16 Long-term contracts in the law and economics literature

Mireia Artigot i Golobardes and Fernando Gómez Pomar

1. Introduction

Contracts increase the likelihood of cooperation in economic and social interaction by making binding commitments credible and less costly. This is also true of cooperation over extended periods of time, even in the presence of other modes of inducing and stabilizing cooperation in protracted interactions. The economic literature dealing with contract issues and, more specifically, on long-term relationships has differentiated between long-term contracts and contracts that extend over a long period of time, often called relational contracts. Relationships that last a long time may be governed by a long contract or by many short contracts, because they do not necessarily require a certain type of contract, a given contract length, or even a formal contract at all. While all long-term contracts are contracts that tend to last a long time and share certain common characteristics, not all contractual relationships that involve significant duration are drafted as long-term contracts.

From a law and economics perspective, three major dimensions of long-term contracts are to be highlighted: the existence of specific investments, their inherent – added – incompleteness, and the complexities of the issues arising from breach and termination.

Specific investments are investments, the value of which depends on whether they are used by parties within the contractual relationship or outside it. A pure specific investment has value for the contract between the parties but is worth nothing outside it, or whenever the contract ends. Hence, in light of the relation-specific value of the investment, the party undertaking such an investment is particularly vulnerable, being subject to the strategic behavior of the other party in the contract and given that outside the contract the value of the investment vanishes. Such specific

---

1 Williamson (1985).
2 Hviid (1999).
4 Becker (1962).
investments strongly influence the parties’ strategies and their incentives to cooperate. This situation is what in economics is often referred to as the holdup problem under which parties would reach a Pareto outcome if they were to cooperate but because of the risk involved in cooperating, they do not reach their first-best solution and therefore do not maximize the returns of the contract between the parties.

A second major characteristic of long-term contracts is the higher degree of uncertainty that is inherent in such a contract, and the impossibility of drafting a complete contract that could foresee and resolve all the potential contingencies that might take place during the life of the contract. Any contract faces the tough challenge of including all necessary clauses to give solutions to potential issues affecting the contractual relationship between the parties. This is an especially important issue in long-term contracts. For that reason, long-term contracts often include renegotiation clauses so that they can be adjusted to any new circumstances faced by the parties and by the contract during its life span. Incompleteness is by no means a unique characteristic of long-term contracts, but it is an especially critical question considering the goals and threats to long-term relationships between parties entering into them.

Finally, long-term contracts typically present increased opportunity for a failure of legal remedies against breach of contract. Due to the long duration of the interaction, the multiplicity of types and events of relevant contractual conduct, the proliferation of instances of potential shading and shirking by the parties, the chances that these non-complying contractual behaviors can be shown, in a sufficiently convincing manner, to a court or other external adjudicator, are significantly lower than in a spot contract with a smaller range of contractually relevant behaviors. Moreover, typical remedies such as specific performance and damages are harder to assess and enforce. These complexities grow even larger inside supply and distribution chains, where the cost of detecting and collecting evidence of these instances of breach, and adequately deterring them

---


6 Salanié (1997) defines a complete contract as one that considers ‘all variables that may have an impact on the conditions of the contractual relationship during its whole duration have been taken into account when negotiating and signing the contract’. It should be noted that under this definition, a contract would be complete, despite not regulating all potential foreseeable circumstances, if there were no change of circumstances that affected the contract.

7 Some authors consider that whenever parties have no need to revise or renegotiate the contract, such contracts could be considered comprehensive contracts. Hart (1995).
increases with the size, territorial, and product scope of the network. The likelihood that specific contract clauses over remedies in the long-term contract can address these problems seems low.

This chapter will be structured as follows: section 2 will present the major differences between long-term and short-term contracts; section 3 will discuss the most important parameters affecting the optimal contract length; section 4 will present other issues inherent in long-term contracts; section 5 will describe the major issues on contract drafting and design; section 6 will discuss the major differences between contract renegotiation and breach; section 7 will present the consequences of termination and breach; section 8 will briefly present the empirical literature on long-term contracts. The chapter will end with some brief conclusions.

2. **Long-term Versus Short-term Contracts**

Despite the particular features and complexities presented by long-term contracts, their legal nature and regulation do not differ significantly from the regulation of short-term contracts or contracts in general.

Contracts are legal instruments that allow parties to establish binding commitments that serve as mechanisms for creating value, as well as for deciding how such value will be divided. Legal systems typically establish – with some variation across legal traditions – a set of requirements in order to have an enforceable contract beyond what could be a simple agreement between two parties. Among these elements there is one that requires that the agreement meets a minimum level of definiteness so that parties, for the contract to be enforceable, need to be in a reasonable position and have enough information to understand and agree on the contract terms, and in the case of breach or disagreement between them, the courts have sufficiently precise information on what the contract required, so as to be able ex post to solve the grievances of the contract parties and award damages, if appropriate.

Protracted relationships may be governed by long-term or by a series of short-term contracts. However, their content is likely to look quite different. While both kinds of contract include basic contract requirements such as a reasonably defined contract object, and normally include the exchange price, long-term contracts may also include governance terms over the long run, re-adjustment provisions, remedies for the long run, and parameters that may affect the parties’ bargaining position within the contractual relationship.

---

The choice between a short contract – or a series of short contracts – or a long-term contract depends on which type of instrument best fits the preferences and nature of the parties' relationship. There will be several dimensions of the contract that will differ depending on whether the contract is short term or long term.

Despite the common legal requirements on many matters for both kinds of contract, issues such as definiteness, uncertainty and specific investments will differ because of the characteristics of the legal instrument. Further, long-term contracts will often generally include many aspects that are not considered necessary in the context of short-term contracts.

For instance, the need to define and specify contract terms will differ depending on the length of the contract, or on whether it will be necessary to make a significant amount of specific investments. So, in a long-term contract, it will probably be necessary to be less precise, and to leave more scope for discretion by the contracting parties in order to be able to adjust the contract terms to the circumstances in relation to the parties' needs. For that reason, long-term contracts tend to be more open-ended. But at the same time, given that parties are aware of the risk arising from specific investments, they will also be more interested in defining the consequences they will face in the case of contract break-up.10

Some voices in the literature have discussed whether long-term contracts require different rules because of the specific problems presented by the specific investments contract parties must make, and the inherent incompleteness that derives from the lack of information available to contracting parties when the contracts are drafted.11

There is no general criterion under which certain relationships should be governed by long-term contracts or under a series of short-term contracts. In other words, there is no optimal contract length at an abstract level, and the optimal contract will depend on the nature of the parties' relationship and the nature and goals of the contract they have entered into.

However, there are certain issues on the performance of short-term and long-term contracts that have been widely discussed in the literature. For example, whether parties have complete or incomplete information available, short-term contracts perform differently from long-term contracts. At the same time, depending on whether the parties' commitment or the role of specific investments are important elements in the relationship, short-

---

term or long-term contracts may be preferred. Fudenberg et al.\textsuperscript{12} identified sufficient conditions – all public information is contractible, there is no information asymmetry when the contract is entered into and contracts are renegotiated – for short-term contracts to achieve the same outcomes as long-term contracts. These elements are developed below.

2.1. Complete Information
Under perfect conditions, parties envisaging a long-term relationship would enter into a long-term contract considering all potential contingencies that they would be able to foresee and contract over. Given that they would be able to anticipate all potential contingencies between them, renegotiation would not be necessary in principle.

It is worth noting that if parties had perfect information and renegotiation were not costly, long-term contracts would be preferred whenever one of the parties’ investments was a sunk cost and the contract would represent a way to smooth consumption.\textsuperscript{13} If renegotiation was not necessary, contract terms could be governed either through a series of short-term contracts or with a single long-term contract. Further, if parties did not place a specific value on commitment and had access to perfect information, short-term contracts would be equivalent to long-term contracts.

Sometimes the need for specific investments by contracting parties is an important element in entering into long-term contracts. One of the major problems of specific investments is that often they are not contractible. However, if there is perfect information, or such investments are observable so that the other contracting party can observe whether these investments are made, Dutta and Reichelstein,\textsuperscript{14} using a single variable in a multi-period context, noted that the hold-up problem present in long-term contracts is significantly mitigated. Hence, when there is perfect information, or specific investments are verifiable, investment incentives will depend on the weight placed on them being verifiable, and not on whether the contract entered into is a short-term or long-term contract.

Therefore, where parties successively negotiate contracts and there is perfect information accessible to both of them, short-term contracts can be as efficient as long-term contracting when they are entering into a new

\textsuperscript{12} Fudenberg et al. (1990). See also Salanié and Rey (1990) and Chiappori et al. (1994).
\textsuperscript{13} Some have suggested that in these cases long-term contracts serve as a substitute for an efficient credit market. Crawford (1988).
\textsuperscript{14} Dutta and Reichelstein (2003).
Hence, under perfect information conditions, even for protracted relationships, short-term contracts may be equivalent to long-term contracts.

2.2. Incomplete Information
When there is incomplete information on the part of either one or both parties, the outcomes under long-term and short-term contracts may vary significantly. Further, the possibilities and costs of renegotiating the contract may be relevant for the performance of short- and long-term contracts.

One of the major advantages presented by short-term contracts is the possibility they offer of inter-temporal smoothing so that parties can adjust to the information asymmetries between them. When some future outcomes are uncertain, when there are risks involved during the life of the contract, or unobservable actions by the parties that may affect the outcome of the contract, short-term contracts seem to present significant advantages. However, the advantages are not always so clear cut.

Short-term contracts do not necessarily solve inter-temporal trade-offs. Rey and Salanié claim that under asymmetric information, commitment – present in long-term contracts – becomes an essential element in the interaction, one that causes long-term contracts to strictly dominate short-term contracts. This advantage ensues from the fact that the incentive problems created by private information not revealed by contracting parties, are generally better overcome through ex ante commitment than through solving conflicts ex post, which will result in inefficiencies. However, because long-term contracts can rely and observe in the second period the returns of the investment made, this could generate incentives to invest, which could result in a better performance by long-term contracts in such a context, as some of the literature shows.

