Strategic Capacity Investment under Holdup Threats: The Role of Contract Length and Width

Durand-Viel, Laure and Villeneuve, Bertrand

Université Paris-Dauphine, CREST

23 November 2009

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

Laure Durand-Viel*
laure.durand-viel@ensae.fr

Bertrand Villeneuve†
bertrand.villeneuve@dauphine.fr.

February 12, 2010

Abstract. This article analyzes the impact of incomplete contracts' length on investment in a bilateral relationship. The seller has the power to set the contract terms whereas the buyer decides on the investment level, which acts as a cap on future demand. Two-part tariffs succeed in implementing the optimal investment and consumption even if commitment is limited, and the length of the contract is irrelevant. Interestingly, this efficient solution is rendered possible by subsidies on consumption during the contract. In other terms, duration matters hugely for the contractual details (the timing of transfers), not for its performance. Under certain circumstances that we discuss, linear pricing may have to be used, which leads to suboptimal investment. We show that longer contracts are less efficient, meaning that a degree of completeness (pricing width) may be strictly complementary to another one (contract length). The buyer's surplus increases with respect to contract length, whereas the seller loses more in profit than the social surplus decreases. A longer contract actually protects expropriable investors rather than the investment itself.

JEL Codes: D42, D45, D92, L95.
Keywords: Long-term Contracts, Incomplete Contracting, Infrastructure Investment.

*Université Paris-Dauphine and CREST.
†Université Paris-Dauphine and CREST. This work was supported by FiME (Laboratoire de Finance des Marchés de l'Energie) and the Chaire EDF CALYON CDC "Finance et Développement Durable - Approches Quantitatives".

Contents

1 Introduction 3

2 General setting 5
2.1 Trade and payoffs . . . 5
2.2 Market power and limited commitment . . . 6

3 Two-part tariffs 6
3.1 Unlimited commitment . . . 6
3.2 Limited commitment . . . 7
3.3 No commitment . . . 8
3.4 Endogenous duration . . . 8
3.5 Investment by the seller . . . 8

4 Linear tariffs: simple cases 8
4.1 Unlimited commitment . . . 9
4.2 No commitment . . . 10
4.3 Investment by the seller . . . 12

5 Linear tariffs with expiry date 12
5.1 A general characterization . . . 12
5.2 When an active seller faces an active buyer . . . 14
5.3 Impact of the contract duration . . . 15
5.4 Investment by the seller . . . 18

6 Conclusion 18

References 19

A Appendix 21
A.1 Linear tariffs with expiry date: general case . . . 21
A.2 Comparative statics w.r.t. investment cost . . . 22
A.3 Comparative statics w.r.t. contract duration . . . 23
A.4 Profit and surplus analysis . . . 24

1 Introduction

This paper considers markets in which the short-run demand elasticity is low because demand is largely pre-determined by specific installed capital with which the buyer is caught. Investment decisions involve not only a qualitative choice (specific technology), but also a quantitative one (capacity). Oil and gas pipelines are the paragon of specific investment, where the initial capacity choice acts as a cap on volumes subsequently traded. In the short run, buyers of energy (oil, natural gas, or electricity) are locked into a previously chosen technology or standard (appliances in Balestra's and Nerlove's 1966 terminology), or tied to a specific supplier (in the case of pipelines). In the long run, fuel substitution will follow the path of capital substitution, making the elasticity much higher.

What happens if market power matters? Markups can rise high with low elasticities, but in the long run price elasticity typically increases.[1] The seller has to be careful to preserve his consumer base over time, and he must adopt reasonable tariffs if he doesn't want to kill the goose that lays golden eggs.

The model studies in detail a bilateral relationship in which a potential or actual investment is specific. Commitment is a critical issue: though reasonable prices may be announced to encourage investment in the specific appliances, these announcements are more or less credible. If commitment is limited, buyers may fear hold-up, *i.e.* unilateral price increases that will extract a rent from the installed capital.

Discrete-continuous choice models in the econometric literature on consumer demand (Dubin-McFadden 1984, Hanemann 1984) usually assume that each buyer makes essentially *one* farsighted decision, namely the technology he adopts. Balestra and Nerlove (1966) build a dynamic model where at any given point in time only one fraction of the market, the non-committed consumers (new consumers or those who replace their appliances), can choose their fuel, and once this choice is made, consumption varies freely in response to price. Not only do these models assume that production is competitive and that consumers are price-takers, they also disregard an important feature of the investment decision that constrains future consumption decisions, the capacity choice.

This feature is also neglected by industrial organization literature on standard competition or aftermarket monopolization (Shapiro 1995, Chen *et al.* 1998, Reitzes and Woroch 2008), which focuses on the optimal strategy of firms trying to attract customers who, once locked in, will be at the mercy of their market power. These models make rather extreme assumptions about the link between equipment and complementary goods. Some assume strict complementarity (one primary good for one aftermarket good, see Carlton and Waldman, 2001), so that there is one decision to be made—equipment. Alternatively, others assume independence—demand *ex post* can be adjusted freely (*e.g.* Borenstein *et al.*, 2000, Morita and Waldman, 2004). In our model, complementarity is asymmetric: the buyer cannot consume more than his capacity allows, but he is free to consume less. This is typically the case with energy-consuming appliances.

Examples could also be taken from the R&D sector. Firms investing in specific production lines that need to pay license fees for intellectual property rights gauge their investment according to the intended intensity of license use. Here again, elasticity is asymmetric: it is easier to slow down production than to exceed capacity. Another example is the franchisor/franchisee relationship: when a franchisee chooses an outlet of a given size and location, his investment caps his future sales capacity, then the payment to the franchisor (in addition to a fixed fee)

[1] See the Le Chatelier principle in Samuelson (1947).

typically consists in royalties calculated as a percentage of his turnover. The main difference with the traditional literature on incomplete contracts, investment and hold-up[2] is that investment itself is a strategy against the seller's potential abuses.

