

# Compensation after Termination of Long-Term Distribution Contracts: An Economic Perspective of EU Law

Fernando Gómez Pomar

Universitat Pompeu Fabra

390

## *Abstract\**

*Termination and compensation after termination are two of the most relevant –albeit not the only ones– dimensions of the long-term contracts through which distribution chains are formed and structured. The EC rules that establish a regime of compensation after termination are studied in this paper, in particular, the Commercial Agents Directive. Then, compensation after termination is analysed from the economic perspective, highlighting the importance of the open-ended nature of a relationship, termination as a disciplining mechanism against non-verifiable breach of distribution contracts, the problem of specific investments and the covenants not to compete. The contrast of the examined EC rules with the theoretical findings of the Law and Economics literature dealing with the matter does not offer a promising view of their likely consequences for the contractual behaviour of the parties.*

## *Summary*

### **1. Introduction**

- 1.1. The problematic nature of the matter**
- 1.2. Crucial issues for understanding the question**

### **2. EC Law basis**

- 2.1. Commercial Agents Directive**
- 2.2. Consumer Credit Directive**
- 2.3. Package Travel Directive**
- 2.4. Consumer Sales Directive**

### **3. The Economic considerations behind compensation after termination**

- 3.1. The importance of the open-ended nature of the relationship**
- 3.2. Incomplete contracts and termination as a disciplining mechanism against non-verifiable breach of distribution contracts**
  - 3.2.1. Incompleteness in long-term distribution contracts and the role of termination**
  - 3.2.2. Empirical evidence on the effects of legal restrictions on termination**
- 3.3. The relevance of specific investments**
  - 3.3.1. The hold-up problem**
  - 3.3.2. Selfish vs. cooperative specific investments**
  - 3.3.3. Are the rules of the Commercial Agents Directive sound in economic terms?**
- 3.4. Covenants not to compete**
  - 3.4.1. Different forms of training and the role of covenants not to compete**
  - 3.4.2. The role of renegotiation in covenants not to compete**
  - 3.4.3. Empirical evidence: Silicon Valley vs. Route 128**
  - 3.4.4. Covenants not to compete and other clauses as barriers to entry**

### **4. Conclusions**

### **5. References**

### **6. Decisions**

---

\* I am grateful to Stefan Grundmann, Karl Riesenhuber, and Jules Stuyck for very detailed and acute comments on an earlier version of the paper. The usual disclaimers apply.

## 1. Introduction

### 1.1. The problematic nature of the matter

A substantial part, if not most, goods and services in a modern economy reach consumers through distribution chains with a varying number of links, and a diverse formal legal shape of the contracts within the chain (franchise, selective distribution, etc.). One of the most relevant questions affecting the contracts building those distribution chains is precisely the issue of compensation after termination of one of these long-term contractual relationships. In Spain, this matter has been one of the most debated topics in the entire field of the Law of distribution in recent years<sup>1</sup>.

When I use the term compensation after termination, I take it broadly, as to be equivalent to any kind of financial payment or other sort of monetary remedy between the parties conditioning, resulting, or otherwise related to the termination of the contract. Even more than any other question on long-term distribution contracts, this question cannot be addressed in isolation. Not only for the obvious reason that termination provides the starting point or the background against which it has to be assessed the existence and the level of the compensation. The influence between the issue of compensation and termination as such, or other items, such as adjustment of the terms of the contract, is reciprocal. Termination cannot be properly understood irrespective of compensation. The imposition of the duty to compensate fixed at a certain -high- level can deter, totally or partially, one party from terminating the contract. Similarly, the absence of any duty to compensate, or compensation in an insufficient amount, can promote or induce termination in a wide variety of circumstances, not all of them desirable from the perspective of the law-maker, or from a social welfare perspective.

Typically, compensation –or lack thereof- payable by one contracting party to the other is one of the key remedies against breach of contractual obligations, and provides one of the core instruments to create incentives for the decision to breach or not to breach those duties, as well as to regulate a variety of contractual behaviours and dimensions (specific investments post-contract, allocation of the risk of losses or future contingencies, and so on). Breach of contract appears thus, almost inextricably with the issue of compensation. Even in a broader sense, the legal decisions concerning compensation after termination are some of the most relevant mechanisms affecting many –if not most- dimensions and elements of the contract, and so their function, role, and consequences cannot be understood as a separate matter from the rest of the rules in place.

In fact, I do not take the *ex post* effects of the compensation mechanisms, if any, present after termination of long-term distribution contracts to be the most relevant issue to be analysed theoretically, nor the most important consequences in economic or practical terms. Bluntly said, money transfers among businesses are of little importance in themselves from the point of view of governing

---

<sup>1</sup> The most controversial issue in Spanish Law has been the use (or the rejection) of analogical reasoning based upon the Spanish *Ley de contrato de agencia* (implementing the European Directive on commercial agents) for contracts outside the specific scope of the *Ley de contrato de agencia* 12/1992 (BOE no. 129, May 29). This problem has substantially misled Courts and scholars as to the adequate treatment of termination and ensuing compensation in franchising and exclusive distribution contracts. See, for opposite views, PAZ-ARES, 1997, and FERNANDO MARTÍNEZ *et. al.*, 2000, p. 464. For instance, in two very recent decisions by the Spanish Supreme Court, analogical application of the compensation provisions in the *Ley de contrato de agencia* has been denied in one case, and affirmed in a very similar one: See, STS, 1ª, 27.10.2005 (Ar. 8043, MP: José Almagro Nosete), denying the

contractual relationships. It is the implications of those payments for the previous behaviour of the contracting parties that make the interesting and difficult –both theoretically and practically– part. Therefore, in what follows I will assume that the contracting parties behave prospectively, and that they know the rules dealing with compensation after termination, and that those rules contribute to shape the incentives to undertake all sorts of contractual behaviour: Drafting contract terms, making investments that increase –for the investing party or for the other contractual partner– the value of performance if it occurs, complying with, or breaching contractual obligations and duties. In all these, as well, I will assume, unless otherwise expressly relaxed, that the contracting parties are firms<sup>2</sup>, and that they are risk-neutral, that is, that they care only about the expected value of alternative courses of behavior, and not about the intrinsic riskiness of each of them<sup>3</sup>. This assumption of risk neutrality obviously does not correspond with many observations of risk-averse behavior in firms, but I think it is tenable enough as the general framework, given that we are dealing here with general rules governing contracts, and not various particular contractual environments or economic sectors. I will, nevertheless, consider some consequences of risk-aversion in section 3.3.3 to fully understand the effects of a kind of compensations schemes for termination of a long-term distribution contract.