The problems presented by private information, though, may be overcome by eliminating the possibility of renegotiation. Even where there is asymmetric information, short-term contracting could implement optimal

---

15 Using a multi-period agency model, where at the beginning of each period one of the parties (the principal) can propose a contract to the other (the agent), on a take-it-or-leave-it basis. Rey and Salanié (1990).
16 Freixas et al. (1985).
17 Rey and Salanié (1990).
18 Rey and Salanié (1990).
19 For an analysis of long-term contracts with private information, see Baron and Besanko (1984).
20 Dutta and Reichelstein (2003).
renegotiation-proof contracts. Leading indicator variables may become useful if the principal is confined to renegotiation-proof contracts even though there are authors who still prefer long-term commitment whenever possible rather than the information provided by indicator variables.

2.3. The Importance of Commitment

One of the elements that may be crucial when determining the choice between short-term and long-term contracts is the importance of commitment in the context of the contract and related to this, whether renegotiation is feasible and if so, costly. Given the higher commitment inherent in long-term contracts, renegotiation is not generally so essential and may be avoided. This is important because renegotiating between parties may be costly and therefore, long-term contracts could imply important savings in transaction costs. But, as mentioned above, even when renegotiation is not costly, long-term contracts will still be preferred whenever one of the parties’ investments represents a way to smooth consumption.

Leaving renegotiation aside, under perfect information, commitment solves the parties’ inter-temporal trade-offs. Therefore, long-term contracts will be preferred to short-term contracts whenever they cannot solve such trade-offs outside the relationship. So, if one of the contracting parties – or both – cannot solve inter-temporal trade-offs outside the contract, some commitment will be necessary to achieve long-term efficiency.

---

21 Malcomson and Spinnewyn (1988).
22 Sliwka (2002).
23 Malcomson and Spinnewyn (1988), in the context of a principal-agent model under asymmetric information, show that, whenever there is no renegotiation, a long-term contract can perform better than a short-term contract when one of the parties – principal or agent – commits to a payoff which is lower than the potential outcome under a short-term contract. Therefore, long-term contracts have some commitment element that cause them to be preferred to a series of short-term contracts.
25 Crawford (1988). Some have suggested that in these cases long-term contracts serve as a substitute for an efficient credit market.
26 There is abundant literature discussing the importance of commitment as a mechanism to solve inter-temporal trade-offs whenever there is no access to credit markets. If both parties had access to perfect credit markets, if they cared only about the present value of the outcome and not about the time frame, there would exist a number of optimal contracts, all of which would have the same present value of the contract outcome in every state of nature, at least one of the contracts would meet all conditions for an optimal short-term contract and therefore would be equivalent to a long-term contract with commitment. See Rey and Salanié (1990).
and long-term contracts will then present an important advantage over short-term contracts. If inter-temporal trade-offs or commitment is not that important, short-term and long-term contracts may be equivalent.\textsuperscript{27} But when there are moral hazard problems involving inter-temporal risk-sharing, long-term contracts will dominate a sequence of short-term contracts.\textsuperscript{28} Accordingly, long-term contracts will in general dominate sequences of spot contracts.

The value of commitment has been widely discussed in the literature. Williamson\textsuperscript{29} suggested that short-term contracts have emerged in order to avoid the difficulties of specifying and enforcing the contingencies inherent in long-term contracts.\textsuperscript{30}

2.4. The Possibility of Renegotiation
As mentioned above, in light of the commitment element inherent of long-term contracts, the need for renegotiation is reduced significantly.\textsuperscript{31} When renegotiation is costly, the transactions costs of long-term contracts will be lower than for short-term contracts and therefore long-term contracts will be preferred.\textsuperscript{32} But if, instead, renegotiation is costless, and parties have perfect information, long-term contracts will still be preferred whenever the contract represents a way of smoothing consumption for at least one of the parties involved.\textsuperscript{33}

2.5. The Importance of Specific Investments
Dutta and Reichelstein\textsuperscript{34} identified the conditions under which optimal long-term contracts induce larger investments and less reliance on

\textsuperscript{27} Rey and Salanié suggest that short-term contracts could be interpreted as loan contracts which could enable the principal of the contract to implement the optimal long-term contract without being constrained by short-term considerations. Rey and Salanié (1990).

\textsuperscript{28} Lambert (1983) and Rogerson (1985).

\textsuperscript{29} Williamson (1985).


\textsuperscript{31} Malcomson and Spinnewyn (1988), in the context of a principal-agent model under asymmetric information show that, whenever there is no renegotiation, a long-term contract can perform better than a short-term contract when one of the parties – principal or agent – commits to a payoff which is lower than the potential outcome under a short-term contract. Therefore, long-term contracts are endowed with some commitment element that may cause them to be preferred to a series of short-term contracts.

\textsuperscript{32} Hart and Holmstrom (1987).

\textsuperscript{33} Crawford (1988). Some have suggested that in these cases long-term contracts serve as a substitute for an efficient credit market.

\textsuperscript{34} Dutta and Reichelstein (2003).
indicator variables compared to short-term contracts, and the conditions under which parties do better with a series of short-term contracts than with a long-term contract. In the context of principal-agent contracts, they conclude that long-term contracts create incentives for the agent to overinvest – or to invest inefficiently – due to an existing moral hazard problem by which the agent knows that he will not assume any risk for overinvesting or for not making the most efficient investment decision on behalf of the principal. In these cases, the principal does better with several short-term contracts that may entail agent rotation; that is, there may be a new agent in the second period. As a consequence, the agent’s incentives to overinvest are controlled because he knows he may be replaced in the next period.

Chiappori et al.\(^\text{35}\) show, in turn, that in order for the performance of short-term and long-term contracts to be equivalent, two conditions are necessary: renegotiation should not be possible so that the commitment value of long-term contracts would be reduced and short-term contracts should allow the smoothing of consumption.\(^\text{36}\) But it should be noted that the length of long-term contracts is not the most important or essential parameter; what is important is to what extent a contract of a given length locks the parties into the relationship.\(^\text{37}\)

### 3. Factors to Consider when Determining the Optimal Contract Length

As explained above, depending on the context, the parties’ interests, their size, characteristics, informational structure, and their attitude towards risk, a long-term contract may or may not perform better than a series of short-term contracts. There is no general optimal contract for all situations but an optimal contract and an optimal length for each kind of relationship.

The relationship between performance and contract length has been widely studied.\(^\text{38}\)

#### 3.1. Contract Length/Contract Price Trade-off

When one of the contracting parties is more risk averse than the other, this party is willing to trade a lower consideration for a longer contract.

---

\(^{35}\) Chiappori et al. (1994).

\(^{36}\) Fudenberg et al. (1990); Malcomson and Spinnewyn (1988); and Rey and Salanié (1990).

\(^{37}\) The length that should be considered is the nominal length of the contract. See Aghion and Bolton (1987).

\(^{38}\) Particularly in the context of sports’ contracts, see Krautmann and Oppenheimer (2002), Maxcy (1997, 2004).
Therefore, higher certainty in terms of a longer contract will compensate for a lower contract price.\footnote{Krautmann and Oppenheimer looked at the relationship between player salaries and contract length, and suggested that when players are more risk averse than club owners, they are willing to receive a lower salary in exchange for a longer contract. Krautmann and Oppenheimer (2002).}

Further, when there is high uncertainty regarding the contract price relative to uncertainty in production, long-term contracts will be preferred to a series of short-term contracts.\footnote{Maxcy (1997; 2004).}

### 3.2. Contract Length and Attitudes towards Risk

The attitude of contracting parties towards risk has also an effect on the contract length. Intuitively, it is easy to foresee that the more risk averse one of the contracting parties is, the more eager this party will be to insure.\footnote{Maxcy (1997; 2004).} This means that when a contract party is risk averse, she will prefer a long-term contract rather than a series of short-term contracts.

### 3.3. Contract Length and Incentives to Invest

In theory, a long-term binding contract should be able to induce more investments from the parties. However, the literature suggests that the evidence is not so conclusive.\footnote{Dutta and Reichelstein (2003).} Even though short-term contracts result in underinvestment, the optimal long-term contract may result in under- or over-investment depending on the importance of agency problems.\footnote{Dutta and Reichelstein (2003).}

What the literature seems to agree on is in the fact that contract length is determined as a trade-off between the costs of entering into a new contract – the costs of re-contracting – and the costs associated with the incompleteness of the contract.\footnote{Williamson (1985) and Aghion and Bolton (1987).}

### 3.4. Contract Length and Parties’ Performance

Contract length may also affect the incentives of the parties to perform, and it could be expected that the longer the contract term, the lower the return on parties’ performance.\footnote{Krautmann and Oppenheimer analyzed the players’ return to performance by contract length and concluded that the returns to performance decline with contract length. Krautmann and Oppenheimer (2002).} Given the unverifiability of performance in regard to many dimensions of the parties’ obligations, however, the use of quasi-rents within the relationship, enhanced by the presence of specific
investments, requires indefiniteness in the contract term, which is easier to obtain with a longer-term contract with the possibility of termination.\textsuperscript{46}

Further, the longer the contract lasts, parties’ performance tends to vary depending on the point reached in the life of the contract. Maxcy et al.\textsuperscript{47} looked for ex ante behavior, before the new contract is signed and ex post behavior, once the contract is signed and concluded that there was high performance in the last year before a new contract was signed and such performance would diminish after the first year after the contract had been signed. So performance may not be smooth across all contract phases.

4. Other Important Issues Inherent in Long-term Contracts

4.1. Long-term Contracts and the Creation of Barriers to Entry

The first to note the competitive importance of long-term contracts for parties entering into them were Aghion and Bolton.\textsuperscript{48} In a well-known 1987 paper, they noted that an implicit effect of long-term contracts is the bilateral monopoly created by this kind of contract between the parties. By locking themselves into a long-term contract, parties significantly reduce the probability of entry by third parties, so that competitors are denied market access. Therefore, the size of the market of the other contracting party is reduced where each contracting party may become a monopolist with respect to each other. This situation may be exploited by the other party in order to maximize the surplus obtained from the other contracting party.

This has led many authors in the economic literature to consider long-term contracts to possess strong anti-competitive elements, because the monopoly created by the parties implies a negative externality of reducing the scope of the competitive alternatives potentially offered to the object of the contract.\textsuperscript{49}

Aghion and Bolton\textsuperscript{50} first claimed that long-term contracts were in

\textsuperscript{46} See, infra, section 6 on contract renegotiation and breach and section 7 on termination and remedies.

