assume) a short-term contract does not make a difference. Indeed, any optimal contract has to implement the same successive identical offers.

The following sections will discuss performance of the market with respect to the first-best allocation. The analysis will focus on a bilateral relationship where the investor is the buyer. The case where the buyer invests before any tariff is proposed is discussed in Subsection 3.3. The case where the seller invests is discussed in Subsection 3.5.

3 Two-part tariffs

3.1 Unlimited commitment ($T = +\infty$)

Implementation of the first-best is straightforward. If $p_0 = c$ and the fixed fee is reasonable (i.e. such that the buyer participates), the buyer invests A^* . To extract the surplus, the

²Liquidity constraints on either player could limit this freedom, which we do not consider.

seller must choose

$$(8) \quad M_0 = \frac{1}{r} (u(A^*) - (c + rk)A^*),$$

which is the present value of the perpetual flow $u(A^*) - (c + rk)A^*$.

This is the generalization of the well-known static result. Note that implementation of the first best via unlimited two-part tariffs doesn't depend on the demand function.

3.2 Limited commitment ($0 < T < +\infty$)

At date T the investment A is sunk and nothing can prevent the seller from exerting hold-up on the buyer by setting at each instant a tariff that captures the entire surplus from the relationship. For $t \geq T$, any tariff scheme (m_t, p_T) such that $p_T \leq P(A)$ (to avoid under-consumption) and $m_t + (p_T - c)A = u(A)$ (to ensure participation) will do. A simple example is $p_T = c$ combined with a fixed fee $m_t = u(A), \forall t \geq T$.

The buyer anticipates that he will obtain no surplus from T on. Assume the price during the first period is constant and denoted by p_0 . After rearrangement, the buyer solves:

$$(9) \quad \max_A \frac{1-e^{-rT}}{r} \left(u(A) - \left(p_0 + \frac{rk}{1-e^{-rT}} \right) A \right) - M_0,$$

where the expression reveals that, due to expropriation ex post, the investment A has to be amortized during the first period. The equivalent flow cost of investment is thus $\frac{rk}{1-e^{-rT}}$ per unit of equipment.

The buyer will invest

$$(10) \quad A = Q \left(p_0 + \frac{rk}{1-e^{-rT}} \right),$$

under the condition that M_0 is small enough to ensure participation.

The seller anticipates that he can obtain the total surplus from T on, whose present value is $\frac{e^{-rT}}{r}(u(A) - cA)$. He solves

$$(11) \quad \begin{aligned} \max_{M_0, p_0} \Pi(M_0, p_0) &= M_0 + \frac{1-e^{-rT}}{r} (p_0 - c)Q \left(p_0 + \frac{rk}{1-e^{-rT}} \right) \\ &\quad + \frac{e^{-rT}}{r} \left[u \left(Q \left(p_0 + \frac{rk}{1-e^{-rT}} \right) \right) - cQ \left(p_0 + \frac{rk}{1-e^{-rT}} \right) \right] \\ \text{s.t. } M_0 &\leq \frac{1-e^{-rT}}{r} \left[u \left(Q \left(p_0 + \frac{rk}{1-e^{-rT}} \right) \right) - \left(p_0 + \frac{rk}{1-e^{-rT}} \right) Q \left(p_0 + \frac{rk}{1-e^{-rT}} \right) \right]. \end{aligned}$$

Clearly, the seller will choose the largest possible fixed fee and the corresponding profit-maximizing price.

Proposition 1. *When the seller can commit to a two-part tariff (M_0, p_0) for a limited*

duration T , he will set a marginal price below his marginal cost:

$$(12) \quad p_0 = c - \frac{e^{-rT}}{1-e^{-rT}}rk,$$

This leads the buyer to undertake the optimal investment $A^* = Q(c + rk)$. The fixed fee allows the seller to capture the entire first-stage surplus:

$$(13) \quad M_0 = \frac{1-e^{-rT}}{r} (u(A^*) - (c + rk)A^*).$$

Then the following tariff allows him to capture the entire second-stage surplus:

$$(14) \quad \text{For all } t > T, \quad m_t = u(A^*) \text{ and } p_T = c.$$

An interesting feature of this game is that the consumption good is subsidized during the contract: it is priced below its marginal cost, the subsidy being the amortization of capital from date T on, i.e. it is an advance compensation for the hold-up period. The shorter the contract duration and the larger the investment cost, the more generous the first-stage unit price reduction must be in order to stimulate investment. Remark that p_0 is negative if T is sufficiently short.

The two instruments of the contract play different roles: p_0 is used to give the right investment incentives to the buyer, and M_0 transfers the surplus to the seller. The seller is willing to set a below-cost unit price in order to encourage investment because this does not prevent him from capturing the entire surplus from the relationship, through the fixed fee in the first stage, then by exerting hold-up after expiry of the contract.

3.3 No commitment

Suppose now the seller cannot commit to a tariff (M_0, p_0) . Strategically speaking, this corresponds to the scenario where the buyer invests before any tariff is proposed. This is a situation of pure hold-up: once the investment is made, the seller will repeatedly set the same tariff (m, p) that captures the entire *ex post* surplus for each period, which means that the buyer cannot recoup his investment costs. As a consequence, the buyer's surplus if he invests is negative. Therefore, there is no equilibrium with positive investment: no commitment and powerful tools is the worst-case scenario.

3.4 Endogenous duration

The first-best investment level can be attained whatever $T > 0$. The interest of the alternative scenarios with respect to T lies in the timing of transfers and the structure of tariffs. In all cases, the whole surplus goes into the seller's hands. In the analysis, the contract duration T is treated as exogenous. Now suppose the seller or the buyer has

the power to determine the duration of the contract. All durations $T > 0$ are equivalent for both players, while no commitment is inefficient, thus the choice will be any strictly positive duration.

3.5 Investment by the seller

Suppose the investment is made by the seller. Fixed fees allow a monopolistic seller to capture the entire surplus both during and after the contract, therefore it is optimal for him to choose the first-best investment level A^* and to always set the variable price $p = c + rk$ to induce a consumption equal to A^* . In fact, whether the investment is undertaken by the seller or the buyer, two-part tariffs yield the efficient level of investment and allocate the surplus to the seller.

4 Linear tariffs with expiry date

Before studying the structure of the equilibrium with linear tariffs, we give two theoretical reasons why they are important in practice.

For the first argument, there is a basic result on two-part tariffs that must be recalled. Take the static standard monopoly. Suppose that the *fixed* part of a two-part tariff is capped by regulation, and let's see what happens when the cap goes down.

- As long as the cap is high enough, the constraint doesn't matter. The seller sets a marginal price equal to the marginal cost and extracts the entire surplus through the fixed fee. This is the best case for the seller and it is socially optimal.
- When the cap on the fixed part is binding and goes down, the seller inevitably decreases the fixed fee and raises the marginal price above marginal cost. The buyer's consumption decreases and his surplus stays at zero.
- As the cap on the fixed part falls further down, the seller fixes the price at the standard monopoly level *before* being forced to a pure linear price. He continues to set the highest permissible fee.