## 1.2. Crucial issues for understanding compensation after termination

In the previous subsection I have already touched upon the complexity of dealing with the matter covered by question 9, namely the numerous and varied set of theoretical dimensions that touch upon it. As was already mentioned, what I take to be the core of the matter is how compensation after termination may affect relevant conducts of the parties to a long-term distribution contract. What are the dimensions that touch upon those incentive effects more closely? To anticipate, in a somewhat cursory fashion, what will be developed in more detail in sections 3 and 4, a brief list could be as follows:

First, that the open-ended nature of a relationship –that is, that termination can occur– is not indifferent for the incentives to cooperate with the contracting partner in achieving full exploitation of the contractual opportunities. Second, that termination can be a powerful tool to discipline subtle and not easily verifiable before a Court or other external adjudicator, instances of breach of contractual duties by distributors. This disciplining or policing aptitude, however, can also be used by the other party to expropriate the distributor of the value resulting from relation-specific investments. Third, that the protection of specific investments and reliance on implicit promises against the risk of imposed renegotiation or expropriation constitutes a possible rationale for imposing compensation, under some conditions. Fourth, that the use in long-term contracts of some clauses, such as exclusivity, and covenants not to compete, are not indifferent to entry by potential challengers, and thus may produce undesirable foreclosure in a market. All those issues just enumerated provide, in my view, the necessary frame of reference to understand legal rules affecting compensation following termination in

---

application, and STS, 1<sup>a</sup>, 21.11.2005 (Ar. 7677, MP: Antonio Gullón Ballesteros) affirming it.

<sup>2</sup> This assumption will probably make my previous one of prospective and –at least legally– informed behavior of the contracting parties more palatable than if I was dealing with consumers.

<sup>3</sup> This assumption refers only to the parties themselves, and it does not imply that in choosing among different options about how to regulate contract behavior, the law-maker may prefer the rule creating less uncertainty in its interpretation by Courts, for instance.

a distribution chain, and eventually, to design some broad legal principles that can serve as building blocks for a more detailed and comprehensive legal regime.

## 2. *EC Law basis*

It may seem strange to start a section dealing with EC Law rules touching upon the matter with a disclaimer of suspected relevance or, at least, suspected generality of the existing EC Law rules on compensation following termination. I think it is worth mentioning in the preliminaries of the summary of EC Law in the matter, the haphazard and problematic basis in the *Acquis communautaire* of the compensation for termination issue, at least from a general perspective.

I am not referring to scarcity or lack of importance of the existing rules as such, which in fact is not the case. There are EC rules, and the rules establish a recognizable regime of compensation after termination in their own scope of application. Artt. 17-19 of Directive 86/653/EEC on self-employed commercial agents address specifically, purposefully, and in detail monetary payments to the agent following termination of the agency relationship. Art. 20 of the same Directive regulates covenants not to compete imposed upon the agent after termination. In turn, And art. 3 of Commission Regulation 1400/2002, on the application of Article 81 (3) of the Treaty to categories of vertical agreements and concerted practices in the motor vehicle sector refer laterally to the issue of compensation after termination in connection with the duration of notice periods for termination of agreements of indefinite term. Art. 5 (b) of Commission Regulation 2790/1999 on the application of Article 81 (3) of the Treaty to categories of vertical agreements and concerted practices (block exemption regulation) also deals with non-compete obligations after the termination of the agreement, setting specific requirements for those covenants not to compete in order for the entire agreement to benefit of the block exemption.

All these may not seem, at first blush, not so scant or inadequate in terms of an EC Law set of rules on which to build an analysis of compensation after termination and related issues. But there are powerful reasons to the contrary, I believe.

On the one hand, the commercial agents Directive can be considered, to some extent, tainted as a general basis (not as a benchmark, however!) for a contractual regime of compensation after termination in long-term contracts. The reason lies primarily in its self-declared objective of protecting commercial agents vis-à-vis their principals, a goal of social protection which has been affirmed by the ECJ in the *Bellone*<sup>4</sup>, *Ingmar*<sup>5</sup>, and *Honyvem*<sup>6</sup> cases, and which, to a non insignificant extent, approximates some of the underlying principles in the Directive to those of employment protection and labour-related legislation.

There seems to be, at least in my view, a fundamental problem in grounding a contractual set of rules for firm behaviour and B2B firm contracting on a mandatory legal regime trying to improve the lot of one of the parties as a question of social policy. If one –as I would- thinks that general, that is, not

---

<sup>4</sup> Case C-215/97, *Barbara Bellone v. Yokohama SpA* [1998] ECR I-2191.

<sup>5</sup> Case C-381/98, *Ingmar GB Ltd v. Eaton Leonard Technologies Inc* [2000] ECR I-09305.

<sup>6</sup> Case C-465/04, *Honyvem Informazioni Commerciali Srl v. Mariella De Zotti* [2006].

circumstance-specific, legal rules to determine and guide contractual behaviour in a variety of economic sectors and social circumstances, ideally should *prima facie* promote, absent external effects towards third parties, the maximization of the contractual surplus available to the parties, a set of rules that purposefully departs from this approach to pursue social protection of one of the parties in a contractual interaction does not seem to be an appealing or promising starting point for attempting a process of generalization.

The preceding remark does not preclude the conceptual solutions underlying the rules in the Commercial Agents Directive –disorgement of benefits rendered to the terminating party by the contractual partner; compensation for some types of specific investments- to be worth exploring in the context of long-term distribution contracts among firms, even acknowledging the differences between both sets of factual and economic circumstances.

As for the block exemption (both general and for the car distribution sector) rules, they do not appear either as a sound basis for developing an analysis whose ultimate purpose is to structure an optimal set of rules in Contract Law. Those rules dealing with the antitrust aspects of certain practices and clauses in vertical relationships provide an inescapable regulatory environment (see Jules paper) for the behaviour of suppliers and distributors in the European market, but they are not, and are not meant to be, a set of optimal rules regarding contractual behaviour.