\textsuperscript{47} Maxcy et al. compared performance for a three-year average contract length in both the year before and the year after a new contract is negotiated. In light of this, they presented a theoretical model in which long-term contracts are desired by sport clubs in order to mitigate both market uncertainty and uncertainty about athletes’ future productivity. See Maxcy et al. (2002).

\textsuperscript{48} Aghion and Bolton (1987).

\textsuperscript{49} Sibley (2002).

\textsuperscript{50} Aghion and Bolton (1987).
Long-term contracts in the law and economics literature

Long-term contracts in the law and economics literature 325

general socially inefficient because they were frequently signed for entry-prevention purposes so that they would block and deter entry. They showed that the contract length will depend on the informational assumption about the incumbent’s costs and finally, whenever there is asymmetric information, the length of the contract will signal the incumbent’s cost.

Long-term contracts would be preferred by sellers who faced a threat of entry into the market in order to prevent the entry of more cost-effective producers.

One of the key assumptions of Aghion and Bolton was that the entrant could observe the incumbent’s costs when making any entry decision. Poitevin\(^1\) showed that by changing this assumption and considering that the incumbent signals both to the buyer and to the entrant, the results of Aghion and Bolton change significantly. If, instead, the entrant cannot observe the incumbent’s costs,\(^2\) Poitevin shows that the nominal length of the contract does not signal the incumbent’s costs since the incumbent always signs a contract regardless of its cost level; second, entry will be completely deterred.\(^3\)

Rasmusen et al.\(^4\) noted that long-term contracts should not be interpreted as exclusionary instruments based on parties being a monopoly and therefore having market power, but as exclusionary agreements that enable an incumbent monopolist to exclude its rivals cheaply by exploiting customers’ inability to coordinate their actions.\(^5\)

Segal and Whinston\(^6\) extended Aghion and Bolton’s analysis and assumed that long-term contracts were complete barriers to entry so that contracting parties could monitor and enforce a fully exclusionary contract. The resulting bilateral monopoly between contracting parties and the risk created by specific investments lead to a contractual dynamic that directly impacts on third parties, such as financial partners or banks, for example.\(^7\) Anticipating such a scenario, contracting parties are hesitant

---

\(^1\) Poitevin (1992).
\(^2\) Poitevin justifies modifying this assumption and considers the case where the entrant cannot observe the incumbent’s costs in the literature on ‘expectational entry deterrence’ such as Milgrom and Roberts (1982).
\(^3\) Poitevin (1992).
\(^5\) In Rasmusen et al.’s model, they consider the scope of entry deterrence by allowing more than one customer, so that fixed costs are covered. Under these assumptions, they found two pure-strategy Nash equilibria in which either enough customers signed an exclusionary agreement and deterred entry or no customers signed the agreements. See Rasmusen et al. (1991).
\(^6\) Segal and Whinston (2000a).
\(^7\) Dailami and Hauswald (2000).
to make relationship-specific investments without adequate contractual protection.58

A common assumption of the above-mentioned papers is that the exclusive contract between the incumbent and a final consumer exerts some externality on third parties so that exclusive dealing is extremely powerful in deterring entry. Fumagalli and Motta59 modified a major assumption of the literature and assumed that one of the parties, instead of being a final consumer, was a firm that used the input bought either from the potential entrant or from the incumbent in order to resell it on a final market.60 Hence, such firms would compete in a downstream market where profits would depend on the one hand on the input price, which would determine their input demand, and on the other hand, on the price paid by other competing buyers. In their model, they showed that downstream competition could eliminate the incumbent’s incentives to exclude based on two effects: first, by making the demand of a single buyer large enough to attract entry, the negative externality that a buyer would exert on others by accepting an exclusive agreement would disappear and second, it could enhance profitability when more efficient than rivals. Consequently, they noted that the potential for using long-term contracts as anti-competitive instruments would significantly depend on the intensity of competition in downstream markets.61

4.2. Specific Investments and Contracting Forms

In a contractual relationship, both parties may need to make certain investments in order to fully exploit the gains from trade. These investments may be necessary for any, or almost any, kind of contractual relationship or they may be specific to the obligations imposed by the contract itself, so that they will not be valuable outside the contract or outside the parties’ relationship. Thus, specific investments, once made, cannot be used in other relationships or businesses or have no value outside the contract or the parties’ relationship. Hence, their value within the contract is always higher than it would be in any other alternative use.

The consequences and risks inherent in making investments in a relationship are very different for the contracting parties depending on whether the investments are general or specific. So, while general investments may be used in other contractual relationships and, therefore, do not entail an

58 Dailami and Hauswald (2000).
60 Fumagalli and Motta (2006).
inherently high risk, specific investments are very risky because they are not valuable in other contractual contexts and become sunk costs.\(^{62}\) These specific investments are especially important in the context of long-term contracts because the discrepancy between the value of the assets within and outside the relationship often tends to be positively associated with the length of the contract. Moreover, in a long-term contract, due to a higher lock-in effect, one would expect more acute conflict between the parties in the contract,\(^{63}\) given that the party who has made higher specific investments is in a more vulnerable position than the other and will be more likely subject to the possibility of holdup.\(^{64}\) Of course, in some, or even many, settings, specific investments are crucial for adequately exploiting the gains from the interaction, and are essentially unavoidable.

Williamson\(^{65}\) distinguished between four different kinds of specific assets: physical capital specificity, which stems from investments that involve tools or other physical assets that have higher value in their intended use rather than in any other use; human capital specificity, which results when individuals enhance their human capital, the value of which is higher in the relationship than outside it; dedicated assets, which are made – in a factory plant, for example – because they have a certain value where they are invested and not in any other place; and site specificity, which refers to the quasi rents generated by savings in inventory and transport costs under vertical integration, for example. A fifth kind of asset specificity, time specificity, refers to assets that must be used in a certain order or under a certain schedule.\(^{66}\)

The most important holdup risks are manifest in breach of contract or in unilateral renegotiation.\(^{67}\) Goldberg points out that exposure to the risk of holdup depends on the access to market alternatives, so that the more difficult it is for contracting parties to have access to alternative markets, the more significant the risk of opportunistic behavior would be. If, instead, parties were to have access to market alternatives, they would not be so vulnerable to other parties’ strategic behavior and this could

---

\(^{62}\) For an application of specific investments in the natural gas context, see Hubbard and Weiner (1986).

\(^{63}\) See Klein (1980); Williamson (1985).

\(^{64}\) De Geest (2006).

\(^{65}\) Williamson (1983).

\(^{66}\) Masten et al. (1991); Pirrong (1993); Williamson (1991).

\(^{67}\) See infra. These risks do not only affect contracting parties, given that they are also transferred to debt-holders of contracting parties who would be indirectly vulnerable to opportunistic and strategic behavior by the contracting parties. See Dailami and Hauswald (2000).
minimize the price divergence between the contract price and the opportunity costs of the parties.\footnote{Goldberg (2000).}

From a structural perspective, Joskow examines the role of specific investments in the choice of vertical integration or entering into long-term contracts, and notes that the relevance of specific investments in vertical relationships strongly determines the decision to integrate vertically or enter into contracts.\footnote{Joskow considered that the allocation of risk between buyers and sellers is also an important factor in the context of vertical relationships: see Joskow (1990).}

In general, there are two instruments that the literature has discussed in order to minimize the expected costs of making relationship-specific investments: vertical integrating\footnote{Joskow (1990).} or entering into a long-term contract.\footnote{In the context of coal markets, see Joskow (1990).}

It should be noted that neither of them is generally better than the other. Coase famously suggested that when the relationship between transaction costs and organizational form is not precise, there is more than one organizational response to a transaction costs problem.\footnote{Coase (1988).}

Further, Holmström and Roberts claim that the relationship between transactions costs and organizational form is many-to-many: there are different governance tasks and various instruments for managing them. Each task can be addressed by more than one instrument, and each instrument can, alone or in combination with others, be used to address more than one task.\footnote{Gilson et al. (2009).} However, the two major instruments will be discussed now.

(i) The vertical integration temptation
There is support for the existence of an important relationship between specificity and the structure of vertical relationships.\footnote{Hart (1988); Williamson (1985); Klein (1980); and Joskow (1988).}

Specific investments made by contracting parties often affect the structure of vertical relationships,\footnote{Joskow (1990).} and even the risk of opportunism will sometimes drive contract parties away from contracts and toward vertical integration.\footnote{Industrial organization theory predicts that when parties in the supply chain have to make transaction-specific investments, they will vertically integrate. See Gilson et al. (2009).}

When the value of the investments made by contracting parties is
mutually dependent, each investor tries to induce the other to invest first in order to extract more favorable terms once an irrevocable commitment has been made and the specific investments are sunk costs. In order to avoid that, many parties understand that placing both assets under the control of a single owner, and therefore vertically integrating, unblocks this situation because the incentives and risk of holdup disappear.77

This is what happened with the famous acquisition by General Motors of Fisher Body, where General Motor merged vertically with Fisher Body, a maker of auto bodies. When this case was analyzed in the early literature, the merger was thought to reflect a market failure or contracts as a result of asset specificity – specific investments – and opportunistic behavior.78 The discussion evolved, with vertical integration between the two companies seen as a response to an interest in improving the coordination of production and inventories, as well as assuring General Motors an adequate supply of auto bodies and access to the talent of Fisher Body officers.79 However, opinion today is divided and there is still discussion regarding the motivation that resulted in the merger between General Motors and Fisher Body.80

(ii) Contractual solutions to the holdup problem

When a contract is used to govern a transaction in which the consequences from holdup are significant due to the presence of relationship-specific investments, contracting parties, aware of the problem, may well try to solve it ex ante and incorporate safeguards in the contract in order to protect these investments from opportunistic behavior from the other party.81

But when parties decide to draft a contract in order to minimize the risk of holdup, the major challenge contracting parties face is drafting a sufficiently complete contract so as to adequately mitigate the risk of appropriability of the specific investments made by one of them.82 Empirical research on this issue shows that the existence and relevance of specific investments is positively correlated with contractual completeness.83

---

77 Gilson et al. (2009).
78 Klein et al. (1978); Williamson (1985); Hart (1995).
80 See Coase (2000, 2006) and Klein (2006), where the author claims that the contract adjustment between General Motors and Fisher Body demonstrates the importance of distinguishing between a threat of an inefficient holdup and the economically efficient way in which it takes place.
81 Poppo and Zenger (2002); Goldberg and Erickson (1987); Joskow (1988).
82 Joskow (1988); Goldberg and Erickson (1987); Poppo and Zenger (2002); Reuer and Ariño (2007).
83 See Poppo and Zenger (2002); Goldberg and Erickson (1987); and Joskow (1988).
But contract drafting does not entirely solve the contract risks or hold-up problem presented by specific investments and the parties' incentives to appropriate these investments made by the contracting parties. The problem of contract drafting and of completeness is that while it solves certain issues, such as being able to address and anticipate potential contract risks, including the risk of appropriability, it also reduces contract flexibility so that specificity sometimes makes it difficult to adjust to potential contract risks that are not foreseeable when the contract is drafted. In this sense, the commitment necessary to create incentives for making specific investments increases the contract surplus but at the same time conflicts with the flexibility needed in long-term transactions.