Contract incompleteness affects both the overall performance of the relationship and the distribution of the generated surplus. To pre-empt possible contract failures, some general rules can be agreed on, without compromising future negotiations about contractual details. For example, in their analysis of natural gas contracts, Crocker and Masten (1985) explain how take-or-pay provisions, by allocating discretionary power to the informed party, ensure that the optimal decisions are taken at investment *and* operations stages.[3] In our view, contract length and width can be seen as key features of the "'constitution' governing the ongoing relationship" (Goldberg, 1976, p. 428). The constitution could be either a prior agreement between the parties, or even a general purpose law governing some commercial relationship. This view motivates our comparative statics analysis of contract length in relation with restrictions on the price structure (width).

Complete contracts are, by definition, unlimited; in contrast, contracts with limited width may be optimized with respect to length in compensation. Conventional wisdom has it that long contracts protect investors. Eventually, one would expect the investment level and social welfare to rise as the commitment duration increases,[4] the extra surplus being shared between the parties depending on their bargaining powers.

Clearly, a time limit may 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). However, even in the absence of uncertainty, long incomplete contracts may have undesirable effects.

In our model, as the contract becomes longer, the buyer's incentives to invest regress while the seller's ability to extract rent diminishes. With linear prices, the negative impact on social welfare of a longer contract is very unequally shared between the parties: the buyer's surplus increases as he produces less surplus but can keep a much larger part of it, whereas the seller is a net loser.[5] A longer contract actually protects expropriable investors rather than the investment itself.

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 then characterized: we start with simple cases (full commitment, no commitment) in section 4 and we expose the case of partial commitment, *i.e.* with expiry date, in section 5. Most of the proofs are relegated to the Appendix.

[2]See Williamson (1971) for an early reference.

[3]Hubbard and Weiner (1986) offer an alternative interpretation of take-or-pay provisions in terms of risk sharing. See Creti and Villeneuve (2004) for a survey.

[4] This is the thesis in the empirical study by Joskow (1988). He finds evidence of a positive relationship between asset "specificity" (as measured by different proxies) and contract length. See also Neumann and Hirschhausen (2008).

[5]Castaneda (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 upfront a lump sum and that the seller invests after acceptance. The buyer thus is reluctant to accept long contracts that would establish extortion for a long time, meaning in turn that the seller's return on them is poor, which depresses investment.

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

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

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:

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

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:

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

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

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

The intertemporal profit of the seller is

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

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

$$A^* = Q(c + rk). \tag{6}$$

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.

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 payment[6]

$$M_0 = \int_0^T m_t e^{-rt} \, dt. \tag{7}$$

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

In the 4[th] 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 seller invests will also be discussed.

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 seller must choose

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

which is the present value of the perpetual flow $u(A^*) - (c + rk)A^*$. This is the generalization of the well-known static result.

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

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:

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

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

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

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

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

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:*

$$p_0 = c - \tfrac{e^{-rT}}{1-e^{-rT}} rk, \tag{12}$$

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:

$$M_0 = \tfrac{1-e^{-rT}}{r}\left(u(A^*) - (c + rk)A^*\right). \tag{13}$$

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

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

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 ($T = 0$)

Suppose now the seller cannot commit to a tariff (M_0, p_0) (in other words, $T = 0$). This is the pure hold-up situation: 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, if the seller cannot commit to a unit price before the buyer invests, 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 lie 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 contract's duration. All durations $T > 0$ are equivalent for both players, while $T = 0$ 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 has no bearing on the outcome.

4 Linear tariffs: simple cases

Let's recall first a basic characteristic of static monopolistic markets. Let's suppose the fixed part of a two-part tariff is capped by regulation, and let's see what happens when the cap goes down. Obviously, if the cap is high enough, the seller achieves the first-best by setting a price equal to his marginal cost and extracts the entire surplus through the fixed fee (see previous sections). If the cap on the fixed part is lower, the seller must decrease the fixed fee and raise the marginal price above marginal cost (the buyer's surplus is still zero). If the cap on the fixed part falls further, the seller continues to set the highest possible fee but stops raising the marginal price. This happens when the marginal price has attained the optimal linear price, usually called monopoly price.

The latter case is the most important for our point as it can be easily extrapolated: tightly capped two-part tariffs yield the same prices, quantities and investment level as linear tariffs, only the surplus sharing is affected. In other words, the marginal effects of the cap and its inframarginal effects are disconnected if the cap is sufficiently low.

A defense of linear pricing relies on its distributive role: guaranteeing the buyer a share of the surplus is an appealing feature of linear pricing. Linear pricing may emerge as a general rule imposed in the buyer's interest simply because in the absence of information on future costs or preferences, it is extremely robust as a protection against full rent extraction.[7]

In this section, the seller is assumed to use linear tariffs. Before solving the game for any T, the extreme cases $T = +\infty$ and $T = 0$, which provide a useful insight into how the game functions, will be analyzed.

4.1 Unlimited commitment: Model *pA*

The timing is simplified, hence the name pA:

1. The seller sets a price p.
2. The buyer invests A at a unit cost k.

After this, at each instant, the buyer faces the predefined price p and purchases q. Obviously, there is no point in a buyer overinvesting with respect to his future consumption, therefore after observing p, the buyer will invest so that at equilibrium A equals q. He buys one unit of equipment for each unit of commodity, so the marginal cost of one commodity unit is $p + rk$. Therefore, his demand function is

$$q(p) = Q(p + rk). \tag{15}$$

The seller equalizes marginal revenue and cost and thus picks p that solves

$$\partial \left(pQ(p + rk) \right) / \partial q = c. \tag{16}$$

The corresponding investment is $A = Q(p + rk)$. The results are summarized in the following proposition:

Proposition 2. *In game pA the equilibrium price and investment level are*

$$p = \frac{\varepsilon c + rk}{\varepsilon - 1}; \tag{17}$$

$$A = Q\left(\frac{\varepsilon(c + rk)}{\varepsilon - 1}\right). \tag{18}$$

Remember the optimal investment level is $A^* = Q(c + rk)$. In model pA, since the demand function is $Q(p + rk)$, the welfare-maximizing price is $p = c$. The equilibrium price is always higher than the social optimum, so that the equilibrium investment level is always suboptimal.