The reason for the scepticism lies not so much in the possible difference in goals separating Competition Law from Contract Law –as I hold the promotion of social welfare in the economic sense to provide the basic objective to both of them-, but for the reason that the rules contained in the block exemption regulations are not easily generalizable as abstract and general rules of contract Law. EC competition Law rules in this area of vertical agreements are increasingly context-specific (as optimal regulatory policies and rules should in principle be), and not clear-cut, universal mandates or prohibitions. They require *ad hoc* estimates of the particular effects of a given practice, clause or behaviour on the actual levels of competition in the affected markets, a task that requires a great deal of empirical assessment carried on in the context of the specific circumstances and the environment requiring the regulatory decision. Contract Law is, almost by definition, albeit not universally so, essentially horizontal across markets and economic environments. Although empirical studies of the real-world consequences of the rules and policy choices underlying general Contract Law are also necessary, the rules have to be drafted and operate at a more elevated level of abstraction than those of Competition Law, not the least by the heavy cost of changing the rules contained in centenary Civil and Commercial Codes.

Finally, other sources in EC Contract Law, which may touch on issues of breach and liability (Consumer Credit Directive, Sales Directive, Package Travel Directive) seem to me of limited use in the analysis of compensation after termination in long-term distribution contracts, and in particular regarding both the issues of unverifiable breach of the obligations of a relational contract, and the treatment of specific investments, given that those European rules typically handle transactions in which the level of non-salvageable specific investments is typically low, and in which the specification of contractual duties and their breach pose no specially onerous burden for parties and Courts in terms of verifiability.

Keeping the above mentioned caveats in mind, I will try, however, to summarize the main rules in EC Contract Law that may affect how to understand and regulate issues of compensation after termination.

## 2.1. Commercial Agents Directive

Undoubtedly, it is the 1986 Commercial Agents Directive<sup>7</sup> the one that deals with contractual scenario that falls within the area of long-term contracts used for making goods and services available in different stages of a distribution chain and in different geographic markets. The Directive, however, applies only where a relationship between a principal and a commercial agent is at stake. A commercial agent is defined in the Directive as a self-employed intermediary who has continuing authority to negotiate the sale or the purchase of goods on behalf of another person, hereinafter called the 'principal', or to negotiate and conclude such transactions on behalf of and in the name of that principal<sup>8</sup>. This leaves outside the scope of the rules of the Directive, including those on compensation following termination, the bulk of contractual arrangements through which distribution chains are structured: distribution (exclusive, selective, and others) agreements, supply agreements, franchising, and so on.

Artt. 17 thru 19 deal with remedial instruments following termination<sup>9</sup> –as foreseen in art. 13 thru 16- and oblige Member States to establish systems ensuring that a commercial agent, after termination of the agency contract, is either indemnified according to the rules broadly set out in art. 17.2, or compensated for losses or damages incurred in accordance with art. 17.3<sup>10</sup>. The entitlement to indemnity or compensation shall also arise if the contract ends due to the commercial agent's death<sup>11</sup>. The entitlement does not arise if termination was produced by the agent –unless that termination was justified by the behaviour of the principal<sup>12</sup>, or by personal circumstances that reasonably prevent the agent from continuing the activity-, if termination by the principal was a legally justified response to the agent's breach of contractual duties, or if the contract was assigned by the agent with the consent of the principal.

<sup>7</sup> Council Directive 86/653/EEC of 18 December 1986 in the coordination of the law of the Member States relating to self-employed commercial agents (OJ L382, 31/12/1986).

<sup>8</sup> Art. 1 (2) and 2.1 of the Directive clarify further its scope, by excluding from its reach: persons who, in their capacity as officers, are empowered to enter into commitments binding on a company or association; partners who are lawfully authorized to enter into commitments binding on his partners; a receiver, a receiver and manager, a liquidator or a trustee in bankruptcy; agents whose activities are unpaid; agents when they operate on commodity exchanges or in the commodity market.

<sup>9</sup> On these remedies after termination of the commercial agency contract, see GOYDER, 2005, p. 186. On the provisions of the different member states implementing the Directive, see BOGAERT and LOHMANN (eds.), 1993.

<sup>10</sup> The Commission has published in 1996 a report on the application of art. 17 in the internal laws of the Member States, and on the choices made by them between indemnity and compensation.

<sup>11</sup> The fact that compensation (including both indemnity and compensation in the strict sense used in the Directive) is to be paid in cases in which there is no influence on the behavior of the principal (like death, illness or retirement of the agent, or even bankruptcy of the principal) makes it clear that it does not pursue –at least exclusively- incentive goals. Again, a rationale of social protection of commercial agents as deserving a treatment not too dissimilar from employees, comes readily to mind.

<sup>12</sup> In case of breach of contract by the principal, moreover, the indemnity received according to art. 17.2, does not preclude the agent from seeking and obtaining damages for contract breach.

Art. 17.2 provides that the agent is entitled to an indemnity if and to the extent that two conditions are met: (i) the agent has brought the principal new customers or has significantly increased the volume of business with existing customers and the principal continues to derive substantial benefits from the business with such customers; (ii) to pay the indemnity is equitable having regard to all the circumstances and, in particular, the commission lost by the commercial agent on the business transacted with such customers (a covenant not to compete is a circumstance that Member States can determine to make the indemnity equitable<sup>13</sup>). This indemnity cannot exceed, in quantitative terms, the agent's average annual remuneration over the preceding five years –or the entire length of the relationship if it was less than five years.<sup>14</sup>

Art. 17.3 stipulates that the commercial agent is entitled to compensation for the damages suffered as a result of the termination of the relationship with the principal<sup>15</sup>. Such damages are considered to have occurred particularly in two set of circumstances: (i) when the agent has been deprived of the commission which proper performance of the agency contract would have procured, while providing the principal with substantial benefits linked to the activities of the agent; (ii) the agent was not able to amortize the costs and expenses incurred for the performance of the contract on the advice of the principal.

Art. 20 deals with covenants not to compete, in a way that severely constrains its use<sup>16</sup>. Such covenants –referred to as restraint of trade clauses- require, for their validity, to be concluded in writing, to be related to the geographical area or the group of customers and the geographical area entrusted to the commercial agent and to the kind of goods covered by his agency under the contract, and to last not more than two years<sup>17</sup>.

## 2.2. Consumer Credit Directive

In a different context, unrelated to long-term contracting in distribution, the Consumer Credit Directive<sup>18</sup> also contains a provision related to the consequences of termination of a –possibly long-term- contract other than the consequences resulting from breach. Art. 8 provides that the consumer shall be entitled to discharge the repayment and other obligations before the expiration of the contract term set in the agreement. In such a case, the consumer, according to the rules laid down by Member States, is entitled to an equitable reduction of the cost of the credit. Although the precise meaning of what is the consumer's entitlement after termination,

---

<sup>13</sup> Such a provision strongly reminds the reader of the logic of an employment relationship. For instance, in Spanish Law, mandatory rules in the field of employment contracts, mandate adequate compensation for the validity of such covenants [together with the existence of a relevant economic interest of the employer, a limit of 2 years: art. 21 *Estatuto de los trabajadores* –Employees Act-, and careful scrutiny by Courts: SSTS, 4<sup>a</sup>, 5.4.2004 (Ar. 3437), 21.1.2004 (Ar. 1727), 2.7.2003 (Ar. 2004\18), 21.3.2001 (Ar. 4106), 18.5.1998 (Ar. 4654)].