The pure linear price is a limit case of the latter regime with the cap at zero. In that important regime, capped two-part tariffs all yield the same marginal price, quantity and investment level as linear tariffs. Only the surplus sharing is affected via the cap-dependent fixed fee (the buyer's surplus increases as the cap decreases). Therefore studying linear prices is not as restrictive as it may seem since the predictions encompass the case of two-part tariffs with relatively tight caps on fixed fees.

This being recalled, a defense of linear pricing relies on its distributive role: it guarantees the buyer a share of the surplus. Linear pricing may emerge as a general rule

for a wide set of contexts imposed in the buyer's interest. In the absence of information on future costs or preferences, it is extremely robust as a protection against full rent extraction.

4.1 A general characterization

We assume now that, at all dates t , the buyer has to pay a linear price p_t , and that prices are guaranteed until date $T > 0$ only.

A priori, during the first stage $t \in [0, T]$, different prices could be used at different dates. In fact, the following lemma proves that this is not necessary:

Lemma 1. *If the seller uses linear prices at each date, it is always optimal for him to restrict himself to a constant price during the commitment period.*

In the first stage (from $t = 0$ to T) the price is p_0 and consumption is q_0 , while in the second stage (from $t = T$ to $+\infty$) the price is p_T and consumption is q_T . The game is sequential, with successive decisions p_0, A, q_0, p_T, q_T . As usual, it will be solved by backward induction, but a second lemma will eliminate further large classes of dominated strategies.

The buyer can either passively invest the capacity corresponding to the monopoly quantity, or use his investment choice to influence the outcome. In addition, the seller can either passively set p_0 at the anticipated marginal willingness to pay of the buyer, or use p_0 strategically to influence his investment behavior, typically to stimulate investment.

This defines the following alternative behaviors. We shall see what they imply and which prevail in equilibrium.

Definition 1. *The buyer is said to be active when his investment choice A induces a response p_T from the seller that differs from the unconstrained monopoly price $\frac{\varepsilon c}{\varepsilon - 1}$. Otherwise he is said to be passive.*

Definition 2. *The seller is said to be active when his price choice p_0 induces a response A from the buyer that differs from $A = Q(p_0)$. Otherwise he is said to be passive.*

To begin with, we will prove that the investment is never oversized in the second stage, and that the second-stage price equates demand with capacity. Otherwise, the seller could slightly adjust his first-period price without affecting the second period; setting the price p_0 closer to the unconstrained monopoly price would enhance profits. The next lemma also proves that the investment can never be oversized in the first stage either.

Lemma 2.

1. *When the buyer is active, $p_T = P(A) > \frac{\varepsilon c}{\varepsilon - 1}$.*

2. When the buyer is passive, $p_0 = p_T = P(A) = \frac{\varepsilon c}{\varepsilon - 1}$, and the seller is also passive.
3. At equilibrium, $A \leq Q(p_0)$.

The lemmas enable us to summarize the four candidate regimes in Table 1.

		Seller	
		Active	Passive
Buyer	Active	$p_0 < p_T$ $p_T > \frac{\varepsilon c}{\varepsilon - 1}$ $A = Q(p_T) < Q(p_0)$	$p_0 = p_T$ $p_T > \frac{\varepsilon c}{\varepsilon - 1}$ $A = Q(p_0) = Q(p_T)$
	Passive	IMPOSSIBLE	$p_0 = p_T$ $p_T = \frac{\varepsilon c}{\varepsilon - 1}$ $A = Q\left(\frac{\varepsilon c}{\varepsilon - 1}\right)$

Table 1: The 4 types of equilibrium.

Theorem 1 (Equilibrium prices in the general case).

- Ⓐ If $\frac{c}{rk} \geq (\varepsilon - 1)e^{rT}$, both parties are passive and

$$p_0 = p_T = \frac{\varepsilon c}{\varepsilon - 1}.$$

- Ⓑ If $(\varepsilon - 1)e^{rT} - \frac{1}{\varepsilon} \leq \frac{c}{rk} < (\varepsilon - 1)e^{rT}$, the buyer is active and the seller is passive, and

$$p_0 = p_T = re^{rT}\varepsilon k.$$

- Ⓒ If $\frac{c}{rk} < (\varepsilon - 1)e^{rT} - \frac{1}{\varepsilon}$, both parties are active and

$$\begin{cases} p_0 &= \frac{1}{1 - e^{-rT}} \left[\left(1 - \frac{\varepsilon e^{-rT}}{\varepsilon + e^{-rT}}\right) \frac{\varepsilon c}{\varepsilon - 1} + \left(1 - \frac{\varepsilon^2 e^{-rT}}{\varepsilon + e^{-rT}}\right) \frac{rk}{\varepsilon - 1} \right], \\ p_T &= \frac{\varepsilon^2(c + rk)}{(\varepsilon + e^{-rT})(\varepsilon - 1)}. \end{cases}$$

Since we are mostly interested in the situation where both players use their opportunities of strategically influencing the other party's behavior, we will concentrate the analysis in the text on the case where the seller and the buyer are active. For the other cases, Figure 1 illustrates how changing T can make the equilibrium prices cross all three regimes (details are given in subsection A.5).

From the theorem, we see that a necessary and sufficient condition for both players to be active for all T is

$$(15) \quad \frac{c}{rk} < \varepsilon - 1 - \frac{1}{\varepsilon}.$$

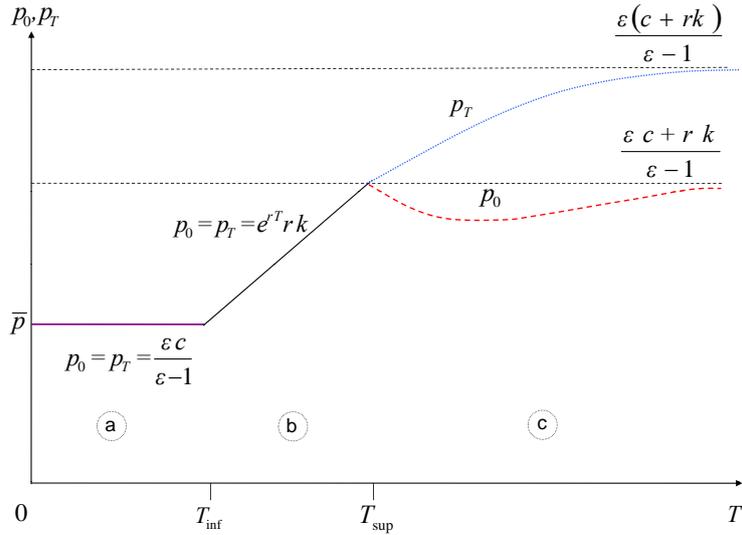


Figure 1: Equilibrium prices as a function of the contract duration (when $\frac{c}{rk} > \varepsilon - 1$).

Note that for the right hand-side to be positive, the elasticity of demand must exceed $\frac{1+\sqrt{5}}{2}$, the Golden Ratio ($\simeq 1.6$). Equilibrium prices as a function of T are given in Figure 2.