[7] Linear prices are robust to asymmetric information about the demand parameter d in the sense that linear tariffs do not depend on d at equilibrium (as the following analysis will show).

4.2 No commitment: Model Ap

The timing is also simplified, hence the name Ap:

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.

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. He has no means to obtain more than the unconstrained monopoly quantity. Even this quantity could actually require too much investment: the buyer weighs up the initial investment cost against the subsequent costs of purchasing the good. Depending on the parameters, he might either passively adjust the investment to the monopoly quantity, or deliberately restrict investment, even though this will induce a higher commodity price. These are the two regimes that are described below.

The model is solved by backward induction.

Choice of q. In the last step, given A and p,

$$\begin{array}{ll} q = Q(p) & \text{if } p \geq P(A), \\ q = A & \text{if } p \leq P(A). \end{array} \tag{19}$$

Choice of p. Anticipating this, the seller chooses p. The optimal price is such that

$$\partial(p\,Q(p))/\partial q = c, \tag{20}$$

if and only if this price effectively leads the buyer to purchase $q = Q(p)$, *i.e.* if it is higher than $P(A)$. In this case, where the seller is not constrained by the capacity, the optimal price is such that the Lerner index equals the inverse of the elasticity of demand: $\frac{p-c}{p} = \frac{1}{\varepsilon}$. This yields the Ramsey monopoly price

$$\overline{p} = \frac{\varepsilon c}{\varepsilon - 1}. \tag{21}$$

This "passive-buyer" regime prevails if and only if $A \geq Q(\overline{p})$. Else, when the buyer has restricted his investment below the monopoly quantity, the seller has to set a higher price in order to equate the corresponding demand to the equipment size: $Q(p) = A$, thus $p = P(A)$. This characterizes the "active-buyer" regime.

To summarize,

$$p = \max\{\overline{p}, P(A)\}. \tag{22}$$

.

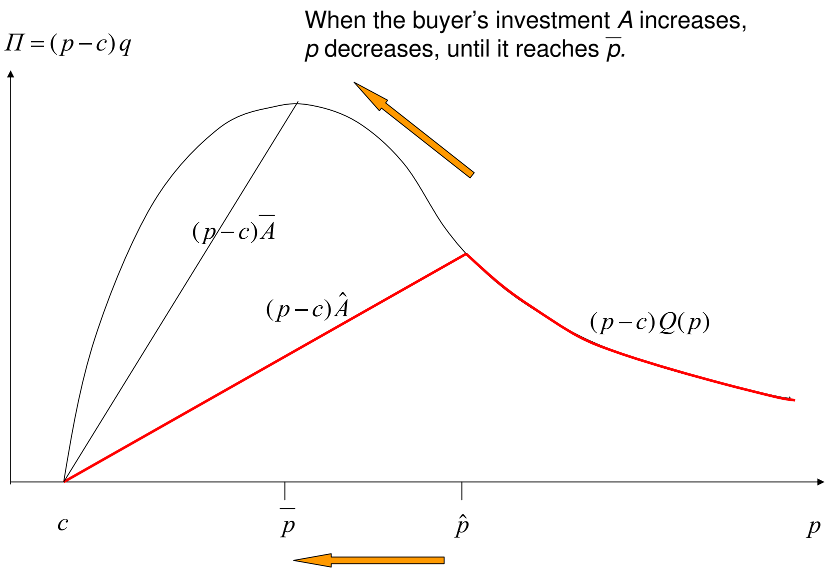

Figure 1: Impact of the investment choice on the price in model Ap.

Choice of A. Solving backwards, we come to the investment choice of the buyer. See Figure 1. We saw that 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 $\overline{p}$: it is never worth investing more than $\overline{A} = Q(\overline{p})$. Actually, when the investment cost is large, the buyer prefers to invest less.

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

$$\begin{aligned} &\max_A S(A) = \frac{1}{r}\left(u(A) - (P(A) + rk)A\right) \\ &\text{s.t. } A \leq Q(\overline{p}) \end{aligned} \tag{23}$$

The solution A is interior when the "active-buyer" regime prevails: the seller adjusts his price so that $q = A$. The optimal investment level equates marginal benefits $(q/r)dp$ (one more unit of equipment causes a permanent reduction dp in the commodity price) with marginal costs k: $(q/r)\,dp = k\,dA$. As $q = A$, this can be rewritten $\frac{1}{r}\frac{dp}{dq} = -\frac{k}{q}$, which yields $p = \varepsilon rk$, and since by assumption $p = P(A)$, the optimal investment level is $A = Q(\varepsilon rk)$. In response, the seller will set[8]

$$p = \varepsilon rk. \tag{24}$$

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

[8]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 $S(A) = \frac{1}{r}(u(A) - P(A)A - rkA)$. 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 $p(A) = \varepsilon rk$.

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

The following proposition summarizes the results.

Proposition 3. *In game Ap, the equilibrium price and investment level are*

$$p = \max\{\tfrac{\varepsilon c}{\varepsilon-1}, \varepsilon rk\}, \tag{25}$$

$$A = \min\{Q\left(\tfrac{\varepsilon c}{\varepsilon-1}\right), Q(\varepsilon rk)\}. \tag{26}$$

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.

4.3 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. The solution of this game is trivial: the buyer will choose $q = \min\{Q(p), A\}$, and anticipating this, the seller chooses simultaneously A and p. Obviously he will not set $A > Q(p)$ because the capacity would be oversized, nor $A < Q(p)$ because he would forgo profit opportunities (the margin $p - c - rk$ is positive), thus necessarily $q = A = Q(p)$ where p maximizes $\Pi(p) = \frac{1}{r}(p - (c + rk))\,Q(p)$. This is the program of a monopolist whose marginal cost is $c + rk$: the equilibrium is characterized by

$$p = \tfrac{\varepsilon(c+rk)}{\varepsilon-1}, \tag{27}$$

$$A = Q(\tfrac{\varepsilon(c+rk)}{\varepsilon-1}). \tag{28}$$

This game is essentially equivalent to the game pA 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.