<sup>14</sup> The indemnity scheme adopted by the Directive has been directly inspired –copied may be closer to the truth- by the system foreseen in art. 89 b) of the German Commercial Code (*Handelsgesetzbuch*), for the *Handelsvertreter*. On this provision, see Hopt, in BAUMBACH/HOPT, 2006, p. 393.

<sup>15</sup> The compensation scheme adopted in the Directive as an alternative to the indemnity scheme is based on the *indemnité du prejudice subi* of French Law.

<sup>16</sup> Again, the resemblance to an employment relationship is strong here. See, note 11 above.

<sup>17</sup> The Directive expressly authorizes Member States to impose additional restrictions on the validity, enforceability, and scope of such covenants.

<sup>18</sup> Council Directive 87/102/EEC of 22 December 1986, for the approximation of the laws, regulations and administrative provisions of the Member States concerning consumer credit (OJ L 42, 12/2/1987).



and correspondingly, what is the right to compensation of the professional lender following termination exercised by the consumer, is not entirely clear, it can be convincingly argued<sup>19</sup> that the rule embodies a principle of pareto-efficient termination with compensation<sup>20</sup>: *“the consumer has the right to terminate when she compensates the other party to the extent that the latter is no worse-off than in the case of full performance.”*

It is clear that if one party fully compensates the other party –in the sense that the latter will be in an entirely equivalent position in terms of welfare that the one in which he would have been if the original plan set out in the contract had been fulfilled- there are no possible gains from blocking<sup>21</sup> the compensating party to put an end to the relationship: by definition the decision to terminate has to be pareto-efficient, because the compensated party is indifferent, and the terminating party must be better-off, or else she would not have opted for termination. The key problem with the implementation of this impeccable theoretical solution, pareto-efficient termination, is how to achieve full compensation of the other party.

In the context of a consumer credit contract, given that termination by the consumer is necessarily accompanied by repayment of borrowed capital, fully compensating the lender implies compensating the fixed costs of drafting and executing a new credit contract with the funds returned<sup>22</sup>, and in case of a fixed interest loan, the difference between the contract interest rate and the current market interest rate at the time of termination<sup>23</sup>. Given the massive use of such credit contracts, and the typically low amount of damages for the lender in a given case, ad hoc determination of damages may be too costly a strategy in terms of administration costs. It may be then practical to allow credit contracts to stipulate damages for the lender in the case of early termination by the borrower as a percentage of early returned capital, perhaps setting a cap on the amount that can be agreed –or set in standard form contracts by the lender<sup>24</sup>.

In fact, the Modified Proposal for a Directive on credit agreements for consumers, amending the current Consumer Credit Directive, of 7 October 2005, contains art. 15, on early repayment, with a more detailed treatment of the matter, concentrating not just in the consumer’s right to early discharge, but on the right of the creditor to a fair and objective indemnity according to the amount or the calculation method set out in the credit agreement. No indemnity shall apply, according to art. 15 of the Modified Proposal, when the period used for the borrowing rate was less than one year, and when early repayment was made under an insurance contract intended to provide a conventional credit repayment guarantee.

<sup>19</sup> See, BIANCA and GRUNDMANN (eds.), 2002, Chapter on questions 7 and 8.

<sup>20</sup> Par. 490b BGB contains a similar rule for credit contracts. Other legal systems use comparable rules for other contract types: See, for instance, art. 1594 Spanish Civil Code, on construction contracts (*locatio conductio operis*, *Werkvertrag*).

<sup>21</sup> A rule blocking termination in such circumstances would produce what in the economic theory of incomplete contracts would be called “inefficient trade”, and in the Law and Economics theory of Contract Law would be called excessive performance (because there was a case of efficient breach). Even if through renegotiation the likelihood of inefficient trade or excessive performance is eliminated or drastically reduced, it is clear that transaction costs would be –inefficiently- consumed in the process.

<sup>22</sup> To profitably use the funds returned would seem to be a required behavior under the principle of mitigation of damages generally applicable in Contract Law (art. 4.504 PECL).

<sup>23</sup> If the loan is a variable-interest loan, typically the old borrower would be paying the market rate, the same as a new borrower would, thus eliminating the differential in interest rates from the harm ensuing from early termination.

<sup>24</sup> This has been the preferred solution of the Spanish *Ley de crédito al consumo* (Consumer Credit Act, implementing the Consumer Credit Directive), setting caps of 3% for fixed-interest loans, and 1,5% for variable-interest loans.

As much as it is wise not to restrict this sort of pareto-efficient termination, it is unwise to allow only apparently pareto-efficient termination, or equivalently, to condition termination –absent demonstrable breach by the other party- on full compensation paid by the terminating party. This would make sense if termination can only respond to the desire to avoid inefficient trade or excessive performance. Such an attitude, however, would in other circumstances, entirely prevent the use of termination as a useful instrument to deter non-verifiable instances of breach of contract by the other party. As will be developed in more detail in the following section, particularly, albeit not exclusively, in long-term, relational contracts, many situations of breach of contractual obligations cannot be ex post shown to be real to a Court in a legal dispute, that is, using the terminology of the economic theory of incomplete contracts, they are non-verifiable dimensions of the contract. The threat of terminating the relationship is a powerful tool –and maybe the only one, in some cases- to discipline the other party on those dimensions of the relationship that are non-verifiable. To subject termination to full compensation would entirely undermine the –desirable, in these cases- deterrence effect of termination, because the breaching party –remember, breach cannot be verified here, and thus cannot be deterred using ordinary contract remedies such as specific performance or damages- would then prefer to breach because he would have the expected portion of the contractual surplus guaranteed if the other party terminates –and if she doesn’t, he gets it anyway- plus the savings in performing the non-verifiable contractual obligations resulting from breach.

In the case of a credit contract, of course, such queries do not affect early termination by the borrower (be it a consumer or a professional) because, almost by definition termination cannot serve a function of disciplining the lender, because he has already performed by giving the capital to the borrower<sup>25</sup>. This is why in this setting checking pre-contract informational advantages of the lender is hugely more important than checking post-contract opportunistic behaviour by the lender.