4.2 The value of commitment

One could think that longer commitment periods are necessarily better for the seller: any price profile that can be implemented with a short commitment duration (say T_1) can be implemented with a longer commitment duration (say T_2). But from this, it doesn't follow that the (so to speak) T_2 -buyer will replicate the behavior of the (so to speak) T_1 -buyer.

Assume that the T_1 -buyer and the T_2 -buyer face the same price profile and that they consider investing A . How do they value a marginal increase in A ? Identically from $t = 0$ to $t = T_1$: price is fixed anyway. Identically also from $t = T_2$ on: the post-commitment price is a function of capacity only. Now, an increase in capacity will yield a price reduction between T_1 and T_2 for the T_1 -buyer, whereas the price remains constant in that interval for the T_2 -buyer. This means that the value of increasing A is necessarily *less* for the T_2 -buyer than for the T_1 -buyer. There is no ambiguity: facing the same price profile, the T_2 -buyer (whose contract is longer) invests less than the T_1 -buyer. This shows that a mere replication doesn't work.

Since $A = Q(p_T)$ is a decreasing function of p_T , which (from Theorem 1) is an increasing function of T in all regimes, the following paradoxical result obtains.

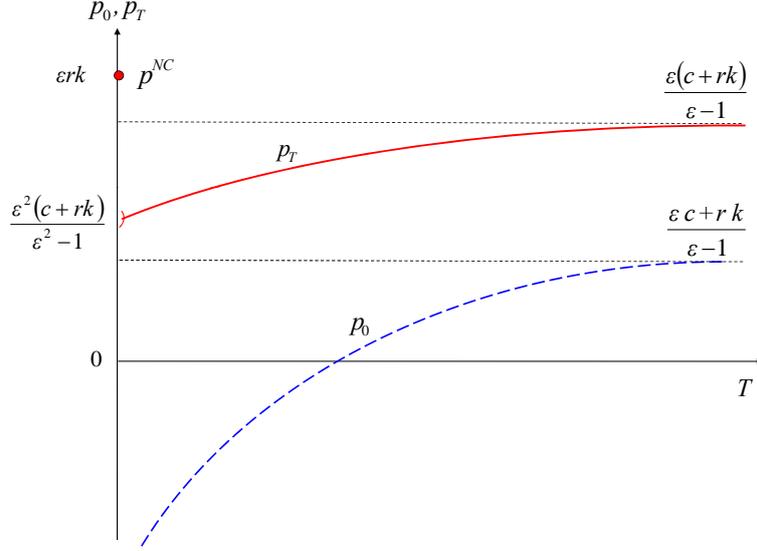


Figure 2: Equilibrium prices as a function of the contract duration (when $\frac{c}{rk} < \varepsilon - 1 - \frac{1}{\varepsilon}$). NC corresponds to the no-commitment scenario (subsection 4.3).

Proposition 2 (Investment). *The equilibrium investment level decreases with respect to the contract duration for all $T > 0$.*

Since the seller cannot commit to refrain from extracting rent after contract expiry, his only means to stimulate investment is to offer at $t = 0$, before A is chosen, the guarantee of a low price until T . The smaller T , the more generous the bargain must be: p_0 can even be negative, i.e. consumption is subsidized during the contract when the contract duration is short. For example, when T is close to zero but still strictly positive, p_0 becomes infinitely negative in the case where $\frac{c}{rk} < \varepsilon - 1 - \frac{1}{\varepsilon}$ (see Figure 2).

Remark. Intuitively, one would think that the equilibrium prices are increasing functions of the investment cost, since a higher k tends to dampen investment and consumption. This is true for the hold-up price p_T , but not necessarily for the contract price. See the analysis in subsection A.6 and Figure 4 therein.

Since social welfare increases with respect to the investment level, Proposition 2 implies that welfare decreases with respect to the contract duration. As for the seller's profit and the buyer's surplus, when $\frac{c}{rk} < \varepsilon - 1 - \frac{1}{\varepsilon}$, they read respectively

$$(16) \quad \Pi = \frac{1}{r} \frac{c+rk}{\varepsilon-1} Q \left(\frac{\varepsilon^2(c+rk)}{(e^{-rT}+\varepsilon)(\varepsilon-1)} \right);$$

$$(17) \quad S = \frac{1}{r} \frac{1-e^{-rT}+\frac{1}{\varepsilon}e^{-rT}}{\varepsilon-1} \frac{\varepsilon^2(c+rk)}{(e^{-rT}+\varepsilon)(\varepsilon-1)} Q \left(\frac{\varepsilon^2(c+rk)}{(e^{-rT}+\varepsilon)(\varepsilon-1)} \right).$$

Proposition 3 (Welfare). *When $\frac{c}{rk} < \varepsilon - 1 - \frac{1}{\varepsilon}$, for all $T > 0$,*

1. Social welfare decreases with respect to T .
2. The seller's profit decreases with respect to T .
3. The buyer's surplus increases with respect to T .

See Figure 3.

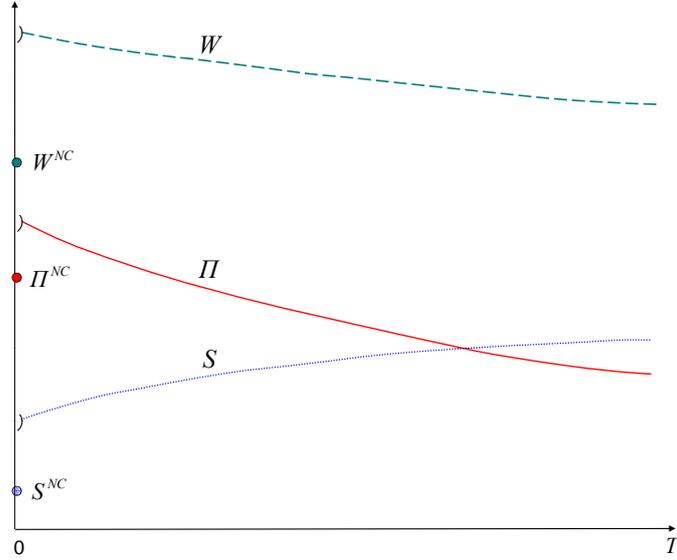


Figure 3: Profit of the seller, surplus of the buyer and social welfare as a function of the contract duration (when $\frac{c}{rk} < \varepsilon - 1 - \frac{1}{\varepsilon}$). NC corresponds to the no-commitment scenario (subsection 4.3).

With linear prices, more commitment through longer contracts is detrimental to the seller and has the effect of depressing investment. In a nutshell, the first price p_0 serves to push investment whereas p_T serves to extract rents: the two successive prices approach a two-part tariff, the first price providing the marginal incentives (investment) and the second price the inframarginal ones (participation). Though the performance of these linear prices is imperfect, this view clarifies the effects of T across agents, and for the economy, as we explain now.