Further, the contractual solution to relationship-specific investments could also entail significant costs in terms of litigation and might not be a perfect instrument. Joskow noted that a long-term contract could help protect specific investments made by buyers and sellers but it is very imperfect given the impossibility of drafting a complete contract that, once signed, and once specific investments are sunk, cannot adapt to changes in market conditions, and therefore not be entirely able to avoid the holdup problem for the party making the investment.

Despite the challenges presented by the contract solution to holdup problems, it is worth noting that the transaction cost theory claims that learning between contracting parties improves contract design and the final outcome and performance of the parties. When firms learn from and have experience of each other they are able to assess the risk of opportunistic behavior by the other, and learn to design more complete contracts. Hence, they do not need to anticipate and contract ex ante over such contingencies, but may cooperate and solve the situation when it appears. Thus, by cooperating and learning from each other, parties gain experience to identify the risks involved in the relationship and how to efficiently address such risks when they occur. This may be particularly relevant in industries with high investment costs and high innovation elements.

Gilson et al. summarize the evidence on vertical disintegration in technology-based industries. Producers cannot innovate in cutting-edge
technology in every field required for the success of their product, and increasingly companies choose to buy innovation from other companies through contracts.91 But specific investments ex ante are significantly high, and collaboration and a long-term relationship between both parties are necessary. This environment is what Gilson et al. characterize as contracting for innovation.

Thus, cooperation ex post is necessary in order to interpret the uncertainties contained in the contract and to renegotiate contract clauses when necessary. The long-term nature of the contract may increase associated uncertainty and risk, albeit it will be compensated by learning about the parties’ propensities to behave opportunistically, which will consequently be significantly reduced and therefore the relationship between the contracting parties will strengthen. Through ongoing cooperation, contracting parties can design governance and dispute resolution mechanisms in long-term contracts that will increase contract surplus.

Cooperation and learning will directly impact the contract design through creating knowledge and routines that raise the parties’ switching costs and through devising a dispute resolution mechanism that builds mutual knowledge of the propensity to reciprocate while deterring opportunistic behavior that could undermine the cooperative equilibrium.92 Consequently, the collaboration process itself raises the costs of taking advantage of the other party’s specific investments.

This account of contracting for innovation fits into the more general hypothesis presented in the transaction cost literature concerning how firms develop governance mechanisms in their inter-firm relationships in order to reduce transaction costs and thus to become more efficient.93

5. Contract Drafting and Design

Truly complete contracts do not exist.94 They are ‘hypothetical contracts that describe what action is to be taken and payments made in every possible contingency’.95 But under incomplete information conditions, it is too costly for contracting parties to foresee and contract over all potential outcomes or contingencies that could take place during the life of the contract. Contracting costs take place both ex ante, when anticipating contract contingencies and outcomes, and ex post, when the contract has to be enforced. It is exceedingly costly to specify all potential states of

91 Gilson et al. (2009).
92 Gilson et al. (2009).
95 Milgrom and Roberts (1992).
the world, as well as to prove that one such state took place. Uncertainty about the future and the cost of writing complete contracts are essential elements when determining the contract length.96

The difficulty of drafting complete contracts does not necessarily mean that parties have no incentives to take into account as many contingencies as possible in order to minimize contract uncertainty as much as they can. Contracts generally include the rights and obligations of each party, the solution to potential contract contingencies, and how the relationship between contracting parties will be structured and governed. And, as firms gain experience, they probably learn to design more complete contracts.97 Still, long-term contracts cannot completely specify in advance all the obligations of both parties over the life of the agreement, and in order to adapt their relationship to changing circumstances they will find it necessary to give one, or both, parties the discretion to respond as new information becomes available.98

Increased contract completeness in long-term contracts, while having positive properties – it may solve issues such as anticipating potential contractual risks, including the risk of opportunistic behavior by parties – also reduces contract flexibility so that the contract becomes more difficult to adjust to risks and new contingencies, and is, thus, no panacea. The commitment necessary to create incentives for parties to make specific investments and maximize the contract surplus could be seriously undermined by the lack of flexibility that is needed in long-term transactions.99

When drafting a long-term contract, parties have to define the contract terms either before, or after, the specific investment is made. Defining contract terms before any specific investment is made requires information that may or may not be yet available. Drafting a contract after specific investments have been made may distort the power structure of the relationship between contracting parties because one of the parties, specifically the one making the specific investment, will not be totally free to decide, in light of the specific investment having already been made. Hence, when drafting a long-term contract, parties have to choose between an uninformed decision, or a subsequent potentially distorted decision.

96 The key question when analyzing the relationship between contract length and contractual incompleteness is determining what contingencies parties should leave out of the contract. Aghion and Bolton (1987).
97 Contracts cannot be completed without having a previous experience of different problems and contingencies arising from former exchanges. See Solis-Rodriguez and González-Díaz (2009).
98 Goldberg (2000).
99 Gilson et al. (2009).
Parties may respond differently to informational problems in terms of the completeness of the contract. There is a taxonomy\textsuperscript{100} of responses that looks at over-completeness, under-completeness and a mix between these two. Over-completeness is one of the possible strategies for parties in a long-term contract, because parties may want to draft an extremely detailed contract to try to anticipate most contingencies. A possible advantage of some over-completeness is the fact that when parties negotiate the terms, they have not made the specific investments necessary for the performance of their contractual obligations, and the solutions and safeguards in the contract may improve the incentives for those investments. At the same time, given the time frame, the specific provisions in the contract will very likely be interpreted and enforced under different conditions and in a different context from that in which they were drafted, and this may cause courts to reinterpret them and often not to enforce them. Sometimes, though, in light of the long nature of the relationship, parties may consider that leaving some contingencies or circumstances open could be a good strategy in order to be able to renegotiate their relationship in light of evolving contract conditions, or to allow a third party, such as a court, to fill the gaps left by them. Under-completeness presents the drawback that when parties renegotiate or where a third party fills the gap, specific investments have already been made, and parties are in a monopoly situation with each other.

Trade-offs of this kind have been explored in the literature. Hubbard and Weiner\textsuperscript{101} model a contract that could solve the bilateral bargaining problem in the natural gas context where there was a high amount of transaction-specific capital with little value outside the relationship. Once the gas well development costs are sunk, a pipeline faces the temptation to appropriate some of the rents from production. In light of this, the producer demands a long-term contract with adjustment clauses beforehand.\textsuperscript{102} Many mechanisms, such as flexible pricing, for example, have been created in order to decrease the return on opportunistic behaviour by one of the contracting parties once the specific investment has been made.\textsuperscript{103}

The inequality in the parties' position concerning the need and cost of specific investments has also been discussed. When they are unequally distributed between parties, each party's negotiation position will depend on whether the contract terms are fulfilled, or whether if the contract is

\textsuperscript{100} De Geest (2006).
\textsuperscript{101} Hubbard and Weiner (1986).
\textsuperscript{102} Generally, guaranteed supply clauses, price and take or pay provisions, because gas well companies are best operated near full capacity. See Hubbard and Weiner (1986).
\textsuperscript{103} Goldberg (2000).
breached, the party who invested more in the relationship will lose more than the other party.

Some argue that it would be desirable for this purpose to have both parties investing symmetrically in the contract relationship so that the opportunistic behavior of parties trying to renegotiate would decrease and that if renegotiation were to take place, the party who invested more would receive the most.

The argument, however, does not take into account that many other factors, such as position in the market or in the industry are also relevant factors when deciding whether to strategically renegotiate the contract terms. Even if parties make specific investments of equal amount, if they occupy unequal bargaining positions or have a different share of the market or a different position in the market, so that they need to create a reputation among others, they may still have incentives to strategically renegotiate or oppose renegotiation.

Dailami and Hauswald analyze how in the face of contractual incompleteness, contract risks are transmitted and allocated between different contracts and investors, in particular in the context of the relationship between the off-take and financial contracts because the former serves as security for the latter, since such long-term supply contracts are necessarily incomplete and subject to opportunistic behavior.

As explained in the previous section, incompleteness is inherent in contracts, and large degrees of incompleteness are pervasive in long-term contracts. Hence, parties will leave out of their contract and therefore out of their negotiations certain terms regulating certain contingencies that may or may not take place. This uncertainty implies that, despite the commitment present in long-term contracts, there is an element of potential conflict and disagreement between parties when one of these unanticipated contingencies takes place. Most contracting parties, aware that conditions change during the life of the contract and of the extremely costly and even impossible option of foreseeing and allocating contract risks, may prefer to adjust the contract to the changed circumstances.

---

106 Dailami and Hauswald (2000).
107 Dailami and Hauswald (2000).
109 Sometimes, even when parties do not expressly state it, it is clear from the contract terms that they intended to do so. See Goldberg (1985).
Joskow distinguishes between the voluntary and involuntary renegotiation that would take place when contractual terms are left unperformed, amounting to a breach of contract.\footnote{Joskow (1990).} In the first case, when both parties want to renegotiate contract terms, it will not be possible to hold parties to the contract.\footnote{See generally Mahoney (2000).} In the second case, the case of breach, the non-breaching party may either enforce the contract whenever possible, or seek damages depending on what remedies are available to the court and how significant the costs of using the legal system are.\footnote{For example, if the remedy were specific performance, the contract could only be enforced if both parties agreed on such a remedy because otherwise specific performance will not be an available alternative. See Jolls (1997).}

Parties may want to renegotiate the contract and agree as to how to resolve the contingency, or accept that there has been a breach of contract so that it will be for courts to determine whether the contract has been breached and what the parties will be required to do as a result. However, it should be noted that the incentives and consequences of incomplete contracts are different depending on whether the solution was reached by and between contracting parties through renegotiation, or was provided by the court.