### 2.3. Package Travel Directive

The relevant rules in the Package Travel Directive<sup>26</sup> dealing with compensation after termination are contained in art. 4.5, and 4.6. The first provision deals with the termination right of the consumer if the organizer significantly modifies the conditions of the contract:

*“If the organizer finds that before the departure he is constrained to alter significantly any of the essential terms, such as the price, he shall notify the consumer as quickly as possible in order to enable him to take appropriate decisions and in particular:*  
*- either to withdraw from the contract without penalty,*  
*- or to accept a rider to the contract specifying the alterations made and their impact on the price.”*

It seems obvious –not just from the text of the provision, but also from the logic behind the grant of termination rights in this particular case- that no compensation is payable to the other party when the consumer exercises the right to termination. What remains to be seen is the more

<sup>25</sup> An exception may be the case of tied (credit and the sale of goods and services) contracts foreseen in art. 11 of the Consumer Credit Directive, in which there is room for post-contract opportunistic behavior by the lender-seller-provider of services.

<sup>26</sup> On the Directive, see WEATHERILL, 2005, p. 98.

controversial issue of the existence of a generalizable principle of Contract Law behind this rule: If one party substantially –but with justification, or otherwise the alteration would be to no effect, or would constitute a breach of existing obligations- modifies the terms of the contract, the other party can terminate the relationship, and the party deciding the modification shall not be entitled to any compensation for the termination of the relationship.

There may be reasons in favour of generalization of such a rule, under the heading of the principle of securing equality of arms, as GRUNDMANN has aptly characterized<sup>27</sup> the grant of termination rights in the Package Travel Directive. I have, however, some hesitations about mandating an absolute rule of no compensation in circumstances –in terms of termination following a change in contract terms decided by the other party- comparable to those mentioned. Two reasons lie behind my doubts. One is that the rule in the Directive applies to a contract which is, true, not a pure spot transaction, but that typically is neither a long-term term relationship, and only very improbably possesses significant relational elements. It is a lagged performance contract, but not a long-term contract. Thus, the need to adjust the contract terms of a package travel to changed circumstances arises mainly in events that have little to do with setting governance mechanisms for an open-ended cooperation within a distribution chain. They would typically cover events that may be akin to those that fall under the general heading of *force majeure* or similar kinds of events<sup>28</sup>. In fact, if one thinks of a paradigmatic long-term relationship, such as the corporate contract, one realizes that the parties are not given exit options for all changes in the contract –the corporate charter- but only for the most relevant ones. Similarly, in marriage, those systems sticking to divorce for cause award it only for serious changes in the terms of the relationship, and not for all of them.

The second reason is that in a typical package travel neither party is likely to make relevant relation-specific investments, the ones we should precisely care about in long-term distribution contracts, in order to come up with producing the right incentives so the parties invest the right amounts. In sum, my view is that the factors that in long-term contracts speak in favour or against compensation after termination have little to do with the circumstances that typically can give rise to the uncompensated termination in art. 4.5 of the Package Travel Directive, and thus this should exercise little influence in the understanding and, eventually, drafting, a desirable legal regime for the type of relationships we concentrate on.

Art. 4.6 of said Directive, in turn mandates:

*If the consumer withdraws from the contract pursuant to paragraph 5, or if, for whatever cause, other than the fault of the consumer, the organizer cancels the package before the agreed date of departure, the consumer shall be entitled:*

*(a) either to take a substitute package of equivalent or higher quality where the organizer and/or retailer is able to offer him such a substitute. If the replacement package offered is of lower quality, the organizer shall refund the difference in price to the consumer;*

*(b) or to be repaid as soon as possible all sums paid by him under the contract.*

*In such a case, he shall be entitled, if appropriate, to be compensated by either the organizer or the*

<sup>27</sup> See, BIANCA and GRUNDMANN (eds.), 2002, chapter on questions 7 and 8.

<sup>28</sup> The reference in 4.6 (ii) to *force majeure* events indicates that this connection may exist.

*retailer, whichever the relevant Member State's law requires, for non-performance of the contract, except where:*

- (i) cancellation is on the grounds that the number of persons enrolled for the package is less than the minimum number required and the consumer is informed of the cancellation, in writing, within the period indicated in the package description; or*
- (ii) cancellation, excluding overbooking, is for reasons of force majeure, i.e. unusual and unforeseeable circumstances beyond the control of the party by whom it is pleaded, the consequences of which could not have been avoided even if all due care had been exercised.*

The compensation in this provision operates in circumstances of lack of performance (due to cancellation by the provider, or involving termination by the consumer after a change of contract terms that is not justified) by the provider of travel services (organizer, retailer, or both). We are, thus, in the area of breach of contract and compensation as a remedy for breach. It refers, moreover, to instances of breach that can adequately be considered as verifiable by a Court or a third-party adjudicator, so we find ourselves squarely in the archetypical case of a remedy of damages. Not that this remedy does not pose problems of its own in a package travel contract (the issue of non-pecuniary damages is of great significance here) but it does not seem to provide a basis for our setting of long-term distribution contracts.

## 2.4. Consumer Sales Directive

The Consumer Sales Directive also contains a rule on termination, namely art. 3.5<sup>29</sup> entitling the consumer to rescind the contract –in addition to requiring and appropriate price reduction– when the consumer is not entitled to the primary remedies of repair and replacement, or when the latter two remedies have not been completed by the seller in a reasonable time or without significant inconvenience to the consumer, and provided that the lack of conformity –the fact that triggers the set of remedies adopted by the Directive– is not minor.

As should be apparent, termination by the consumer in this circumstance operates as a remedy for breach and thus as an incentive device –coupled with others– to induce adequate levels of quality in the provision of consumer goods by professional sellers. Being the sale of consumer goods the paradigm of a spot contract, and being specific investments by either party seldom made with respect to mass-marketed goods, it seems clear that this provision has very limited, if any, implications for the realm of long-term distribution contracts.

## 3. The economic considerations behind compensation after termination

Termination and compensation after termination are two of the most relevant –albeit not the only ones– dimensions of the long-term contracts through which distribution chains are formed and structured. Suppliers of components or inputs and manufacturers, manufacturers and distributors at different levels, franchisors and franchisees, they all engage in contractual relationships varying in degree of legal formalization, scope, intensity of the interrelation, and initial term of duration. The contractual

---

<sup>29</sup> See, Massimo C. BIANCA, Commentary to art. 3.5, in BIANCA and GRUNDMANN (eds.), 2002, p. 164.

links between the different agents in the distribution chain (manufacturer, wholesaler, distributor, retailer, and consumer) and apart from the retailer-consumer relationships, (not uniformly, though) to be long-term in duration and nature<sup>30</sup>.