At the beginning of Section 4, we recalled how a cap on the fixed fee benefits the buyer. Longer durations are similar to lower (i.e. more advantageous) caps: the longer the commitment, the smaller the impact of the hold-up period is. In compensation, as the second price loses impact when T increases, the seller has to mix in the first price two objectives (incentives and rent extraction), which is the typical inefficiency of plain static linear prices. What about the buyer? As T increases, the buyer's surplus increases even though the surplus to be shared is smaller.

4.3 No commitment

An alternative scenario which doesn't fit in the previous structure, has to be considered: no commitment at all. The timing is simplified:

1. The buyer invests A at a unit cost k .
2. At each date t , the seller makes a take-it-or-leave-it offer p_t to the buyer. If the buyer accepts, he buys q_t .

In effect, this case differs, as we shall see, from the case $T \rightarrow 0$ in the previous game. Indeed, whenever $T > 0$, and whatever small T may be, there are a price before and a price after A is chosen.

Clearly the same price and quantities (denoted respectively by p and q) will be chosen at all dates. The buyer anticipates that he will suffer hold-up as soon as the investment is realized, but extortion is imperfect as linear prices are used.

The following proposition summarizes the results.

Proposition 4 (No commitment at all). *The equilibrium price and investment level are*

$$(18) \quad p = \max \left\{ \frac{\varepsilon c}{\varepsilon - 1}, \varepsilon r k \right\},$$

$$(19) \quad A = \min \left\{ Q \left(\frac{\varepsilon c}{\varepsilon - 1} \right), Q(\varepsilon r k) \right\}.$$

Since in both cases the capacity is fully used ($q = A = Q(p)$), the outcome is ex post efficient; however, ex ante efficiency is not achieved. Indeed, the equilibrium price is always higher than the social optimum $p = c + rk$, which implies that the equilibrium investment level is always suboptimal.

No commitment and “some” commitment ($T > 0$) are qualitatively different. In other terms, to encourage investment, the smallest contract is the best (Proposition 2), no contract at all is the worst case. See NC (no commitment) prices, profits and welfare in Figures 2 and 3. In particular, note that p^{NC} is larger than any p_0 or p_T .

4.4 Investment by the seller

An alternative assumption could be that the investment is undertaken by the seller. Let us examine the following game:

1. The seller chooses the investment size A and sets price p .
2. The buyer purchases q indefinitely.

Commitment is not relevant here, because the optimal price for the seller does not change over time. There is no need to create incentives and extract rents in two separate episodes. This means that contract duration does not matter.

The seller solves the program of a monopolist whose marginal cost is $c + rk$: the equilibrium is characterized by

$$(20) \quad p = \frac{\varepsilon(c+rk)}{\varepsilon-1},$$

$$(21) \quad A = Q\left(\frac{\varepsilon(c+rk)}{\varepsilon-1}\right).$$

This game is essentially equivalent to the previous game with $T = +\infty$ with buyer investment: the outcomes (investment level, profit and surplus of the parties) are the same. The seller simply increases the commodity price by rk to pass on his investment cost.

Does this solve the seller's problem? In fact, the seller would prefer to let the buyer invest. The reasoning is as follows. If the seller invests, he can invest at the level that he wishes, but the linear prices limit his ability to extract the rent thus generated. Since the investment cost is difficult to recoup, his incentive to invest in the first place is limited. Conversely, if the buyer invests, the linear prices protect his effort (he is able to keep a fraction of the surplus generated). So the incentives that the seller gives to the buyer are leveraged by the buyer's own incentives, which makes the seller's contribution to investment (through rebates) relatively small. This allows the seller to extract a bigger rent through the second price p_T .

5 Conclusion

In this paper, we have examined a situation where contracts are incomplete in several respects: both the pricing structure and the duration of commitment are restricted. All of the pricing schemes that we have analyzed share the following feature: some commitment is always better than none. The reason is intuitive: the seller can offer a significant rebate that gives the buyer perfect or at least fairly good incentives to invest. The commitment period secures the indispensable thrust.

A longer contract never gives higher investment incentives. Tariffs with a sufficient "width", such as two-part tariffs, are powerful enough to make the contract length irrelevant. With linear tariffs, the investment level is a strictly decreasing function of the contract length. When the contract duration is very short, the seller chooses to subsidize investment heavily to maximize the surplus that will subsequently be extracted. A longer commitment postpones the hold-up period, and the seller prefers to extract rent during the contract period by setting higher prices, with a negative effect on investment incentives. Social welfare, the seller's profit, and the buyer's surplus do not vary in parallel when the contract length changes. Contrary to the seller's profit and social welfare, the buyer's surplus increases with length. A long commitment period means that there is a cap on the rent extracted by the seller, which benefits the buyer.

The link between contract duration and investment incentives in a noncompetitive environment has not been thoroughly established by economic theory. One reason is that duration is not just another parameter in a contract, complete or incomplete. Its interaction with other dimensions can be quite counterintuitive. Our model offers a plausible theoretical interpretation for empirical results that link contract duration and investment incentives. In an empirical study of coal contracts, Joskow (1987) finds evidence of a positive relationship between asset “specificity” —as measured by different proxies— and contract length (see also Neumann and Hirschhausen, 2008). This theory is not developed beyond the reasoning that specificity is a weakness and that duration is a protection. Specificity is weaker if the seller can use the productive capacity to deliver the product to several buyers. In this case, our model suggests that the seller’s market power can be used to seek the shortest possible contracts. Conversely, if the asset is very specific, for example, if there is a single buyer, market power is rebalanced in favor of the buyer, and the seller must concede a longer contract.

References

- Borenstein, Severin, Jeff Mackie-Mason and Janet Netz (2000): “Exerting Market Power in Proprietary Aftermarkets”, *Journal of Economics and Management Strategy*, 9(2), 157-88.
- Carlton, Dennis and Michael Waldman (2010): “Competition, Monopoly, and Aftermarkets”, *Journal of Law, Economics & Organization*, 26, 54-91.
- Castaneda, Marco A. (2006): “The Hold-up Problem in a Repeated Relationship”, *International Journal of Industrial Organization*, 24(5), 953-970.
- Chen, Zhiqi, Thomas Wayne Ross and William T. Stanbury (1998): “Refusals to Deal and Aftermarkets”, *Review of Industrial Organization*, 13, 131-151.
- Crawford, Vincent P. (1988): “Long-Term Relationships Governed by Short-Term Contracts”, *American Economic Review*, 78, 485-499.
- Creti, Anna and Bertrand Villeneuve, Bertrand (2004): “Long-term Contracts and Take-or-pay Clauses in Natural Gas Markets”, *Energy Studies Review*, 13, 75-94.
- Crocker, Keith J. and Scott E. Masten (1988): “Mitigating Contractual Hazards: Unilateral Options and Contract Length”, *RAND Journal of Economics*, 19(3), 327-343.
- Goldberg, Victor P. (1976): “Regulation and Administered Contracts”, *Bell Journal of Economics*, 7(2), 426-448.
- Guriev, Sergei and Dmitriy Kvasov (2005): “Contracting on Time”, *American Economic Review*, 95, 1369-1385.
- Joskow, Paul L. (1987): “Contract Duration and Relationship-Specific Investments: Empirical Evidence from Coal Markets”, *American Economic Review*, 77(1), 168-185.
- MacLeod, W. Bentley and James M. Malcomson (1993): “Investments, Holdup, and the Form of Market Contracts”, *American Economic Review*, 83, 811-837.
- Morita, Hodaka and Michael Waldman (2004): “Durable Goods, Monopoly Maintenance, and Time Inconsistency”, *Journal of Economics and Management Strategy*, 13(2), 273-302.
- Neumann, Anne and Christian von Hirschhausen (2008): “Long-Term Contracts and Asset Specificity Revisited: An Empirical Analysis of Producer-Importer Relations in the Natural Gas Industry”, *Review of Industrial Organization*, 32(2), 131-143.
- Reitzes, James D. and Glenn A. Wroch (2008): “Competition for Exclusive Customers: Comparing Equilibrium and Welfare Under One-Part and Two-Part Pricing”, *Canadian Journal of Economics/Revue canadienne d'économique*, 41(3), 1046-1086.
- Rey, Patrick and Bernard Salanié (1990): “Long-term, Short-term and Renegotiation: On the Value of Commitment in Contracting”, *Econometrica*, 58, 597-619.
- Shapiro, Carl (1995): “Aftermarkets and Consumer Welfare: Making Sense of *Kodak*”, *Antitrust Law Journal*, 63, 483-511.
- Williamson, Oliver E. (1971): “The Vertical Integration of Production: Market Failure Considerations”, *American Economic Review, Papers and Proceedings*, 61(2), 112-123.