6.1. Renegotiation

If parties had perfect information, it might be possible for contracting parties to define and determine the circumstances of renegotiation and even to exclude it, if parties would prefer to do so.\footnote{Hence, if parties had complete information, they could draft contracts that could even exclude – if at all possible – the possibility of renegotiation, and therefore would be renegotiation-proof.} Hence, if parties had complete information, they could draft contracts that could even exclude – if at all possible – the possibility of renegotiation, and therefore would be renegotiation-proof.\footnote{Contracts which are renegotiation proof can still improve the ex ante versus ex post problem – ex post sharing of the contract surplus may not be ex ante efficient – presented by the possibility of renegotiation by designing the environment of the renegotiation in the unverifiable state for a third party and where the information between parties is symmetric. See Hart and Moore (1988) and Dewatripont (1989).}

Given that it is not possible for contracting parties to draft complete contracts, they will enter into incomplete contracts so that renegotiation will be necessary whenever an unforeseen contingency takes place.\footnote{When information costs are high, parties may emphasize ex post efficiency rather than ex ante efficiency and will seek to balance both elements and draft a formal contract with vague standards so that it may be renegotiated ex post and adjusted to the contingencies that may take place during the life of the contract. See Gilson et al. (2009).}
Based on the commitment of the parties to each other, each party makes specific investments and relies on renegotiating the contingencies that are not specifically addressed in the contract.\(^{116}\) There may be scope for renegotiation when ex ante efficiency determines ex post inefficiency.\(^{117}\)

The parties’ incentives for renegotiating the contract might not be symmetric when the parties’ reputation is considered. So reputation may constrain one party’s incentives to renegotiate or unilaterally adjust contract provisions so that the likelihood of opportunistic behavior by this party is low. As a consequence, the other contracting party, having less reputational interests to protect, may even benefit from the greater discretion of the more reputed party in determining the desired performance.\(^{118}\)

The prospect of renegotiating presents advantages and disadvantages for parties. On the positive side, it has the clear advantage of creating flexibility to achieve ex post efficiency without incurring high costs ex ante in trying to foresee all potential contingencies between the parties.

If renegotiation is taken into account once the contract is drafted, ex post opportunistic behavior may be either reduced or even totally eliminated. Gilson et al. suggest that in order to eliminate the risk of hold-up, renegotiation should be regulated so that parties would determine how to share the benefits and surplus created by the contract and therefore assure both ex ante and ex post efficiency, whenever possible.\(^{119}\)

Once parties have renegotiated and the uncertainty is solved, they can perform the contract, modify the terms of the contract, withdraw from the transaction, or write a new contract. The renegotiation solution can achieve an ex post efficient result through negotiation as the Coase theorem\(^{120}\) predicts: if the contract were to be profitable for one of the parties, the other party, who would not be interested in enforcing the contract, could bribe the other contracting party in order to withdraw the contract and therefore not to enforce it.

On the downside, renegotiation also presents serious disadvantages. The choice of whether to renegotiate contract terms voluntarily will depend on how likely this renegotiation is going to be. As mentioned earlier, if there is room in the contract for renegotiation, it will most

---

118 This result was obtained by Arruñada et al. (2001) in the context of the automobile distribution industry in contracts between manufacturers and auto dealers.
119 Gilson et al. (2009).
120 Coase (1937).
likely take place, and hence, contract breach and related litigation will be avoided. It is difficult to see how the legal system, under freedom of contract, can prevent renegotiation – which is a form of contracting – to happen if parties leave room for it in the contract. But if renegotiation is likely and parties are aware of it, long-term contracts would be a very inefficient instrument for parties because the value of commitment – which is one of the good qualities of contracts in general, and especially one of the added values of long-term contracts – will be substantially eroded, and therefore parties would not have proper incentives to make efficient specific investments.

First, there is the moral hazard problem arising from the level of effort that an agent might choose. If the agent chooses first the level of effort, the principal will not be able to make the agent choose the optimal level of effort, and renegotiation will result in inefficiencies.121

Also, renegotiation raises the chances of holdup for the party making the investments and undermines the incentives to make specific investments in the first place.122 This causes serious inefficiencies and undermines the ex ante advantages presented by renegotiation. When there is uncertainty, parties seek to minimize contracting costs and balance the benefits of commitment and flexibility with the costs of uncertainty and the risk of potential holdup.123

Further, whenever parties are aware that renegotiation will take place, the literature124 has shown that the revelation of information between contracting parties slows down because of the trade-off between incentives to renegotiate and the incentives to reveal private information.

Finally, as can easily be foreseen, renegotiation is costly. Goldberg125 has noted that parties, aware of this, anticipate the potential costs of renegotiation and introduce price mechanisms in long-term contracts in order to avoid the costs of renegotiation, thereby increasing the expected value of the long-term contract. The larger the uncertainty or the variance of contract outcomes, the more resources would be devoted to the contract drafting effort.126 Hence, anticipating future renegotiation costs may increase the contract drafting process.127

121 Chiappori et al. (1994).
122 Gilson et al. (2009).
123 Gilson et al. (2009).
127 Goldberg and Erickson (1987).
6.2. Court Adjustment

Long-term contracts include a component of commitment and cooperation so that often parties will settle their differences so that they can maintain a constructive relationship and preserve the businesses’ goodwill and reputation. But reaching a renegotiated solution between themselves sometimes will not be possible. Hence, whenever contracting parties are not able to reach an agreement and fill the contract terms in order to address the unforeseen contingency, a court will have to adjust, interpret or fill the contract gaps.

Whether and how courts should interpret incomplete contracts is a highly debated matter. Schwartz suggests that any analysis of contractual interpretation should answer two questions: first, whether courts use broad or narrow evidentiary bases in determining the meaning of the contract’s language and second, whether courts should always admit the possibility that the parties wrote in a private language and therefore should be entitled to provide an interpretation of the incomplete contract.

The first issue refers to the question of what evidence should be admitted in order to interpret the incomplete contract. When courts decide contract cases where contracts are incomplete, courts generally pursue three strategies: protecting process values, interpreting language and filling the contract gaps so that they supply terms when the parties’ contract fails to provide for the dispute that divides them. The norms and language that should be used in order to decide contract cases is not unanimous among law and economics scholars: while some consider that courts should use norms that transcend the relationship, such as fairness, others understand that such norms should be provided by normative desirable terms that parties should be free to vary. But it should be noted that when contracts are incomplete as a consequence of parties’ asymmetric information, courts should not follow the above-mentioned strategies and treat incomplete contracts as if they were complete, so they interpret the contract as written.

---

129 It is often argued that the party making the highest specific investment faces a higher cost the more specific is the investment because of its exposure to the risk that the other contracting party will walk away from the contract, especially when there are other suppliers available in the market. Therefore, one of the parties – the one assuming a lower specific investment – has an implicit real option through breach of contract. See Dailami and Hauswald (2000) for a discussion of this situation in the context of the Ras Gas Project.
130 Schwartz and Scott (2003).
Some authors claim that the goal of the court’s interpretation of contracts should be to facilitate the efforts of contracting parties to maximize the joint gains – the contractual surplus – from transactions. Another theory, the negative claim theory, suggests that this is all courts should do.\(^{134}\) In the US, a majority of jurisdictions applies a literal interpretation of contract terms, mostly based on an application of the plain meaning rule.\(^{135}\) In Europe, the approach suggested by the European Commission in the Principles of European Contract Law is to give effect to the intention of the parties regardless of whether this intention is reflected by the works used.\(^{136}\) Scott\(^{137}\) justifies formal interpretations because they offer the best prospect of maximizing the value of contractual relationships, especially considering that contract interpretation often finds competent parties together with incompetent courts.\(^{138}\)

Nevertheless, there is a general tendency in the literature to favor restricting the evidence that courts may use to interpret a contract based on the argument that parties should be allowed to save costs from contract interpretations on minimal evidentiary bases even if, in any given case, the odds of an accurate interpretation would be higher with a broader base.\(^{139}\)

Once a court has determined the evidence admissible to interpret the contract terms, the second issue is whether parties’ language in the contract

\(^{134}\) Schwartz and Scott (2003).
\(^{135}\) Scott (2000).
\(^{136}\) Article 5.101(1) of the European Principles of Contract Law provides that

(1) A contract is to be interpreted according to the common intention of the parties even if this differs from the literal meaning of the words.
(2) If it is established that one party intended the contract to have a particular meaning, and at the time of the conclusion of the contract the other party could not have been unaware of the first party’s intention, the contract is to be interpreted in the way intended by the first party.
(3) If an intention cannot be established according to (1) or (2), the contract is to be interpreted according to the meaning that reasonable persons of the same kind as the parties would give to it in the same circumstances.

The European Principles of Contract Law can be found at: http://frontpage.cbs.dk/law/commission_on_european_contract_law/PECL%20engelsk/engelsk_partI_og_II.htm.

\(^{137}\) Scott (2000).
\(^{138}\) Scott (2000).
\(^{139}\) Schwartz and Scott (2003).
before the court or the court itself should provide the interpretation. Schwartz and Scott\textsuperscript{140} argue that the parties’ sovereignty in the contract requires courts to delegate to them the choice of the contract’s substantive terms and the interpretative theory that should be used to enforce those terms.\textsuperscript{141} Firms and contracting parties are suited to creating their own contracts, while the state is best suited to create the broad structure within which parties’ contracts fit. Hence, their roles are different and they should act accordingly.