Though contracts within the distribution chains vary widely in actual commercial practice, they tend to share some important common features –with some exceptions, however, particularly regarding franchising: selectivity (the search for a contracting party is lengthy and costly, and entails the use of screening mechanisms or probationary periods); informality (many contracts are not in writing, or when they are in writing, their content is extremely sketchy), which determines a substantial and often almost entirely unconstrained interpretation and renegotiation by the contracting parties; the use of framework agreements, typically developed, specified or executed by more detailed or specific contracts for particular projects, invoices, or items; open-ended character (many contracts have no specified term of duration, or when they do, they are typically renewable automatically).

All these features fit quite nicely with the most relevant dimensions of the relational (or long-term, in the economic sense) contract notion, which has attracted considerable attention and effort from the Economic literature on contracts and organizations. For our task of understanding the role and effects of legal rules on long-term contracts in distribution chains, and building a desirable set of rules, several lessons can be drawn from the stock of analytical and empirical findings accumulated by the economic analysis of long-term contracts.

Moreover, I believe that these lessons are the more relevant ones in order to address the questions surrounding the legal treatment of contracting in distribution chains. I am not implying that economic analysis alone can provide definitive answers to the underlying issues confronting the legal regime of long-term distribution contracts. Legal rules and institutions in isolation do not capture the full picture of cooperation in relational contracts. In fact, it has to be acknowledged that the big questions concerning the actual effects of contract Law on long-term contracts remain to some extent unanswered by economic theory in a fully convincing and comprehensive way. But I am extremely skeptical concerning the ability of legal rules and institutions in isolation to capture the full picture of cooperation in relational contracts, and not the least in distribution chains. Disregarding the variables and dimensions that economic analysis has unearthed as relevant for achieving efficient outcomes in relational contracts will be done by legal analysis at own risk. Inducing efficient behavior by the firms integrated in distribution chains seems to me the preferred goal that legal rules can, at best, approximate in this context. To focus exclusively on the “traditionally legal” dimensions might induce a design of rules and legal instruments for contracts in distribution chains that interfere the goal of inducing efficient behavior, with the undesired result of reducing, rather than increasing, the level of cooperation in the market, and, ultimately, social well-being.

---

<sup>30</sup> The long-term character is sometimes (not infrequently, in the supplier-large retailer relationship) the result of short-term contracts that are renewed at the end of each contractual period, over a long time horizon.

### 3.1. The importance of the open-ended nature of the relationship

In distribution chains, if (a big if, but that concerns relatively little the reach of Contract Law within distribution chains) consumers have full information on the goods and services provided to them through the distribution channels, and there is competition among chains, what is required for an efficient allocation of resources is that the different agents acting as parties in the various contract links in the chain cooperate to produce the adequate combination of attributes of the product or service: quantity and quality of inputs, of pre- and post-sale service, of advertising and market information effort, and so on. The goal would then be to maximize cooperation –or the contract surplus, to use a more precise, or jargon-laden, term- among the parties in the contracts in the distribution chain (supplier-manufacturer, manufacturer-distributor, franchisor-franchisee, and so on)<sup>31</sup>.

Many naïve analysts may think that the fact that contracts in distribution channels are long-term, so that parties have much to gain from continuous and iterated cooperation to achieve the maximum available surplus can be a powerful motivator or incentive instrument for the good.

But as the game-theoretic analysis of economic interaction shows, if there are incentives not to cooperate in isolated or one-shot interaction, simple repetition of the game is not enough. The economic idea of how a long-term relationship can –abstracting from external forces such as the Law, or social norms- can sustain cooperation among parties with divergent desires, goes beyond the mere repetition of a one-shot interaction. We need open-endedness.

Let's assume that we have a relationship between a manufacturer and a distributor. In this relationship, the manufacturer, Player 1, sends goods to the distributor, Player 2, but customers value pre-sale advice and service. The distributor can take two actions regarding these pre-sale activities, high effort or low effort. The efficient action from the point of view of total contract surplus is to provide high effort, though high effort is personally more costly for the distributor. The pay-off for the distributor is determined prior to the provision of pre-sale services, through the margin on the price paid to the manufacturer.

An interaction like this can be represented by a one-shot dynamic game, such as the so-called trust game<sup>32</sup>: Here, not trust and betray are non-cooperative actions, whereas trust and honor are the cooperative actions.

The game begins with a decision node for Player 1, who can choose either to Contract or Not to Contract with Player 2. If player 1 chooses Contract, then the game reaches a stage in which Player 2 can choose either to exercise High or Low service effort. If Player 1 chooses Not to Contract, then the game ends (in fact, Player 1 chooses not to initiate a relationship).

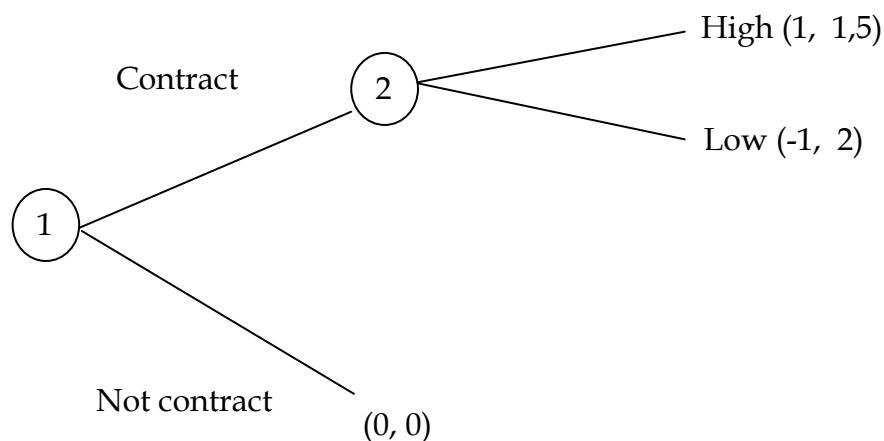
If Player 1 chooses not to contract with the distributor, Player 2, both players' payoffs are zero. If Player 1 chooses to contract and to trust Player 2, however, then both players' payoffs are one if

---

<sup>31</sup> The assumptions of competitive structure and full consumer information guarantee then that consumers would be the ultimate beneficiaries of achieving the best available contractual surplus in each link within the distribution chain.