A Appendix

A.1 Proof of Proposition 1

Proof in the text.

A.2 Proof of Lemma 1

Consider a contract of length T , such that the producer commits to a price profile $\{p_t\}_{0 \leq t \leq T}$. We assume that for all real number x , sets of dates like $\{t | p_t \leq x\}$ are Lebesgue-measurable (a measure denoted by d), which is not restrictive in our context and it is technically more rigorous for the argument. The reasoning will be developed in three steps.

Step 1. First we shall prove that at equilibrium $A \leq Q\left(\frac{\varepsilon c}{\varepsilon - 1}\right)$. Suppose instead that $A > Q\left(\frac{\varepsilon c}{\varepsilon - 1}\right)$: then the post-commitment price is $p_T = \frac{\varepsilon c}{\varepsilon - 1}$, and A is oversized in the period following contract expiry. We denote by I the measurable set of dates in $[0, T]$ for which $p_t \leq p(A)$ (i.e. those dates for which consumption equals A).

(i) All prices above $p(A)$ (i.e. for $t \in [0, T] \setminus I$) can be changed with profit to $\frac{\varepsilon c}{\varepsilon - 1}$: some increase and the others decrease, depending on their position relative to $\frac{\varepsilon c}{\varepsilon - 1}$. These changes have no effect on the choice of A , consumption at those dates being unconstrained by A before and after the change for all $t \in [0, T] \setminus I$. If $d(I) = 0$, then we have shown that the price schedule was not optimal. If $d(I) > 0$, we have to proceed with the proof.

(ii) We denote the “weight” of I by $\mu(I) = \int_I e^{-rt} dt$. Prices in I can be replaced by the weighted mean

$$(22) \quad p_I = \frac{\int_I p_t e^{-rt} dt}{\mu(I)},$$

without changing the seller’s profits or the buyer’s investment. Indeed, p_I is in the convex hull of $\{p_t | t \in I\}$, and thus verifies $Q(p_I) \geq A$: the constraint is still binding.

(iii) We have established in (i) and (ii) that any price profile can be replaced (with profit) with another price profile where price is a constant $\frac{\varepsilon c}{\varepsilon - 1}$ over $[0, T] \setminus I$ and a constant $p_I < \frac{\varepsilon c}{\varepsilon - 1}$ over I .

The buyer will equate q_t with A for $t \in I$ only, and will be strictly unconstrained at other dates, so that his program can be rewritten, for the only part that depends on A :

$$(23) \quad \max_A \mu(I) \cdot \left[u(A) - \left(p_I + \frac{k}{\mu(I)} \right) A \right].$$

At equilibrium the buyer’s best choice is

$$(24) \quad q_t = A = Q\left(p_I + \frac{k}{\mu(I)}\right).$$

If he anticipates this reaction, the seller wants to approach as much as possible the interior solution of the corresponding profit maximizing program. This interior solution is

$$(25) \quad \frac{\varepsilon c + \frac{k}{\mu(I)}}{\varepsilon - 1} > \frac{\varepsilon c}{\varepsilon - 1}.$$

In any case, the initial pricing choice appears not to be optimal. This proves step 1 by contradiction.

Step 2. Then we prove that consumption cannot be lower than chosen capacity during the contract period, in other terms that $q_0 = A$. Indeed, suppose that for some t in $[0, T]$, $q_t < A$, which means that $Q(p_t) < A$. As $A \leq Q\left(\frac{\varepsilon c}{\varepsilon - 1}\right)$, necessarily, $p_t > \frac{\varepsilon c}{\varepsilon - 1}$. A reduction in p_t is beneficial for the producer in the short run, because the current price will be brought closer to the monopoly price so that profit will increase, and it will have no detrimental effect in the long run, because investment will not be affected. Indeed, as long as capacity is not binding at time t , the current price has no other effect than on current consumption. As a consequence, any contract price profile offered by the producer is necessarily such that the buyer will consume A for all t in $[0, T]$.

Step 3. Finally, we prove that the seller can never be worse off by restricting himself to a constant contract price. Since $\{p_t\}_{0 \leq t \leq T}$ is such that $q_t = A$ for all t in $[0, T]$, any profile $\{p'_t\}_{0 \leq t \leq T}$ included in the convex hull of $\{p_t | t \in I\}$, and such that

$$(26) \quad \int_0^T p'_t e^{-rt} dt = \int_0^T p_t e^{-rt} dt,$$

will produce the same choice A by the buyer and the same consumption at each date.

In effect, the buyer's first-order condition w.r.t. A is

$$(27) \quad \int_0^T (u'(A) - p_t) e^{-rt} dt + \frac{d}{dA} \int_T^{+\infty} (u(A) - p(A)A) e^{-rt} dt - k = 0,$$

which is also true if $\{p'_t\}_{0 \leq t \leq T}$ is used instead of $\{p_t\}_{0 \leq t \leq T}$. Thus the producer's profit would be the same if he had set the following constant price:

$$(28) \quad p_0 = \frac{\int_0^T p_t e^{-rt} dt}{\int_0^T e^{-rt} dt}.$$

To conclude, the largest possible profit can always be achieved with a constant price.

A.3 Proof of Lemma 2

The proof of this lemma is derived from Lemma 1's.

1. From Step 1, $A \leq Q\left(\frac{\varepsilon c}{\varepsilon - 1}\right)$. We know that $p_T = \max\left(\frac{\varepsilon c}{\varepsilon - 1}, P(A)\right)$. Therefore $p_T = P(A) \geq \frac{\varepsilon c}{\varepsilon - 1}$. Finally, since the buyer is active, the last inequality is strict.