In order to have a successful intervention by courts in providing a contract solution that parties have not managed to attain by themselves, a certain level of definiteness will be needed.\textsuperscript{142} Generally, when contracts include clear terms and define parties’ obligations, parties may be able to avoid many disputes of interpretation, for example.\textsuperscript{143} But clear provisions will also be useful when courts need to act, because if courts could not enforce the contract terms, some efficient transactions may be deterred and inefficient ex post negotiations may take place.\textsuperscript{144} Court adjustment may be appropriate in some circumstances that are sufficiently identifiable.\textsuperscript{145} Whenever contract clauses are not definite or clear enough, some courts will disregard them and invoke the parties’ good faith when determining the solution of the parties’ dispute.\textsuperscript{146}

The contract literature disputes whether court-imposed solutions possess advantages. Dawson\textsuperscript{147} opposes court adjustment of long-term contracts because he understands that courts lack sufficient standards to redesign the contract so that it reflects the parties’ ex post agreements based on what they would have agreed ex ante. Dawson understands that courts enjoy an excessively unlimited discretion to create a new contract and further, that the parties’ duty to adjust can override express contract terms.

Different factors may predispose courts to activism when adjudicating

\textsuperscript{140} Schwartz and Scott (2003).
\textsuperscript{141} Schwartz and Scott (2003).
\textsuperscript{142} Joskow (1990).
\textsuperscript{143} Hillman (1987).
\textsuperscript{144} Joskow (1990).
\textsuperscript{145} Dawson (1984).
\textsuperscript{146} See Goldberg (2000) regarding contract clauses concerning quantity boundaries. Goldberg claims that contract tailoring by parties may create an incentive for them to take into account and hence internalize their reliance interests, which could be done much more efficiently by parties than using the good faith standards of the courts.
\textsuperscript{147} Hillman analyzed the different approaches in the literature to the court interpretation and adjustment of incomplete contracts. See Hillman (1987).
incomplete contracts.\textsuperscript{148} process values are offended in the contract formation or in the court of performance; enforcement of the contract adversely affects third parties; the contract directs a result that is substantially unfair to one of the parties; and the contract is incomplete and the court can complete it with a term that parties will accept and courts will be able to apply. But the court’s scope of discretion is not unlimited. The parties’ bargaining over a contract term and the parties’ purpose in including a certain provision are crucial in determining whether the court adjustment should trump an express contract provision.\textsuperscript{149}

Further, another important issue beyond the scope of this chapter is to determine how courts should approach a contract dispute in order to properly decide the issues at stake.

7. The Content of the Contract: Breach, Termination, Remedies, and Non-Compete Clauses

7.1. The Problem of Verifiability of Non-performance
We have already highlighted the essential incompleteness of long-term contracts. Even if incomplete contracts are the inescapable rule in any ordinary setting of economic interaction, where long-term contracts are the norm, the extent of incompleteness is even stronger and more decisive. Long-term relationships have, by their very nature, a more extended time horizon than spot or short-term contracts. The number, influence, complexity, and difficulty and cost in anticipation, of contingencies that can, in one way or another, have an effect on contractual outcomes, dramatically increases. This is why economists, when approaching long-term contracts, have routinely assumed that contracts between the parties are incomplete.

This makes the long-term relationship typically a relational contract: many of the relevant actions cannot be foreseen and specified when the contract is signed, and it is in the course of the ongoing relationship that the parties will adopt those actions, based upon a set of incentives arising from factors (personal, institutional) that differ from the formal contract and legal rules in contract law.

Of course, a relational contract in this sense cannot be enforced purely

\textsuperscript{148} Schwartz (1992).

\textsuperscript{149} Courts face difficult challenges because good faith is the parameter applicable to parties when they interpret a contract term and such a principle is based on business expectations but courts face the challenge of determining the content or interpretation of a contract term in circumstances different from the ones of the moment when the contract was entered. See Hillman (1987).
as written\textsuperscript{150} by a court or arbiter.\textsuperscript{151} This does not mean that no contractual clause nor contractual behavior by the parties is able to be legally enforced. Some instances of indisputable breach of contract – departures from the cooperative equilibrium, or lack of performance – can be detected by the contractual partner, and verified in front of a court, and thus can be deterred by the use of legal remedies, such as specific performance and damages.

However, there is always the possibility of a wide variety of cases of non-cooperative behavior within the contractual relationship that pose insurmountable measurement problems, particularly when those behaviors are multi-dimensional. The chances that non-complying contractual behaviors can be shown, in a sufficiently convincing manner, to a court or other external adjudicator, is very low. Further, the amount and scope of unverifiable breaches of contract would tend to increase not only with the chances that the breaching party would escape undetected or not be subject to legal contractual remedies, but also with the chances that consumers would not detect or punish the defecting distributor.

Parties to a long-term contract can resort to several alternatives that can serve as – imperfect – substitutes for perfect court enforcement of contractual remedies against verifiable breach of contract.\textsuperscript{152} First, the contract can use, instead of a non-verifiable dimension, some verifiable proxy for the desired contractual behavior, and thus make use of legal enforcement of this proxy, given that courts could use legal remedies for its breach. Second, the contract can contain clauses that tend to decrease the benefits, and to increase the costs, of behavior of the other party deviating from the cooperative pattern. In other words, they can try to introduce clauses in the contract that serve as mechanisms facilitating its self-enforcing character.

Of particular importance among the self-enforcing mechanisms that are an alternative to legally formal enforcement, and specially relevant for the legal treatment of termination and compensation after termination, is

\textsuperscript{150} On many occasions nothing is written, and the contractual intention has to be inferred from the behavior of the parties, prior or posterior to the initiation of the relationship.

\textsuperscript{151} This is the reason why the quest, in this area of relational contracts, of the economic literature has been how to design self-enforcing relational contracts, that is, contracts in which the parties are induced to adopt the best available actions for the common good, based on their own strategies, but checked by reciprocity, reputation, or other intrinsic motivators. This is, as well, the reason why economists often view with mixed feelings the function of the law in this sort of setting.

the use of future quasi-rents for the distributor linked to the continuation of the relationship:153 if the terms of the contract are adjusted so that one of the contracting parties expects to earn quasi-rents on its investments, if and only if the existing relationship goes on, that party has a powerful motive to remain into the contract, and thus to avoid the kind of negative behavior that can trigger the end of the relationship.

It should be noted, though, that in order to make this instrument of relation-specific quasi-rents work as an incentive mechanism to achieve cooperation in dimensions outside what can be verified by a court, preserving the effectiveness of the threat of termination is crucial, a point that has been underlined by several commentators.154

7.2. Specific Investments and Breach Remedies
As mentioned earlier, specific investments are relationship-related investments so that their value depends on whether parties are within or outside the contract terms. The nature of such investments, as summarized above, poses important challenges for optimal investment decisions. The presence of specific investments also affects decisions to perform or breach, and to terminate or to go on with the relationship. It is easy to observe that the holdup problem makes the party with higher specific investments more vulnerable in all dimensions of the relationship.155 Therefore, whenever parties make investments that are mostly relation specific, the costs of ending the contract are high for the party making the investment and, therefore, this party will have a much lower incentive to terminate because this will be very costly. In contrast, if investments are not relation specific and therefore may be used outside the relationship, the incentives to terminate the contract will increase.

The presence of specific investments also affects decisively the working of the standard legal remedies against breach of contract, and thus, given the pervasive presence of such investments in long-term contracts, the effects of breach remedies in these may be quite different from what happens in spot contracts.

For the analysis on parties’ incentives to perform and to invest, a further distinction within specific investments is necessary. The economic

153 The pioneering analyses along this line are by Rubin (1978); Klein (1980); and Klein and Leffler (1981).
154 Klein (1995) underlines the importance of termination-at-will for the effectiveness of self-enforcing mechanisms, and how legal constraints – mandatory severance payments or compensation, or good cause requirements – severely limit this option. See also, Paz-Ares (1997, 2003).
theory literature dealing with incomplete contracts has distinguished two pure types of such investments. Selfish investments are investments that benefit the party making the investment, and not the other party: if the buyer makes the investment, it just increases the value of performance for the buyer, without decreasing production costs for the seller; if the seller makes the investment, it just decreases its production costs, without increasing the value of performance for the buyer.

Cooperative investments\textsuperscript{156} are investments that confer direct benefits on the other contracting party, and not on the party making the investment. If the buyer makes the investment, it just decreases the production costs for the seller, without increasing the value of performance for the buyer; if the seller makes it, it just increases the value of performance for the buyer, without decreasing production costs for the seller. There are also hybrid investments that benefit both contracting parties, the investor and his partner, although we will disregard this complication in what follows.\textsuperscript{157}

For selfish investments, the solutions explored in the economic literature have evolved in two directions. First, to design mechanisms in an incomplete contract that can achieve an efficient outcome, both in terms of trade and in terms of specific investment. Most contributions explore ingenious procedures in the bargaining conditions in the renegotiation phase of the contract, so that the party making the investment receives the full value of the investment. Some opt to place external conditions on the renegotiation phase,\textsuperscript{158} others for the use of options and appropriate strike prices for their exercise.\textsuperscript{159} Other approaches rely on contracts determining intermediate quantities to trade, so after renegotiation it turns out that the investing party receives in some cases more and in some cases less than the marginal social value of the investment, and if the quantity is adequately chosen in the contract, both effects may cancel out and in the end induce efficient levels of investment. These approaches, however, have only limited applicability to the design and operation of legal rules in the contract setting.\textsuperscript{160}

\textsuperscript{156} The first treatment of cooperative investments is that of MacLeod and Malcomson (1993). For the standard treatment of these investments, see Che and Chung (1999); Che and Hausch (1999). For a very interesting recent contribution on contract remedies and specific investments, showing that the dismal result – concerning expectation damages – of Che and Chung (1999) is but an extreme case, see Stremitzer (2008).

\textsuperscript{157} In fact, few papers explore hybrid investments: Che and Hausch (1999); Segal and Whinston (2000b); Göller and Stremitzer (2009).

\textsuperscript{158} Chung (1991).

\textsuperscript{159} Nöldeke and Schmidt (1995).

\textsuperscript{160} Edlin and Reichelstein (1996).
The second strand of the literature on selfish investments specifically deals with legal remedies against breach, and with their impact on the investment decision by the parties (also on the decision to perform or to breach the contract, but let us leave that aside).