As a matter of fact, the identity of the investor is not crucial. The real issue is whether the game played is essentially a "two-stage game" (like this game or game pA) where the seller plays first and the buyer has the final word, or a "three-stage game" (like game Ap) where in an additional preliminary step the buyer takes a first move by choosing A.

5 Linear tariffs with expiry date

5.1 A general characterization

Now the linear case with positive and finite values of T will be treated. Two stages can be distinguished: 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 contract duration T (*i.e.* the length of the first stage) acts as a weight given to the first stage as opposed to the second stage, so that the basic games Ap and pA seen in the previous section will correspond to extreme values (close to zero or to infinity) of the contract duration in this "pAp" game.

As in game Ap, 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:

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.

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 first two lemmas will eliminate large classes of strongly dominated strategies.

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:

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

Proof. If $A > Q(\frac{\varepsilon c}{\varepsilon-1})$, then A is not constraining in the second stage, and the seller can set the unconstrained monopoly price $\frac{\varepsilon c}{\varepsilon-1}$. Therefore, $A \leq Q(\frac{\varepsilon c}{\varepsilon-1})$ and $p_T = p(A) \geq \frac{\varepsilon c}{\varepsilon-1}$. Since by definition of active buyer, $p_T \neq \frac{\varepsilon c}{\varepsilon-1}$, the lemma obtains. □

Lemma 2. *When the buyer is passive, $p_0 = p_T = P(A) = \frac{\varepsilon c}{\varepsilon-1}$, and the seller is also passive.*

Proof. If the buyer is passive, $p_T = \frac{\varepsilon c}{\varepsilon-1}$. Since $p_T = \max\{P(A), \frac{\varepsilon c}{\varepsilon-1}\}$, this implies $A \geq Q(\frac{\varepsilon c}{\varepsilon-1})$. Suppose $A > Q(\frac{\varepsilon c}{\varepsilon-1})$. The choice of A only affects the first-stage game, and so does p_0. This game is solved like model pA, the difference being that trade takes place from $t = 0$ to T only: the buyer will equate q_0 with A, so that his program can be rewritten

$$\max_A \frac{1-e^{-rT}}{r}\left[u(A) - \left(p_0 + \frac{rk}{1-e^{-rT}}\right)A\right], \tag{29}$$

and at equilibrium the buyer's best choice is

$$q_0 = A = Q\left(p_0 + \frac{rk}{1-e^{-rT}}\right). \tag{30}$$

Anticipating this the seller chooses

$$p_0 = \frac{\varepsilon c + \frac{rk}{1-e^{-rT}}}{\varepsilon-1} > \frac{\varepsilon c}{\varepsilon-1}. \tag{31}$$

But this means that $A = Q(p_0) < Q(\frac{\varepsilon c}{\varepsilon-1})$, which contradicts the assumption. Therefore, $A = Q(\frac{\varepsilon c}{\varepsilon-1})$. Anticipating this, the seller will set p_0 at the monopoly price. □

In words, the investment capacity can never be oversized in the second stage. 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.

Now we turn to the first stage. The next lemma proves that the investment can never be oversized in the first stage either:

Lemma 3. *At equilibrium, $A \leq Q(p_0)$.*

Proof. Assume that $A > Q(p_0)$. In model pA there was no reason to overinvest with respect to demand. Here the capacity could be adjusted to the anticipated second-stage demand and oversized relative to first-stage demand: $A > q_0 = Q(p_0)$, or equivalently $p_0 > P(A)$. Since (from Lemmas 1 and 2) $p_T = P(A) \geq \frac{\varepsilon c}{\varepsilon - 1}$, this requires $p_0 > \frac{ec}{e-1}$. But if the capacity were strictly oversized in the first stage, this would mean that the capacity was chosen according to the second stage, therefore a marginal change in p_0 would not affect the investment size, nor would it affect the second-stage game. Therefore, if $p_0 > \frac{\varepsilon c}{\varepsilon - 1}$, the seller would have an incentive to deviate to a lower price, closer to his profit-maximizing price $\frac{\varepsilon c}{\varepsilon - 1}$: this would increase his first-stage profits without altering the second-stage game. Consequently, $A > Q(p_0)$ cannot be an equilibrium. $\square$

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

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

Table 1: The 4 types of equilibrium in game pAp.

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 on the case where the seller and the buyer are active.

The other cases are treated in the Appendix.

5.2 When an active seller faces an active buyer

Whenever the buyer is active, he solves:

$$\max_A \frac{1-e^{-rT}}{r}(u(A) - p_0 A) + \frac{e^{-rT}}{r}(u(A) - P(A)A) - kA, \tag{32}$$

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 3 enables us to replace consumption in the first period by A.

We find

$$A = Q\left(\frac{p_0 + \frac{r}{1-e^{-rT}}k}{1 + \frac{1}{\varepsilon}\frac{e^{-rT}}{1-e^{-rT}}}\right). \tag{33}$$

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

$$\max_{p_0}\left[\frac{1-e^{-rT}}{r}(p_0 - c) + \frac{e^{-rT}}{r}\left(\frac{p_0+\frac{r}{1-e^{-rT}}k}{1+\frac{1}{\varepsilon}\frac{e^{-rT}}{1-e^{-rT}}} - c\right)\right] Q\left(\frac{p_0 + \frac{r}{1-e^{-rT}}k}{1 + \frac{1}{\varepsilon}\frac{e^{-rT}}{1-e^{-rT}}}\right). \tag{34}$$

This yields the optimal contract price

$$p_0 = \frac{1}{1-e^{-rT}}\left[\left(1 - \frac{\varepsilon e^{-rT}}{\varepsilon+e^{-rT}}\right)\frac{\varepsilon c}{\varepsilon-1} + \left(1 - \frac{\varepsilon^2 e^{-rT}}{\varepsilon+e^{-rT}}\right)\frac{rk}{\varepsilon-1}\right]. \tag{35}$$

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 (33) and (35) 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

$$\frac{c}{rk} < (\varepsilon - 1)e^{rT} - \frac{1}{\varepsilon}. \tag{36}$$

Therefore, a necessary and sufficient condition for both players to be active *for all* T is

$$\frac{c}{rk} < \varepsilon - 1 - \frac{1}{\varepsilon}. \tag{37}$$

Note that for the right hand-side to be positive, the elasticity of demand must exceed $\frac{1+\sqrt{5}}{2} \simeq 1.6$, the Golden Ratio.