2. Let's denote $Q\left(\frac{\varepsilon c}{\varepsilon - 1}\right)$ by \bar{A} .

If the seller is passive, then by definition, $A = \bar{A}$ and $p_T = \frac{\varepsilon c}{\varepsilon - 1}$. We reason by contradiction by assuming that $p_0 < p_T$ (which implies in turn that $q_0 = A$).

Given that $A = \bar{A}$, the maximum p_0 such that the buyer doesn't want to reduce his investment is defined by the first-order condition

$$(29) \quad \int_0^T (u'(A) - p_0) e^{-rt} dt + \frac{d}{dA} \int_T^{+\infty} (u(A) - p(A)A) e^{-rt} dt - k = 0$$

evaluated for $A = \bar{A}$.

This maximum price is denoted by p_0^* . Its expression is much simplified by using two facts: first, $u'(\bar{A}) = p(\bar{A}) = \frac{\varepsilon c}{\varepsilon - 1}$; second, $p'(\bar{A})A = -p(\bar{A})/\varepsilon$. Equation (29) gives

$$(30) \quad p_0^* = \frac{\varepsilon c}{\varepsilon - 1} + \frac{e^{-rT}}{\varepsilon(1 - e^{-rT})} \left(\frac{\varepsilon c}{\varepsilon - 1} - e^{rT} r \varepsilon k \right).$$

Clearly, if $p_0 < p_0^*$, then p_0 is not optimal for the seller: increasing it will not decrease investment A (thus second-period profits are preserved) and first-period consumption stays at \bar{A} with a higher price (profits are increased). So we have proved that $p_0 \geq p_0^*$ for the rest of the proof.

If $p_0^* \geq \frac{\varepsilon c}{\varepsilon - 1}$, then $p_0 \geq \frac{\varepsilon c}{\varepsilon - 1}$, which we excluded by assumption.

If $p_0^* < \frac{\varepsilon c}{\varepsilon - 1}$, we have to show that the seller wants a higher p_0 than p_0^* (and not just stay at $p_0 = p_0^*$), which implies that A will not stay at \bar{A} .

When p_0 passes from p_0^* to $p_0^* + dp_0$, given that A will exhibit a first-order variation (it diminishes), the variation of the second-period profit is of second order (we had a repetition of the static maximum profit). To reach a conclusion, all we need is to evaluate the variation of profit for the first period. It is proportional to

$$(31) \quad \bar{A} + (p_0^* - c) \frac{dA}{dp_0},$$

for $A = \bar{A}$ and $p_0 = p_0^*$.

The implicit definition of $A(p_0)$ is given by (29). After rearrangement of the terms, this gives

$$(32) \quad p(A) \left[1 + \frac{1}{\varepsilon} \frac{e^{-rT}}{1 - e^{-rT}} \right] - \frac{rk}{1 - e^{-rT}} = p_0.$$

We can use again $u'(A) = p(A) = \frac{\varepsilon c}{\varepsilon - 1}$ and $p'(A)A = -p(A)/\varepsilon$ on this implicit definition of $A(p_0)$ to find $\frac{dA}{dp_0}$. This yields

$$(33) \quad \frac{dA}{dp_0} = \frac{1}{p'(A) \left[1 + \frac{1}{\varepsilon} \frac{e^{-rT}}{1 - e^{-rT}} \right]}.$$

The sign of (31) can now be computed: it is also the sign of

$$(34) \quad p_0^*(1 - \varepsilon) + \frac{rk}{1 - e^{-rT}} + \varepsilon c.$$

Clearly, it is positive if $p_0^* < \frac{\varepsilon c}{\varepsilon - 1} + \frac{rk}{1 - e^{-rT}}$, which is the case we are in. Therefore the seller doesn't want to choose $p_0 = p_0^*$. We have the contradiction proving that $p_0 \geq p_T = \frac{\varepsilon c}{\varepsilon - 1}$.

The case $p_0 > \frac{\varepsilon c}{\varepsilon - 1}$ can be excluded since p_0 would be uselessly too large: lowering it to $\frac{\varepsilon c}{\varepsilon - 1}$ would not decrease investment and would allow the seller to cash in the monopoly surplus all over the life of the investment.

3. When the seller is passive, by definition $A = Q(p_0)$. When the buyer is passive, from Lemma 2 (point 2), $A = Q(p_0)$. When both the buyer and the seller are active, $A = Q(p_T) \neq Q(p_0)$. We will now prove by contradiction that $A > Q(p_0)$ cannot be an equilibrium. Indeed, it implies $p_0 > p_T > \frac{\varepsilon c}{\varepsilon - 1}$. But it is clearly profitable for the buyer to reduce p_0 : in the first period where $q_0 = Q(p_0)$ he would get closer to his static monopoly profit, and in addition, this would induce the buyer to invest more, bringing second-period consumption closer to the seller's preferred level \bar{A} . Therefore, in an equilibrium with active buyer and active seller, necessarily $A < Q(p_0)$.

A.4 Proof of Theorem 1

Whenever the buyer is active, he solves:

$$(35) \quad \max_A \frac{1-e^{-rT}}{r} (u(A) - p_0 A) + \frac{e^{-rT}}{r} (u(A) - P(A)A) - kA,$$

where the first term is the present buyer utility from the contract, the second term is the present utility for the subsequent period, and the third is investment cost. The fact that the buyer is active enables us to replace p_T with $P(A)$; Lemma 2 (1.) enables us to replace consumption in the first period with A .

We find

$$(36) \quad A = Q \left(\frac{p_0 + \frac{r}{1-e^{-rT}} k}{1 + \frac{1}{\varepsilon} \frac{e^{-rT}}{1-e^{-rT}}} \right).$$

Since the seller is active, $p_0 < p_T$: the seller offers a lower price in the first period so that the buyer is incited to invest more, then once the capacity is fixed he sets a higher price. To choose p_0 , the seller solves

$$(37) \quad \max_{p_0} \left[\frac{1-e^{-rT}}{r} (p_0 - c) + \frac{e^{-rT}}{r} \left(\frac{p_0 + \frac{r}{1-e^{-rT}} k}{1 + \frac{1}{\varepsilon} \frac{e^{-rT}}{1-e^{-rT}}} - c \right) \right] Q \left(\frac{p_0 + \frac{r}{1-e^{-rT}} k}{1 + \frac{1}{\varepsilon} \frac{e^{-rT}}{1-e^{-rT}}} \right).$$

This yields the optimal contract price

$$(38) \quad p_0 = \frac{1}{1-e^{-rT}} \left[\left(1 - \frac{\varepsilon e^{-rT}}{\varepsilon + e^{-rT}} \right) \frac{\varepsilon c}{\varepsilon - 1} + \left(1 - \frac{\varepsilon^2 e^{-rT}}{\varepsilon + e^{-rT}} \right) \frac{rk}{\varepsilon - 1} \right].$$