<sup>32</sup> David KREPS, in ALT and SHEPSLE (Eds.), 1990.

Player 2 provides high service effort, but Player 1 receives -1 and Player 2 receives a pay-off of two if Player 2 decides to shirk and save in service costs by providing low effort.



The Trust Game can be easily solved by backwards induction or rollback—that is, by working backwards through the game tree, one node at a time:

If player 2 gets to move (*i.e.*, if Player 1 has chosen Contract) then Player 2 can receive either a payoff of 1.5 by providing high effort, or a payoff of 2 by choosing low effort. Since 2 exceeds 1.5, Player 2 will choose low effort if given the opportunity to do so. Knowing this, Player 1's initial choice amounts to either not initiating the relationship (and so receiving a payoff of zero) or contracting with the distributor (and so receiving a payoff of -1). It is clear then that the manufacturer prefers not to enter into the relationship with the distributor. Notice that (Contract, High) is the strategy that leads to higher joint surplus of the parties, 2.5, higher than with (Contract, Low), which is 1, and with (No contract, Low) –the Nash equilibrium in this game– which is zero. So cooperative behaviour cannot be sustained even if both parties would be better-off by establishing and honouring the distribution relationship.

Unfortunately, having a future of potential cooperation does not in itself improve matters and increases welfare for the parties. Imagine that the interaction can take place, in exactly the same terms as shown above, a very large number of times. That is, the difference is that we have a repeated game, played a very large number of periods,  $T$ . Would caring about the future make the distribution relationship sustainable?

Let's think first, using again the perspective of backwards induction, of the last period. In the  $T$ th period, Player 2 has no reason to provide high effort, because he can obtain a higher pay-off by choosing low effort, given there is no future to care about in this last period. Anticipating this, Player 1 will not enter the relationship in the last period,  $T$ . Then, the last possible effective period of the relationship is the  $T-1$ th period. Here, Player 2 knows it is the last period, and therefore he would behave accordingly, providing low effort. Anticipating this, Player 1 will not play the  $T-1$ th round, making the  $T-2$ th the last plausible period. And similarly the dismal logic

of the game unravels down to period 1, and both players find themselves exactly in the same position as in the one-shot interaction. Finite repetition does not improve cooperation in economic exchange. Cooperation remains unachievable when the parties know their relationship will last for a fixed and known length of time<sup>33</sup>.

But if the same interaction can be repeated an infinite number of periods or, more realistically, that it can be played an unknown –by the parties– number of periods, so the interaction can end at any round, with a given probability which the parties (or in this example, the distributor) are unable to control.

In such a scenario, it is possible to show that if the parties care enough about the future (i. e., the discount factor at which they discount pay-offs ahead in the future is not too large), reciprocity strategies<sup>34</sup> in the repeated interaction can lead to cooperative outcomes<sup>35</sup>.

This powerful result that reciprocity can achieve cooperation has been crucial to the relevance of long-term contracts in recent economic theory, because it implies that even in the absence of other external mechanisms to achieve cooperation in economic exchange, most notably, even in the absence of complete, formal, and legally enforceable contracts, in long-term relationships, the parties themselves can attain, under certain conditions, satisfactory cooperative outcomes. But this result stresses the importance, in relationships in which, for different reasons<sup>36</sup>, the parties cannot, or do not want to, rely on legal remedies to induce cooperative behaviour, and to discipline uncooperative behaviour, of the open-ended nature of the contract. The last-period curse which we have illustrated before, makes apparent that the common practice in long-term contracting of not fixing a pre-specified duration, is not the product of inadvertence or naïvete, quite the contrary, it is a conscious choice intended to provide incentives through reciprocation and reputation. But for this mechanism to work effectively, it is necessary that the parties can make use of the disciplining or reciprocation strategies that serve to check the private incentives to deviate from cooperative behaviour, basically, termination of the contract at any time<sup>37</sup>. In fact, legal systems seem to acknowledge this function, by allowing parties to enter into open

---

<sup>33</sup> This result is known in game theory as the unravelling result, and was demonstrated by Reinhard Selten in his theorem on unique subgame perfect equilibria.

<sup>34</sup> Although many economists use terms such as trust, or reputation, in fact they refer mostly to something that is best captured by the term reciprocity. For instance, an infinitely repeated interaction, such as the one described in the text, is equivalent to one-shot two-players interactions among all members in a given society, where each player is perfectly informed about the reputation (the past history of cooperation or defection), of each and every other player. This shows that reputation is, in conceptual terms, not very different from reciprocation.

The two reciprocity strategies most used in the game-theoretic literature are the following: (i) The grim trigger strategy, that states for each player: Choose a cooperative action until the other player deviates from cooperation; once the other player has defected, play defect forever after. (ii) The Tit-for-tat strategy, that states for each player: Choose a cooperative action if the other player cooperated in the previous period; defect if the other player defected the previous player. It seems that this attitude of reciprocal altruism has some biological roots, or at least is not exclusive of humans: TRIVERS, 1971, p. 35. A delightful application of these ideas to economic and social cooperation, in SEABRIGHT, 2004.

<sup>35</sup> See Joel WATSON, 2002; Robert GIBBONS, 2001.

<sup>36</sup> Some of them will be explored in the next subsection.

<sup>37</sup> Termination is, unfortunately, a double-edged weapon, because it can serve, as asserted in the text, as an indispensable part of a reciprocation strategy to check uncooperative actions by the other party. Termination, or the threat to terminate, however, in certain circumstances can also serve as an instrument of uncooperative behaviour to take advantage of the other party to appropriate a bigger share of the contractual surplus. This makes the legal regime of termination and its consequences particularly difficult and, to some extent, even potentially contradictory.



ended, potentially of infinite duration, relationships, but allowing them to retain the right to terminate at any time<sup>38</sup>.

### 3. 2. Incomplete contracts and termination as a disciplining mechanism against non-verifiable breach of distribution contracts

If parties could write perfectly complete contracts specifying all relevant actions<sup>39</sup> for the parties, the pareto-optimal actions, and Courts could costlessly enforce them, of course, the short or long nature of the relationship, and the use or not of reciprocity strategies, is of no importance: Contract Law can force cooperation through the use of the appropriate legal sanction. When individuals and firms can write a complete contract that determines contractual behaviour in the whole set of possible contingencies, and the basis of each contractual determination can be verified by the legal enforcer, typically, a Court of Law, the role of the legal system would be (at least in the eyes of an economist) is essentially mechanistic: the role of the Law would be substantially equivalent to blind enforcement of each and every term in the complete contract set out by the parties.