Proposition 4 (Both players active for any contract duration)**.** *When* $\frac{c}{rk} < \varepsilon - 1 - \frac{1}{\varepsilon}$,

$$\text{If } T > 0, \begin{cases} p_0 &= \frac{1}{1-e^{-rT}}\left[\left(1 - \frac{\varepsilon e^{-rT}}{\varepsilon+e^{-rT}}\right)\frac{\varepsilon c}{\varepsilon-1} + \left(1 - \frac{\varepsilon^2 e^{-rT}}{\varepsilon+e^{-rT}}\right)\frac{rk}{\varepsilon-1}\right], \\ p_T &= \frac{\varepsilon^2(c+rk)}{(\varepsilon+e^{-rT})(\varepsilon-1)}). \end{cases} \tag{38}$$

If $T = 0$, p_0 *is irrelevant and* $p_T = \varepsilon rk$.

The equilibrium is discontinuous at 0, meaning that no commitment ($T = 0$) and “some” commitment ($T > 0$) are qualitatively different.

5.3 Impact of the contract duration

Since $A = Q(p_T)$ is a decreasing function of p_T, which (from Proposition 4) is an increasing function of T, the following paradoxical result obtains.

Proposition 5 (Investment)**.** *The equilibrium investment level decreases with respect to the contract duration for all* $T > 0$.

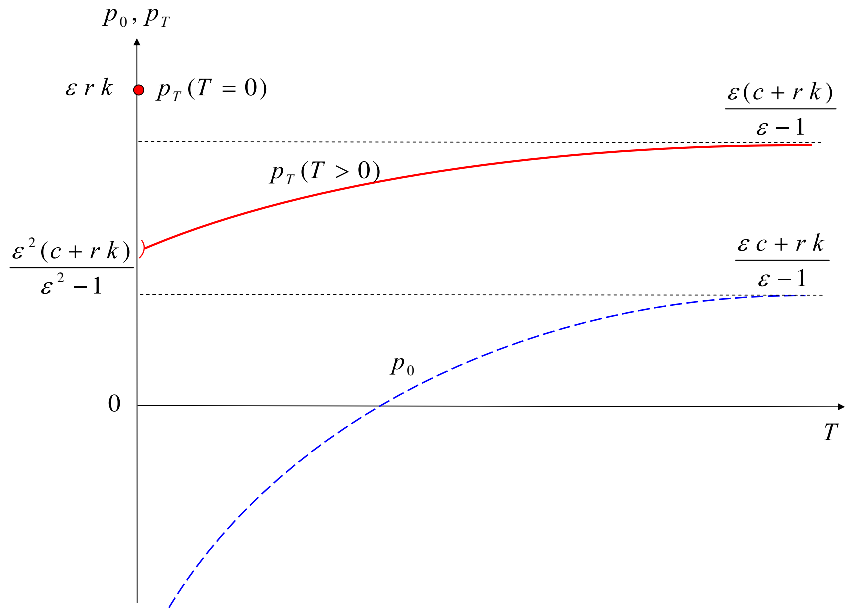

Figure 2: Equilibrium prices as a function of the contract duration (when $\frac{c}{rk} < \varepsilon - 1 - \frac{1}{\varepsilon}$).

In other terms, to encourage investment, the smallest contract is the best, no contract at all is the worst case.

Since the seller cannot commit to refrain from hold-up 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. When T is close to zero but still strictly positive, p_0 becomes infinitely negative (see Figure 2). But when $T = 0$ the game is radically altered, because the seller has no possibility to induce a higher investment than the level corresponding to the monopoly price: the investment suddenly falls. This is reflected in a discontinuity in the second-stage hold-up price $p_T = P(A)$ (upward jump when the duration goes from $T > 0$ to $T = 0$).

Since social welfare increases with respect to the investment level, Proposition 5 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

$$\Pi = \frac{1}{r}\frac{c+rk}{\varepsilon-1} Q\left(\frac{\varepsilon^2(c+rk)}{(e^{-rT}+\varepsilon)(\varepsilon-1)}\right); \tag{39}$$

$$S = \frac{1}{r}\frac{1-e^{-rT}+\frac{1}{\varepsilon}e^{-rT}}{\varepsilon-1}\frac{\varepsilon^2(c+rk)}{(e^{-rT}+\varepsilon)(\varepsilon-1)} Q\left(\frac{\varepsilon^2(c+rk)}{(e^{-rT}+\varepsilon)(\varepsilon-1)}\right). \tag{40}$$

The following proposition summarizes.

Proposition 6 (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.*

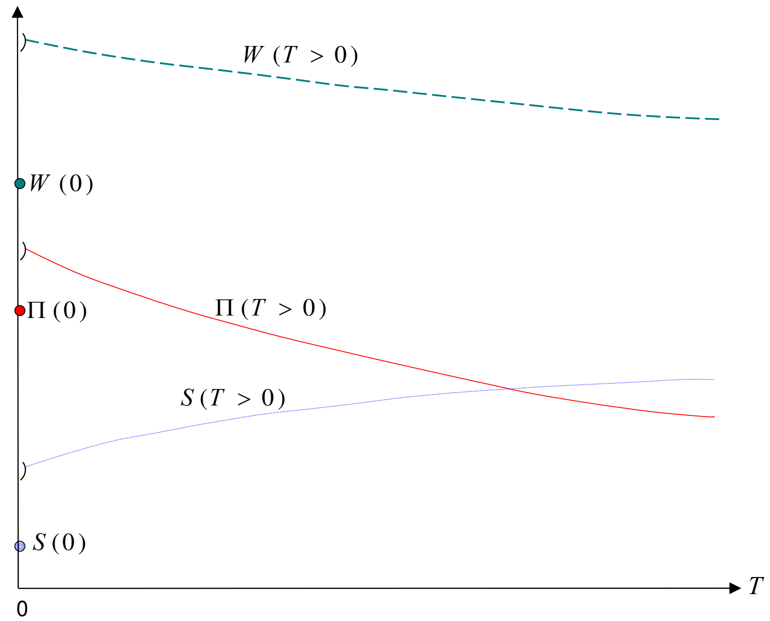

Figure 3: Profit of the seller, surplus of the buyer and social welfare as a function of the contract duration (when $\frac{c}{rk} < \varepsilon - 1 - \frac{1}{\varepsilon}$).