Finally, we need to check under which conditions the buyer and the seller are actually active. To ensure $A < Q(\frac{\varepsilon c}{\varepsilon - 1})$ and $A < Q(p_0)$, we calculate from equations (36) and (38) that the ratio of the production and investment costs must verify respectively $\frac{c}{rk} < \varepsilon e^{rT}$ and $\frac{c}{rk} < (\varepsilon - 1)e^{rT} - \frac{1}{\varepsilon}$. Clearly the latter inequality implies the former: both players are active if and only if

$$(39) \quad \frac{c}{rk} < (\varepsilon - 1)e^{rT} - \frac{1}{\varepsilon}.$$

To complete the proof of (c) in the theorem, we use equation (36),

$$(40) \quad p_T = \frac{p_0 + \frac{r}{1-e^{-rT}} k}{1 + \frac{1}{\varepsilon} \frac{e^{-rT}}{1-e^{-rT}}}.$$

For case (b), we have both $p_0 = p_T$ and equation (40). Thus the price is $r e^{rT} \varepsilon k$. From Lemma 2 (1. and 2.), we can deduce that a necessary and sufficient condition for the buyer to be active is $p_T > \frac{\varepsilon c}{\varepsilon - 1}$.

Case (a) remains to be proved. When the buyer is passive, he knows that the seller will set $p_T = \frac{\varepsilon c}{\varepsilon - 1}$. The price prevailing after contract expiry is

$$(41) \quad p_T = \min \left\{ \frac{p_0 + \frac{r}{1-e^{-rT}} k}{1 + \frac{1}{\varepsilon} \frac{e^{-rT}}{1-e^{-rT}}}, \frac{\varepsilon c}{\varepsilon - 1} \right\}.$$

This enables us to write the following necessary and sufficient condition for having a

passive buyer:

$$(42) \quad p_0 - \frac{\varepsilon c}{\varepsilon - 1} < \frac{c - (\varepsilon - 1)e^{rT}rk}{(\varepsilon - 1)(e^{rT} - 1)},$$

Since this situation is characterized by $p_0 = \frac{\varepsilon c}{\varepsilon - 1}$, the left hand-side is equal to zero, and the following result obtains: when

$$(43) \quad \frac{c}{rk} \geq (\varepsilon - 1)e^{rT},$$

the buyer is passive—and so is the seller.

QED.

A.5 Comparative statics w.r.t. contract duration: a complement

The case where $\frac{c}{rk} < \varepsilon - 1 - \frac{1}{\varepsilon}$ was detailed in subsection 4.1. Now assume the opposite holds. Let

$$(44) \quad \begin{cases} T_{\text{inf}} &= \frac{1}{r} \ln \left(\frac{c}{(\varepsilon - 1)rk} \right), \\ T_{\text{sup}} &= \frac{1}{r} \ln \left(\frac{\varepsilon c + rk}{\varepsilon(\varepsilon - 1)rk} \right). \end{cases}$$

The three cases mentioned in Theorem 1 correspond respectively to $T \leq T_{\text{inf}}$, $T_{\text{inf}} < T \leq T_{\text{sup}}$ and $T > T_{\text{sup}}$, and to zones ①, ②, and ③ in Figure 1.

By assumption, $T_{\text{sup}} > 0$; T_{inf} is positive when $\frac{c}{rk} > \varepsilon - 1$. In what follows, we assume this is also true (else Case ① is simply empty).

For the buyer, being active, i.e. reducing A below the monopoly quantity, implies an immediate benefit (lower investment expenses), and from T on, a cost (higher p_T).

For the seller, being active, i.e. reducing p_0 below the marginal willingness to pay of the buyer during the contract, implies an immediate cost (lower first-stage profits), and from T on, a benefit (larger volumes due to increased investment).

① $T \leq T_{\text{inf}}$: both parties are passive. For the buyer, the cost of being active is coming up too early, and he prefers to stay passive and invest $A = Q\left(\frac{\varepsilon c}{\varepsilon - 1}\right)$.

② $T_{\text{inf}} < T \leq T_{\text{sup}}$: only the buyer is active. The cost of being active is delayed, it is worth reducing A : equalizing marginal cost and marginal benefit ($k = \frac{e^{-rT}}{r} \frac{P(A)}{\varepsilon}$) yields $A = Q\left(re^{rT}\varepsilon k\right)$.

③ $T > T_{\text{sup}}$: both parties are active. When T becomes large, capacity risks to decrease too much if the seller stays passive. He becomes active and offers a lower p_0 to stimulate investment. Decreasing p_0 is a good strategy as long as T is not too large. But as the contract length approaches infinity, the hold-up period becomes too remote: the seller seeks to preserve profits made during the contract by increasing p_0 .

See Figure 1.

Ⓒ $k > k_{\text{sup}}$: **both parties are active.** When k becomes sufficiently large, setting $p_0 < p_T$ in order to increase A becomes worthwhile, since the positive volume effect in both stages offsets the first-stage loss. In this case only, the contract price and the hold-up price differ. The contract price can even be a decreasing function of k , and become negative when k is high, if the contract length is short enough: the seller accepts a loss that will be compensated by larger sales volumes after contract expiry.

A.7 Proof of Proposition 2

Proof in the text.

A.8 Proof of Proposition 3

Profit of the seller. The seller's profit can be expressed as a function of the contract duration:

$$\textcircled{a} \quad T \leq T_{\text{inf}} : \quad \Pi(T) = \frac{1}{r} \frac{c}{\varepsilon-1} Q\left(\frac{\varepsilon c}{\varepsilon-1}\right);$$

$$\textcircled{b} \quad T_{\text{inf}} < T \leq T_{\text{sup}} : \quad \Pi(T) = \frac{1}{r} (re^{rT} \varepsilon k - c) Q\left(re^{rT} \varepsilon k\right);$$

$$\textcircled{c} \quad T > T_{\text{sup}} : \quad \Pi(T) = \frac{1}{r} \frac{c+rk}{\varepsilon-1} Q\left(\frac{\varepsilon^2(c+rk)}{(e^{-rT}+\varepsilon)(\varepsilon-1)}\right).$$

The seller's profit is a non-increasing function of the contract duration: he always prefers shorter contracts, but actually any contract duration between 0 and T_{inf} is equivalent for him.

Surplus of the buyer.

$$\textcircled{a} \quad T \leq T_{\text{inf}} : \quad S(T) = \frac{1}{r} \left(\frac{\varepsilon c}{(\varepsilon-1)^2} - rk \right) Q\left(\frac{\varepsilon c}{\varepsilon-1}\right);$$

$$\textcircled{b} \quad T_{\text{inf}} < T \leq T_{\text{sup}} : \quad S(T) = \frac{1}{r} \frac{1-e^{-rT} + \frac{1}{\varepsilon} e^{-rT}}{\varepsilon-1} re^{rT} \varepsilon k Q\left(re^{rT} \varepsilon k\right);$$

$$\textcircled{c} \quad T > T_{\text{sup}} : \quad S(T) = \frac{1}{r} \frac{1-e^{-rT} + \frac{1}{\varepsilon} e^{-rT}}{\varepsilon-1} \frac{\varepsilon^2(c+rk)}{(e^{-rT}+\varepsilon)(\varepsilon-1)} Q\left(\frac{\varepsilon^2(c+rk)}{(e^{-rT}+\varepsilon)(\varepsilon-1)}\right).$$

The buyer's surplus function is not monotonous with respect to T : it is first constant, then decreasing, then increasing and it tends to a finite limit. In addition, the thresholds in T are not always positive, so that depending on the value of the parameters, cases Ⓐ and Ⓑ can be empty.