The pioneering contribution here is by Shavell.\(^{161}\) He addresses two scenarios. The first is the one in which the investing party is the party that may be the victim of breach. He then shows that expectation damages induce excessive specific investment by the potential victim of breach. The reason for this effect of over-reliance lies in the fact that expectation damages fully insure the investing party against the possibility of losing the return on the investment, more than is optimal from the point of view of the joint welfare of the parties: even when there should be (and there is) no trade, that is, when the contract should not be performed, the investing party gets the full return from the investment. Reliance damages perform even worse than expectation damages, that is, they induce even more over-investment. The reason is that to the full insurance motive to over-rely (reliance damages fully insure the investing party because in all possible future states of the world, he obtains at least restitution of the cost of the investment) now has to be added a performance inducement function: by investing more in specific assets, the party directly increases the damage award the other party has to pay in case of breach, thus increasing the incentives of the latter to perform. These results have been to an important degree confirmed by experimental tests of contracting behavior in a controlled laboratory setting.\(^{162}\)

The second scenario appears when the investing party is the one who can take the decision to breach or to perform the contract. In this case, expectation damages induce efficient investment: the breaching party is the residual claimant of the value of the investment (the reduction in cost of production, for instance), because the damage award he has to pay (the value of performance to the other party) does not depend on the level of investment. Reliance damages also perform worse than expectation damages, although in a different direction than in the first scenario. Given that reliance damages generally induce too little performance with respect to the efficient level, the investing party will get a return on the specific investment that is less than optimal, and therefore the incentives to invest would be too low.\(^{163}\)

Shavell’s analysis in the first scenario (the investing party is not the one

\(^{161}\) Shavell (1980).
\(^{162}\) Sloof et al. (2003).
\(^{163}\) Shavell (1980).
making the breach–perform choice) was extended along two different lines. First, that post-breach renegotiation by the parties does not alter the inefficient investment incentive under both expectation and reliance damages, and also the ranking of the two remedies: reliance damages are less attractive than expectation damages in order to provide less inefficient investment incentives.\footnote{Rogerson (1984).} Second, that with an appropriate instrument, one can transform the first scenario into the second one, that is, that the investing party is the one that can take the decision to breach or to perform. The instrument is a large up-front payment from the buyer to the seller (assuming the seller is the party who can invest), which ensures that the buyer would never want to breach – he would get performance for a price close to zero, because the up-front payment is sunk when the performance decision arises. Then, any breach would come eventually from the investing party, who has efficient breach and investment incentives under the expectation damages remedy.\footnote{Edlin (1996).} It is true, though, that using up-front payments can have problems of their own (basically liquidity problems on the part of the prospective buyers, or the non-investing party more generally), but it clearly shows how the investment problem can be solved, for selfish reliance expenditures, by concentrating the decision to breach, and the decision to invest, in one and the same party.

The question is more complex still for cooperative specific investment. It can be shown that if the parties cannot commit not to renegotiate the contract, there is no incomplete contract, however complex, that can induce efficient incentives.\footnote{Che and Hausch (1999) and Maskin and Moore (1999).} In fact, the dismal result is that a contract is no better for the parties than no contract at all.\footnote{Stremitzer (2008) shows that if the contract can be conditioned (that is, courts are able to verify) on whether performance is above or below a given threshold – of quality, as paramount example – under certain conditions first best can be attained.}

For these cooperative investments, the role of legal remedies against breach has also been explored, both absent renegotiation, and when ex post renegotiation is feasible, usually considering that the investing party is not the one that can take the breach–perform decision.\footnote{Che and Chung (1999).} In the first case, with no renegotiation, expectation damages perform very poorly, because as a remedy for breach, they induce zero cooperative investment. Reliance damages, in turn, do much better according to Che and Chung: although at the price of some distortion – in the direction of

\begin{footnotes}
\footnote{Rogerson (1984).}{\footnote{Edlin (1996).}{\footnote{Che and Hausch (1999) and Maskin and Moore (1999).}{\footnote{Stremitzer (2008) shows that if the contract can be conditioned (that is, courts are able to verify) on whether performance is above or below a given threshold – of quality, as paramount example – under certain conditions first best can be attained.}{\footnote{Che and Chung (1999).}}}\\
\end{footnotes}
excessive breach – in the breach–perform decision, they provide much better incentives for specific investments, and overall improve contractual surplus over expectation damages. With efficient renegotiation ex post, expectation damages continue to perform as poorly as before, but reliance damages can now achieve efficient incentives, both to perform and to incur cooperative specific expenditures.

There is also a defense of expectation damages in this setting, albeit not ordinary expectation damages, but bilateral expectation damages. This implies that the party who can breach can be subject to paying expectation damages to the other party, the investor; but the latter can also be liable in front of the former if the level of investment falls short of the level determined in the contract. If this is the case, bilateral expectation damages do also induce efficient trade and efficient cooperative investments. For this result to hold, it is necessary not only that the actual level of investment can be verified before the court (a condition for reliance damages to operate as remedy against breach, to be sure), but also that the parties can fix in the contract the efficient level of cooperative investment, which is a much more implausible – though not impossible – assumption.

Finally, in a recent paper, it has been argued that the extremely poor – zero cooperative investment – efficiency performance of expectation damages (the preferred remedy for breach of contract in common law, and also heavily used in civil law jurisdictions in many contexts) is due to an implicit assumption that the contract does not contain any threshold of performance, and that the court cannot imply one either. If, on the contrary, the contracting parties or the courts are able to compare performance with a verifiable legally binding threshold (over the relevant dimension of performance), expectation damages generally induce positive – albeit suboptimal – levels of cooperative investments, and under certain conditions (those of the so-called maximum quality or Cadillac contracts), they can even provide incentives for efficient investments of a cooperative nature. This chapter also favors, when renegotiation is possible, the use of an optional remedial regime for the non-investing party, consisting of a choice between specific performance and termination – with restitution of payments made, if any – if performance of the investing party falls

---

172 It is well known that absent renegotiation, specific performance is likely to produce inefficient trade and thus undesirable results: Shavell (2004).
173 The use of optional regimes has not often been considered in the literature, with the exception of Avraham and Liu (2006, 2008).
below the legally enforceable threshold. Above the threshold, only specific performance would be available for either party.

7.3. Covenants Not to Compete

Long-term contracts often have effects even after the contract has ceased to be in force and the relationship has ended. Covenants not to compete are frequently observed as contract clauses in long-term contracts, and at face value they serve to control the post-contract behavior of one or both parties. From a theoretical perspective, they are typically instruments that are related to contract investments, especially investments in training, and know-how transfers that one party to the contract makes and the other party enjoys. In this sense, covenants not to compete may be regarded as an enforcement mechanism in the implicit agreement between parties to pay back investments in general training and know-how.

However, the differences highlighted earlier regarding the different kind of investments – general or specific – are still present. So if the investment in human capital is perfectly specific, the employee could not take with him the increased human capital and increase his productivity in another company. In addition, there could be a double holdup problem depending on who funds the costs of the training.

However, regardless who makes the investment, there is always one party – the non-investor – who would credibly threaten to breach the implicit promise sustaining the investment in specific training.

As explained earlier, the possibility of renegotiation also matters for the outcome when specific and general training are at issue. For example, when dealing with specific investments in specific training, both parties have a clear incentive to over-invest. As long as the employee’s training

---

174 In this sense, they could be qualified as cooperative investments. See Nölife and Schmidt (1995). The first analysis of cooperative investments is by MacLeod and Malcomson (1993). For standard analysis of these investments, see Che and Chung (1999); Che and Haush (1999). For a very interesting recent contribution on contract remedies and specific investments, showing that the result regarding expectation damages – of Che and Chung (1999), is an extreme case, see Stremitzer (2008).

175 And may be a better remedy than a liquidated damages clause, due to problems of limited assets or personal bankruptcy of the employee. See Rubin and Shedd (1981).

176 Labor economics literature characterizes this problem as the trade-off between salary and training. The employee accepts lower wages – but not so low as to reflect the true cost of training – for a while, and then the employer, once the enhanced productivity is in place, pays a salary above the opportunity cost of the employee, but below the full value of the trained employee. See Lazear (1998).
value is more valuable to a different third party, parties in the original contract have an incentive to over-invest, especially considering the contract could be renegotiated ex post, in the case of a contractual bid by a third party, because a high level of investment increases the value from the third party to obtain release of the employee, because the training is not worth as much when the employee moves to another partner.\textsuperscript{177} The higher the specific investments that the potential breaching party makes, the greater are the switching costs to the new contracting party, because the more valuable is the existing relationship for the parties – and not the outside option, because the investment creating the extra value is specific. The third party needs to compensate the terminating party for any switching costs, in order to induce him to terminate.\textsuperscript{178} So, specific investments, even without renegotiation, tend to be excessive when decided by the party who can also decide on termination.\textsuperscript{179}

Covenants not to compete, however, may be socially preferable to the other options the parties may use to benefit at the expense of the entrant, in the sense that the scope of the covenant may be adjusted (by the court, for instance, ex post), simply to cover those outside alternatives for the employee who really benefits from the general training, and excluding from the enforceable scope other outside options for which such training is worthless. Reducing the scope of the covenant not to compete to just the industry may be efficient, because it ensures enough incentives for investing in training, but eliminates the excessive incentives to invest that respond to the purpose of extracting rent even from those who value the employee but not the training.\textsuperscript{180}