With linear prices, more commitment through longer contracts is detrimental to the seller *and* has the effect of depressing investment.

In a nutshell, the reason is that the first price p_0 serves to push investment whereas p_T serves to extract rents. We have a clear analogy with the static monopoly case recalled at the beginning of Section 4: 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.

The attempt to approach the optimum is less successful with long commitments. As T increases, the hold-up period gets farther away, and pushing investment through price rebates in the first period becomes costly to the seller. Of course the rebate can be smaller if it can be sustained for a longer time; still, the profitable period is rejected to a farther future, which imposes a clear opportunity cost. As a consequence, a large T induces a smaller A, which means lower social surplus and lower profits.

What about the buyer? As T increases, the buyer's surplus increases. In other terms, as social surplus decreases by one unit, the buyer's surplus does increase by more than one.

At the beginning of Section 4, we recalled the protective role for the buyer of a cap on the fixed fee. Longer durations are similar to lower (*i.e.* more constraining) caps: the longer the commitment, the more limited the hold-up period. In compensation, as the second price loses impact, the seller has to mix in the first price two objectives (incentives and rent extraction), which is the typical limit of plain static linear prices.

5.4 Investment by the seller

If the seller invests, there is no need to create incentives and extract rents in two separate episodes: he can and will replicate the same price in the first and second stages. This means that contract duration does not matter. The contract in effect necessarily gives the solution of game pA seen in Section 4, which is equivalent to the case $T = +\infty$ in the pAp game.

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. This makes the seller's contribution to investment (through subsidies) relatively small and he uses the second price p_T to extract a bigger rent.

6 Conclusion

Long-term contracts on specific assets have been repeatedly under the scrutiny of the competition authorities in Europe. The divide is between the pro-competition line of argument (foreclosure establishes durably abusive market power, *e.g.* Aghion and Bolton, 1987) and the pro-investment line of argument (contracts offer protection against expropriation). The 2007 Energy Sector Inquiry states that "long-term supply agreements seem to foreclose the availability of crucial inputs for actual or potential competition" (p. 66). On the other hand, the 2004 Directive on Security of Natural Gas Supply emphasizes their positive impact on investment, underlining that "long-term contracts have played a very important role in securing gas supplies for Europe and will continue to do so" (recital 11). Though the latter view is largely supported by the energy industry, 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.

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 the pricing schemes that we have analyzed share the following feature: some commitment ($T > 0$) is always better than none ($T = 0$). The reason is intuitive: the seller can offer a significant rebate that gives the buyer perfect or at least fairly good incentive 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 (provided $T > 0$). With linear tariffs, the investment level is a strictly decreasing function of the contract length. A shorter subsidy period concentrates the effort, but it doesn't diminish the incentives that can be conveyed. Moreover, a shorter commitment limits the postponement of the profitable hold-up period, mechanically limiting the opportunity cost. Accordingly, when the contract duration is very short, the seller will choose to subsidize investment heavily to maximize the surplus that will subsequently be extracted. In this sense, length and width of the contract are complements.

With linear tariffs, social welfare, the seller's profit and the buyer's surplus do not vary in parallel when contract length changes. The analogy is the following: in a static monopoly model, it is possible to increase the buyer's surplus by restricting the use of two-part tariffs. One option is to put a cap on the fixed part, or even to eliminate it. Albeit inefficient, pure linear pricing may afford the buyer more surplus than other solutions. In our model, the first period is the one during which investment incentives are set up by the seller. The seller tries to make the buyer invest and consume as if prices were exactly equal to marginal costs. The second period in contrast acts like the fixed part, serving to extract surplus: requiring that the commitment period be long means that there is a cap on the rent extracted by the seller, which benefits the buyer.

The analogy is imperfect, if only because two linear prices cannot strictly replicate nonlinear tariffs. Nevertheless, it simply shows that a rule forcing linear prices and longer commitment on prices could be set up as a "universal" means to protect buyers against sellers' market power. This would be at the expense of investment. Depending on the political game played between the various interest groups in a society, a rule guaranteeing linear prices and forcing a certain commitment will be imposed or opposed.

References

Aghion, Philippe and Patrick Bolton (1987): "Contracts as a Barrier to Entry", American Economic Review, 77(3), 388-401.

Balestra, Pietro and Marc Nerlove (1966): "Pooling, Cross Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas", Econometrica, 34(3), 585-612.

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 (2001): "Competition, Monopoly, and Aftermarkets", NBER Working Paper 8086.

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.

Council of the European Union (2004): Council Directive 2004/67/EC of 26 April 2004 concerning measures to safeguard security of natural gas supply.

Creti, Anna and Bertrand Villeneuve (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 (1985): "Efficient Adaptation in Long-term Contracts: Take-or-Pay Provisions for Natural Gas", American Economic Review, 75, 1083- 1093.

Crocker, Keith J. and Scott E. Masten (1988): "Mitigating Contractual Hazards: Unilateral Options and Contract Length", RAND Journal of Economics, 19(3), 327-343.

Dubin, Jeffrey.A. and Daniel L. McFadden (1984): "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption", Econometrica, 52, 118-131.

European Commission (2007): DG Competition report on energy sector inquiry, SEC(2006) 1724, 10 January 2007.