First suppose $\frac{c}{rk} \leq \varepsilon - 1 - \frac{1}{\varepsilon}$: only case Ⓒ exists, and since the surplus is a strictly increasing function of T in this case, the buyer always prefers the longest possible contract.

Now if $\varepsilon - 1 - \frac{1}{\varepsilon} \leq \frac{c}{rk} \leq \varepsilon - 1$, we are either in case Ⓑ or in case Ⓒ, so the buyer's surplus is highest either for $T = 0$, or when $T \rightarrow +\infty$. Comparing the two values yields the following result: the buyer's surplus is highest for $T \rightarrow 0$ if and only if

$$(45) \quad \frac{c}{rk} \leq (\varepsilon - 1) \varepsilon^{\frac{1}{\varepsilon-1}} - 1.$$

Finally, when $\frac{c}{rk} \geq \varepsilon - 1$, all three cases exist. Clearly the buyer's surplus is maximal either for $T \rightarrow 0$, or when $T \rightarrow +\infty$. Let $y = \frac{c}{rk}$. After rearrangement, we find that the

buyer prefers the shortest possible contract whenever

$$(46) \quad \left(1 - \frac{(\varepsilon - 1)^2}{\varepsilon y}\right) \left(1 + \frac{1}{y}\right)^{\varepsilon - 1} \geq 1.$$

The derivative with respect to y of the left hand-side has a unique root $\varepsilon(\varepsilon - 1)$. The LHS is increasing then decreasing, it tends to $-\infty$ when y tends to zero, and to 1 when y tends to $+\infty$, thus it is equal to 1 for a unique value of y . Therefore there is a unique y^* such that the buyer's surplus is maximal when $T \rightarrow +\infty$ whenever $y \leq y^*$, and it is maximal for $T \rightarrow 0$ whenever $y \geq y^*$.

In addition, the LHS is higher than 1 for $y = \varepsilon(\varepsilon - 1)$, which means that $0 < y^* < \varepsilon(\varepsilon - 1)$. As a consequence, when the elasticity parameter ε is close to 1, the threshold y^* is close to zero, so that for all values of the ratio $\frac{c}{rk}$ that are bounded away from zero, the buyer prefers the shortest possible contract.

A.9 Proof of Proposition 4

The model is solved by backward induction.

Choice of q and p . From Lemma 2 (1. and 2.), $q = Q(p)$ with $p = \max\{\frac{\varepsilon c}{\varepsilon - 1}, P(A)\}$.

Choice of A . Solving backwards, we come to the investment choice of the buyer. See Figure 5. In the “active-buyer” regime the price $P(A)$ decreases with respect to A . Thus as long as this regime prevails, a larger investment yields a bigger price reduction. But when the investment attains the monopoly quantity, this effect stops because the seller would never decrease the price below the monopoly price $\frac{\varepsilon c}{\varepsilon - 1}$: it is never worth investing more than $\bar{A} = Q\left(\frac{\varepsilon c}{\varepsilon - 1}\right)$. Actually, when the investment cost is large, the buyer prefers to invest less.

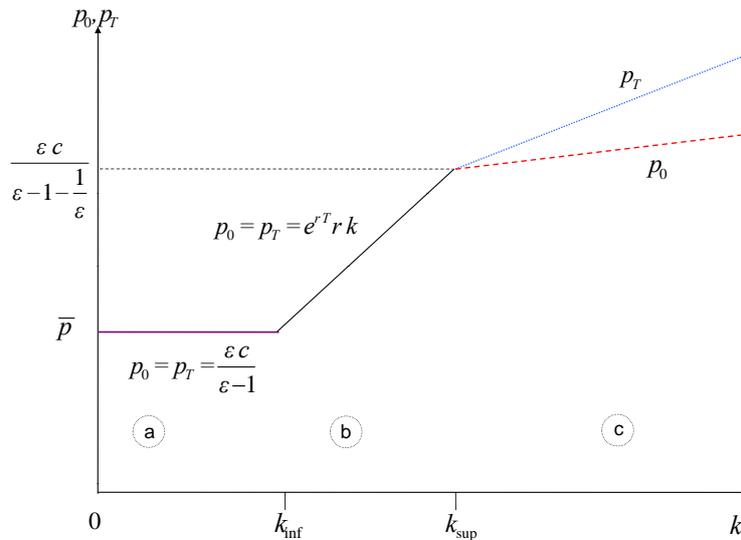


Figure 5: Impact of the investment choice on the price in model Ap with $\bar{p} = \frac{\varepsilon c}{\varepsilon - 1}$.

Thus, the buyer's surplus-maximization program can be expressed as

$$(47) \quad \begin{aligned} \max_A S(A) &= \frac{1}{r} (u(A) - (P(A) + rk)A) \\ \text{s.t. } A &\leq Q\left(\frac{\varepsilon c}{\varepsilon - 1}\right). \end{aligned}$$

The solution A is interior when the active-buyer regime prevails: the seller adjusts his price so that $q = A$. Simple calculations yield

$$(48) \quad p = \varepsilon rk.$$

This situation combines the features of monopoly and monopsony. In the regime where $p = P(A)$, the buyer chooses A as a monopsonist to maximize

$$(49) \quad S(A) = \frac{1}{r} (u(A) - P(A)A - rkA).$$

The first-order condition can be rewritten $\frac{P(A)+rk-u'(A)}{P(A)} = \frac{-P'(A)A}{P(A)} \equiv \frac{1}{\varepsilon}$. However, while in the monopsony case the price equals the marginal cost, here the seller's market power allows him to set a price equal to the marginal utility of the buyer ($P(A) = u'(A)$). Thus the equation can be rewritten $p(A) = \varepsilon rk$.

The seller is constrained by this investment choice if he cannot set his monopoly price $\frac{\varepsilon c}{\varepsilon - 1}$ because $Q\left(\frac{\varepsilon c}{\varepsilon - 1}\right) > Q(\varepsilon rk)$, or equivalently, $\frac{\varepsilon c}{\varepsilon - 1} < \varepsilon rk$. This is the case when $\frac{c}{rk} < \varepsilon - 1$.

Conversely, when $\frac{c}{rk} > \varepsilon - 1$ then $\varepsilon rk < \frac{\varepsilon c}{\varepsilon - 1}$, thus investing $Q(\varepsilon rk)$ will not prevent the seller from setting his monopoly price. The buyer would do better to invest passively $Q\left(\frac{\varepsilon c}{\varepsilon - 1}\right)$.