8. Empirical Literature
Long-term contracts have provided a fertile environment for empirically testing hypotheses on contracting behavior. The empirical bent of the transaction cost economics literature has provided an added impulse to empirical studies on long-term contracts and the incentives they generate

\begin{itemize}
\item \textsuperscript{177} Posner and Triantis (2001).
\item \textsuperscript{178} See Chung (1998).
\item \textsuperscript{179} This may be an additional reason to make the employer pay for the investment in training when it is the employee who can terminate. If it is the employer (or, more realistically, the manufacturer in a distribution contract) who is more likely to terminate, the opposite result would be desirable, however.
\item \textsuperscript{180} See Posner and Triantis (2001). Courts or external adjudicators cannot, however, restrict their attention to covenants not to compete, and disregard the other alternatives (liquidated damages or penalties, for instance), because the parties would resort to them if covenants not to compete are controlled to avoid the effect of extracting value from the third parties.
\end{itemize}
to make certain choices. Take the decision to vertically integrate as one example. Lafontaine and Slade\(^\text{181}\) suggest that when transaction costs are important, firms will choose governance structures to reduce the likelihood and cost of haggling and exploitation of the other firms. It should be noted that the importance of the different transactions costs – downstream or upstream – does not equally affect the decision to vertically integrate. So as the importance of local or downstream effort grows, integration becomes less likely, whereas as the importance of company-wide or upstream effort grows, integration becomes more likely.\(^\text{182}\) Regarding the effect of higher monitoring costs on vertical integration, the empirical literature is not unanimous.\(^\text{183}\) Lafontaine and Slade note the differences between transaction cost theories and property right theories regarding the incentives of firms to integrate vertically. Transaction costs theories, developed by Williamson\(^\text{184}\) and Klein, Crawford, and Alchian,\(^\text{185}\) noted that when the problems associated with transaction costs are important, governance structure will seek to minimize the likelihood and cost of negotiation and exploitation.\(^\text{186}\) Transaction cost theories predict that vertical integration will be more likely when transaction costs are complex and involve specific investments such as durable specific assets, unverifiable quality of those assets, uncertain environment or when the quasi-rents are generated by a relationship. Property right theories, on the other side, were developed by Grossman and Hart,\(^\text{187}\) Hart and Moore\(^\text{188}\) and Hart,\(^\text{189}\) and focused on the relationship between specific assets, incomplete contracts and ex post bargaining. Property rights literature predicts that vertical integration can result in a reduction of incentives to make investments.\(^\text{190}\)

But in light of the abundance of the literature regarding transaction costs and vertical integration,\(^\text{191}\) and given that some of it is essentially driven by

\(^{181}\) Holmstrom and Roberts (1998); Gibbons (2005); Whinston (2003); and Lafontaine and Slade (2007).

\(^{182}\) Holmstrom and Roberts (1998); Gibbons (2005); Whinston (2003); and Lafontaine and Slade (2007).

\(^{183}\) Holmstrom and Roberts (1998); Gibbons (2005); Whinston (2003); and Lafontaine and Slade (2007).


\(^{185}\) Klein et al. (1978).

\(^{186}\) Lafontaine and Slade (2007).


\(^{188}\) Hart and Moore (1990).

\(^{189}\) Hart (1995).

\(^{190}\) Lafontaine and Slade (2007).

\(^{191}\) Holmstrom and Roberts (1998); Gibbons (2005); Whinston (2003); and Lafontaine and Slade (2007).
Long-term contracts in the law and economics literature

concerns about organizational theory, and not by an interest in the legal or regulatory environment of long-term contracts, we will focus only on the legal restrictions on termination of long-term distribution contracts, which has become a highly contested legal issue in many jurisdictions (USA, Spain, and others) and is one of the key issues in the harmonization exercise of European private law dealing with long-term contracts (article IV. E.-2:301 and following of the Draft Common Frame of Reference).192

The empirical evidence concerning the effects on the behavior of contracting parties in a long-term contract of the legal rules that restrict or impose legal conditions on terminating the contract is rich and ample.193 This evidence refers essentially to franchising,194 but there does not seem to be a powerful reason to doubt that its main findings would not be applicable to other contractual arrangements in distribution chains that share issues of controlling opportunism by distributors (and, eventually, also by manufacturers).

The first and best-known piece of empirical evidence concerning termination of long-term distribution contracts is Brickley et al.195 They hypothesized that laws restricting franchisor termination rights would lead to less franchising because this would lead to less profitable franchising, making other arrangements (such as franchisors running the units directly) more profitable by comparison.

In turn, Beales III and Muris196 looked at whether data on franchise terminations and non-renewals support the efficiency or the opportunistic explanation for terminations.197 Their results neither support nor present cause to reject the opportunism hypothesis: the estimated coefficients are often of the wrong sign or statistically insignificant. However, they did obtain a robust, significant, and negative coefficient on the ‘growth in outlets’ variable. This suggests that, if opportunism or expropriation by the franchisor is a factor, its effect is diluted by the franchisor’s interest

192 See, critically, Gomez (2009).
193 See Brickley et al. (1991); Beales III and Muris (1995); Williams (1996); Lafontaine and Shaw (2005); Brickley et al. (2006); Klick et al. (2007).
194 The reason for this lies in the fact that the studies are based on US experience, where state legislation interfering with termination at will has concentrated on franchise contracts. Moreover, it seems that franchise plays a somewhat larger role in US distribution compared with the European context.
196 See Beales III and Muris (1995).
197 Efficient termination would be one in which the franchisor detects a breach of quality provision duties by a franchisee while opportunistic termination is defined as any non-efficient termination. See Beales III and Muris (1995).
in maintaining its reputation in order to attract additional quality franchisees.

Williams also examined termination rates of franchise contracts, in a sample of over 1,000 contracts over a four-year period, and found no evidence of termination being influenced by a franchisor appropriating for himself those units that, whether through franchisees’ sales effort or for other reasons, turned out to be particularly profitable.\(^{198}\)

Klick et al.\(^{199}\) also used data on franchising chains to assess the relative importance for termination of the disciplining and expropriation stories. They examined state laws limiting franchisor termination rights to identify the effect of termination at will on both the decision to franchise and on franchisor expansion generally. Their results tend to support the view that the disciplining effect of termination on a franchisee’s non-cooperative behavior seems to outweigh opportunities for franchisor abuse and expropriation of value that termination at will may allow.

Lafontaine and Shaw\(^{200}\) have investigated whether data sustain the proposition that franchisor opportunism is an important factor behind the rate of termination, and found no result consistent with that prediction.

Brickley et al.\(^{201}\) sought to assess the ‘exploitation’ theory of franchising, concentrating on clauses regulating contract duration that are typically crucial for the chances that franchisees recover relation-specific investments made in contemplation of the contract being in place for some period of time. Specific investments make the franchisee vulnerable, because the termination of the contract will not allow the franchisee to recover the specific, and thus non-salvageable, investment. The longer the contract term, the higher are the chances of complete recovery of investment by the franchisee.

Using a large sample of franchising firms, Brickley et al. analyzed the effects on contract duration clauses of several factors.\(^{202}\) If the exploited franchisee view were correct, we would expect that the larger and more

\(^{198}\) In fact, the main factors driving termination rates appeared to be a desire to transfer the unit (frequently, by the franchisee herself) and to close units underperforming due to poor franchisee performance or a disadvantageous location. Williams (1996).

\(^{199}\) See Klick et al. (2007).

\(^{200}\) See Lafontaine and Shaw (2005).

\(^{201}\) See Brickley et al. (2006).

\(^{202}\) Among those factors were the number of years the franchisor has been in operation; the number of sites the franchising network comprises (that is, the franchisor’s size); the average total initial investment of a franchisee entering the franchise network; the number of weeks of off-site training of a franchisee’s personnel. See Brickley et al. (2006).
sophisticated the franchisor, the more exploitative the contract terms, and
the shorter the contract duration, will be. Again, if the naïve franchisee
image were correct, the level of specific investments would not raise con-
tract duration, given that exploitative franchisors would try to appropri-
ate the value of the non-amortized specific investments incurred by the
franchisee.

Empirical results show that the four factors are positively and sig-
ificantly correlated with the length of the contract term: both the level of
the investments by the franchisee, and the size and the experience of the
franchisor tend to increase contract duration,203 contrary to the predic-
tion of the ‘exploitation’ hypothesis.204 There is thus evidence to indicate
that franchisors are responsive to the level of specific investments by
franchisees, and are more responsive as they become bigger and better
established. Such results furthermore provide indirect evidence that the
threat posed by opportunistic and exploitative behavior on the part of
franchisors is not in reality a particularly worrisome problem205 or, at
least, is sufficiently marginal so as not to show up in the data.

A final important issue discussed in the empirical literature on franchise
contracts is the relationship between the legislation restricting termination
at will and the number of terminations. Contrary to what one would intui-
tively expect, legislation restricting termination at will increases, rather
than decreases, the number of terminations.

The explanation for this finding206 advanced by some commentators
is that unconstrained termination at will induces franchisors to be more
forgiving of minor (even if verifiable) instances of breach by the franchisee.
Being forgiving at the beginning is not too costly for a franchisor, given
that she always retains the ability to terminate without any restriction
as soon as she observes that her benevolence has not been repaid with
cooperative behavior by the franchisee. On the contrary, if the deci-
sion to terminate is legally constrained, the franchisor will terminate on
the first occasion she can so that the franchisor (or the principal, more

203 See Blair and Lafontaine (2005), who also find that larger franchisors tend
to offer longer contracts, on average, than smaller ones.
204 These results hold irrespective of the fixed effects of the particular industry
in which the franchisor operates. Brickley et al. (2006).
205 It is true, however, that Brickley and his co-authors also find a positive
effect of legal restrictions on franchise termination (in the state where the fran-
chisor has its headquarters) on contract duration clauses: Brickley et al. (2006).
They hypothesize that this effect is due to the increased bargaining power such
legislation gives franchisees upon termination of the contract, thus reducing the
value of short-term contracts for the franchisor.
generally) will not be inclined to act forgivingly in front of a first minor breach if there is sufficient evidence that termination would be deemed an acceptable punishment of a franchisee’s breach.

9. Conclusion
In light of the many different issues raised by long-term contracts, and the multiplicity of approaches and results in the literature, drawing general conclusions is virtually unfeasible. Long-term contracts present specific issues such as – added – incompleteness, specific investments, and the important difficulties arising from the difficulties in observing and verifying non-cooperative behavior. Short-term contracts are not exempt from these issues, but in some cases, they may appear as a useful alternative to contracting parties. It does not seem possible to establish from a general perspective whether parties should enter into short-term or long-term contracts. Which type of contract will best fit the parties’ needs will depend on their goals, position, information available and the time frame of their relationship.

From a general perspective, contracts require cooperation from contracting parties, and usually this is not self-enforcing. However, in light of the open-ended nature of long-term contracts, ongoing cooperation between parties is of prime relevance. Long-term contracts involve significant risks for parties, given that they involve a lower degree of certainty, and they may raise added problems concerning specific investments, given the chances for renegotiation. At the same time, they may yield higher levels of commitment and cooperation, so that compliance with their contractual obligations actually improves, and without the need to rely on formal remedies for breach.

Bibliography


Poitevin, Michel (1992), ‘A Note on “Contracts as a Barrier to Entry”’, Mimeo, University of Montreal.