Goldberg, Victor P. (1976): "Regulation and Administered Contracts", Bell Journal of Economics, 7(2), 426-448.

Hanemann, W. Michael (1984): "Discrete/Continuous Models of Consumer Demand", Econometrica, 52(3), 541-561.

Hubbard, Glenn and Robert Weiner (1986): "Regulation and Long-term Contracting in US Natural Gas Markets", Journal of Industrial Economics, 35, 71-79.

Joskow, Paul (1988): "Price Adjustment in Long-term Contracts: The Case of Coal", Journal of Law and Economics, 31, 47-83.

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. Woroch (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.

Samuelson, Paul (1947): *Foundations of Economic Analysis*. Cambridge, MA: Cambridge University Press.

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

The case when $\frac{c}{rk} < \varepsilon - 1 - \frac{1}{\varepsilon}$ was detailed in subsection 5.1. Here the general case is treated.

A.1 Linear tariffs with expiry date: general case

From Lemmas 1 and 2, we can deduce that a necessary and sufficient condition for the buyer to be active is $p_T > \frac{\varepsilon c}{\varepsilon - 1}$. In addition, equation (33) tells us that when the buyer is active, he anticipates that the seller will adjust the price to the investment ($p_T = P(A)$). From equation (33),

$$p_T = \frac{p_0 + \frac{r}{1-e^{-rT}}k}{1 + \frac{1}{\varepsilon}\frac{e^{-rT}}{1-e^{-rT}}}. \tag{41}$$

Conversely, 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

$$p_T = \min\left\{\frac{p_0 + \frac{r}{1-e^{-rT}}k}{1 + \frac{1}{\varepsilon}\frac{e^{-rT}}{1-e^{-rT}}}, \frac{\varepsilon c}{\varepsilon - 1}\right\}. \tag{42}$$

This enables us to write the following necessary and sufficient condition for having an active buyer:

$$p_0 - \frac{\varepsilon c}{\varepsilon - 1} > \frac{c - (\varepsilon - 1)e^{rT} rk}{(\varepsilon - 1)(e^{rT} - 1)}, \tag{43}$$

The buyer is passive if and only if the opposite inequality holds. 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

$$\frac{c}{rk} \geq (\varepsilon - 1)e^{rT}, \tag{44}$$

the buyer is passive—and so is the seller.

Now suppose $\frac{c}{rk} < (\varepsilon - 1)e^{rT}$: the buyer is active. Two subcases have to be analyzed: passive seller and active seller. Let us start with the case where the seller is active. A necessary and sufficient condition for having an active seller is $p_0 < P(A)$. In this case, as computed previously (see equation 35),

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

The condition $p_0 < P(A)$, where A is given by equation (33), can now be rewritten $\frac{c}{rk} < (\varepsilon - 1)e^{rT} - \frac{1}{\varepsilon}$, which yields the following necessary and sufficient condition for both parties to be active:

$$\frac{c}{rk} < (\varepsilon - 1)e^{rT} - \frac{1}{\varepsilon}. \tag{45}$$

Accordingly, the equilibrium involves an active buyer and a passive seller if and only if $(\varepsilon - 1)e^{rT} - \frac{1}{\varepsilon} \leq \frac{c}{rk} < (\varepsilon - 1)e^{rT}$. In this case, $p_0 = p_T = P(A)$, and combining equations (41) and (35) yields the equilibrium prices and investment level.

The results are summarized in the following proposition:

Proposition 7 (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*

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

A.2 *pAp*: Comparative statics w.r.t. investment cost

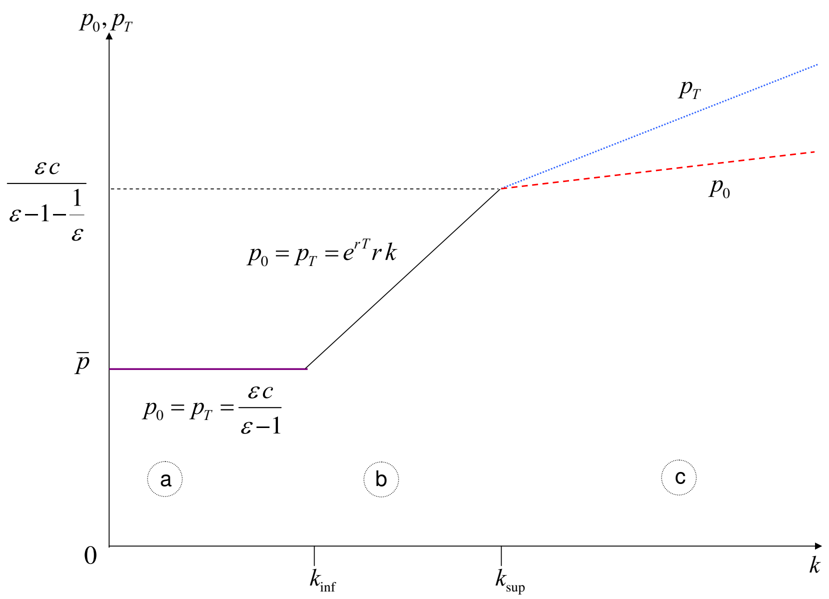

Figure 4: Equilibrium prices as a function of the investment cost.

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: it may be worth the seller's while to set a low p_0, thus sacrificing first-stage profits, to induce a higher investment and more second-stage sales at the hold-up price $p_T = P(A)$.

Let

$$\begin{cases} k_{\inf} &= \frac{e^{-rT}}{r}\frac{c}{\varepsilon-1}, \\ k_{\sup} &= \frac{e^{-rT}}{r}\frac{c}{\varepsilon-1-\frac{1}{\varepsilon}e^{-rT}}. \end{cases}$$

For all T, $0 < k_{\inf} < k_{\sup}$. The three cases mentioned in Proposition 7 correspond respectively to $k \le k_{\inf}$, $k_{\inf} < k \le k_{\sup}$ and $k > k_{\sup}$, and and to zones ⓐ, ⓑ, and ⓒ in Figure 4.

ⓐ $k \le k_{\inf}$: both parties are passive. In the second stage, the seller will be able to set his unconstrained monopoly price $p_T = \frac{\varepsilon c}{\varepsilon - 1}$ only if the invested capacity is sufficiently large. Therefore, as long as setting his monopoly price $p_0 = \frac{\varepsilon c}{\varepsilon - 1}$ in the first stage induces a sufficiently large investment $A \ge Q(\frac{\varepsilon c}{\varepsilon - 1})$, the seller has no incentive to deviate, and he will set in both stages $p_0 = p_T = \frac{\varepsilon c}{\varepsilon - 1}$. This will be the case when the investment cost k is not too large: the buyer responds passively to the contract price by setting $A = Q(\frac{\varepsilon c}{\varepsilon - 1})$.

When k becomes larger, the buyer compensates the higher unit cost of equipment by reducing his investment to $A < Q(\frac{\varepsilon c}{\varepsilon - 1})$. The seller then necessarily sets in the second stage the corresponding hold-up price $p_T = P(A) > \frac{\varepsilon c}{\varepsilon - 1}$. In the first stage, two possible choices are available to the seller, knowing that $q_0 = \min\{A, Q(p_0)\}$. Either he passively sets p_0 such that at equilibrium $p_0 = p_T = P(A)$ and in particular $p_0 = P(q_0)$ (case ⓑ below), or he strategically sets the contract price below the marginal willingness to pay of the buyer ($p_0 < P(q_0)$): he forgoes first-stage profits in order to induce the buyer to invest in a larger equipment capacity, thus preserving second-stage profits (case ⓒ below).

ⓑ $k_{\inf} < k \le k_{\sup}$: only the buyer is active. For intermediate values of k, the investment level is still close to the seller's preferred level $Q(\frac{\varepsilon c}{\varepsilon - 1})$, therefore setting $p_0 = p_T > \frac{\varepsilon c}{\varepsilon - 1}$ yields only a second-order loss in both stages compared to the initial situation with a low k, whereas setting $p_0 < P(A) = p_T$ would cause a first-order loss $(P(A) - p_0)\, A$.

ⓒ $k > k_{\sup}$: 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.3 *pAp*: Comparative statics w.r.t. contract duration

The case where $\frac{c}{rk} < \varepsilon - 1 - \frac{1}{\varepsilon}$ was detailed in Subsection 5.1. Now assume the opposite holds. Let

$$\left\{ \begin{array}{rcl} T_{\inf} & = & \frac{1}{r} \ln\left(\frac{c}{(\varepsilon-1)rk}\right), \\ T_{\sup} & = & \frac{1}{r} \ln\left(\frac{\varepsilon c + rk}{\varepsilon(\varepsilon-1)rk}\right). \end{array} \right. \tag{46}$$

The three cases mentioned in Proposition 7 correspond respectively to $T \le T_{\inf}$, $T_{\inf} < T \le T_{\sup}$ and $T > T_{\sup}$, and to zones ⓐ, ⓑ, and ⓒ in Figure 5.

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

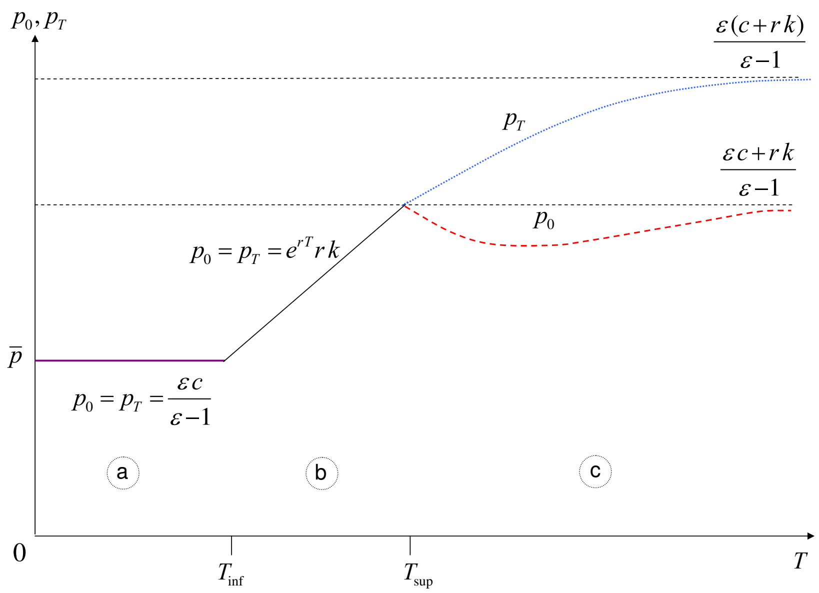

Figure 5: Equilibrium prices as a function of the contract duration (when $\frac{c}{rk} > \varepsilon - 1$).

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

ⓑ $T_{\inf} < T \leq T_{\sup}$**: 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 $A = Q\left(re^{rT}\varepsilon k\right)$.

ⓒ $T > T_{\sup}$**: 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. When T tends to infinity, the contract price tends to p_{pA} .

A.4 *pAp*: Profit and surplus analysis

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

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

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

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

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.

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

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

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

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 $T \to +\infty$. Comparing the two values yields the following result: the buyer's surplus is highest for $T = 0$ if and only if

$$\frac{c}{rk} \leq (\varepsilon-1)\varepsilon^{\frac{1}{\varepsilon-1}} - 1. \tag{47}$$

Finally, when $\frac{c}{rk} \geq \varepsilon - 1$, all three cases exist. Clearly the buyer's surplus is maximal either for $T = 0$, or when T tends to infinity. Let $y = \frac{c}{rk}$. After rearrangement, we find that the buyer prefers the shortest possible contract whenever

$$\left(1 - \frac{(\varepsilon-1)^2}{\varepsilon y}\right)\left(1 + \frac{1}{y}\right)^{\varepsilon-1} \geq 1. \tag{48}$$

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 $T \to +\infty$ whenever $y \leq y^*$, and it is maximal for $T = 0$ 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.