The problem for us, in the real world, is that the complete contract notion is a conceptual construct. Contracts in the real world are never complete, given the extremely exacting requirements that pareto-optimality, completeness of coverage, and verifiability of information pose on complete contracts.

A number of reasons have been identified by economic theorists supporting the proposition that contractual incompleteness (and thus the preponderance of relational contracts) is the rule and not the exception: The prevalence of non-verifiable information concerning many relevant contractual behaviors; the inevitability of imperfect specification of actions that will take place in the future, and in all possible states of the world in the future; difficulties in measuring and evaluating the cooperativeness of contractual behaviors, given the multidimensionality of complexity of many of them; advantage, under certain circumstances, of internal (including reciprocity-based) motivators for cooperation over external mechanisms such as formal and legally enforceable contracts<sup>40</sup>.

Lawyers tend to add other factors to the long list of reasons that make drafting complete contracts chimerical in our imperfect world, basically the unavoidable ambiguities and uncertainties brought about by the use of ordinary language, the one in which contracts are written, and bounded rationality on the part of the parties, that prevents them from perfectly anticipating all future contingencies, or fully exploiting the gains from trade between the parties in all future states of the world<sup>41</sup>.

---

<sup>38</sup> Art. 15 Commercial Agents Directive is a good example of this attitude.

<sup>39</sup> Actions that, by definition, would be pareto-optimal: If contracting partners jointly agree on them ex ante it can only be because they maximize joint welfare of the parties, that is, they are pareto-optimal.

<sup>40</sup> See Martin BROWN, Armin FALK, and Ernest FEHR, 2002, for experiments showing how explicit and enforceable contractual clauses may crowd out reciprocity and altruism in contracting. On these general issues, Colin CAMERER, 2003.

<sup>41</sup> Benjamin HERMALIN, Avery KATZ, and Richard CRASWELL, "Law and Economics of Contracts", in A. Mitchell POLINSKY and Steven SHAVELL (eds.), 2006.

### 3.2.1. Incompleteness in long-term distribution contracts and the role of termination

If incomplete contracts are the rule in any ordinary setting of economic interaction, when long-term contracts, in distribution chains or in other areas, are the norm, the force and the extent to which the factors leading to incompleteness is even stronger and more apparent. 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 the other, have an effect on the contractual outcomes, dramatically increases. How can a contract adequately foresee a substantial number of the future events that can affect the kind and quality of inputs, the conditions of demand and production functions, the contingencies affecting the firm structure of the contractual partners, and so on? No way, is the short but truthful answer.

This is why economists, when approaching long-term relationships, have routinely assumed that contracts between the parties are incomplete, that is, they do not specify the complete set of optimal actions by the contractual parties. 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 the set of incentives arising from factors (personal, institutional) different from the formal contract and the legal rules in contract Law. The relational contract can be based on information that cannot be verified by a Court of Law, or controlled by formal legal rules and procedures. The relational contract can be based on information only at the disposal of the parties as it becomes available, maybe only as a result of a change of circumstances.

Of course, a relational contract in this sense cannot be enforced purely as written<sup>42</sup> by a Court or an arbiter<sup>43</sup>. This does not mean that any contractual clause or any contractual behavior by the parties is able to be legally enforced. Some instances of undisputable breach of the contract – departures from the cooperative equilibrium in the long-term relationship – can be detected by the contractual partner, and verified in front of a Court (or there is a substantial probability that this can happen), and thus can be deterred by the use of legal remedies, such as specific performance and damages. For example, if the franchisee in a fast-food franchise chain cheats on the quality of the raw materials to save costs, it may be the case that the fall in quality is detected by consumers and by the franchisor (let's think of repeated cases of food poisoning), so that the

---

<sup>42</sup> In 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. Very commonly, only something very sketchy and obligatorily incomplete has been agreed, leaving enormous space for future adjustment of terms through more or less formal renegotiation, or to future decisions, again, more or less formal, by one party.

<sup>43</sup> 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 radical incompleteness of the relationship challenges, at least at the theoretical level, the role of Law and formal, enforceable contracts, in long-term relationships: Does the use of contract Law promotes –reinforces the incentives created by other means– or undermines –weakens the positive effect of other means– cooperation in long-term contracts? The response from economic theorists is mixed and cautious. The Law, and Contract Law in particular, can be a powerful instrument to create incentives for cooperative behavior in economic exchange, but in theory it is not a priori certain whether contract Law overall improve or weaken the self-enforcing nature of relational contracts.

latter can sue for damages and obtain compensation. If (a big if, however) damages correspond to the decrease in value of the franchise chain due to the breach of contract by the franchisee, deterrence of the undesirable and non-cooperative contractual behavior is achievable using the ordinary contract remedies.

Unfortunately, this is not always the case with a wide variety of cases of non-cooperative behavior within the contractual relationship. Many sorts of contractual behavior pose unsurmountable measurement problems, particularly when those behaviors are multi-dimensional. Maybe, concerning most of these, the contracting partner can observe the departure from expected behavior, or even only a signal that is correlated with a higher likelihood of non-cooperative behavior of the other party. The chances that these non-complying contractual behaviors can be shown, in a sufficiently convincing manner, to a Court or other external adjudicator, is very low. How can a manufacturer demonstrate, satisfying the level of proof legally required, that the distributor provides helpful and adequate advice to prospective buyers, or that the personnel of the distributor is cooperative with respect to the manufacturer's representatives, or that marketing effort through leaflets or other written material is satisfactory? May be only with cameras or other recording devices, or for cases of most egregious violations of the expected contractual efforts, giving rise to consumer complaints. Moreover, the cost of detecting and collecting evidence of these instances of breach, increase with the size, territorial, and product scope of the distribution chain, thus becoming more serious the more integrated and expansive the relevant consumer markets. The amount and scope of unverifiable breaches of contract would tend to increase not only with the chances that the breaching party would scape undetected or not subject to legal contractual remedies, but also with the chances that consumers would not detect or punish the defecting distributor (percentage of non-repeat customers, informational deficiencies of consumers, and so on).

It is true that parties to a long-term distribution contract can resort to several alternatives that can serve as -imperfect- substitutes of perfect Court enforcement of contractual remedies against verifiable breach of contract<sup>44</sup>.

First, the contract can use, instead of a non-verifiable dimension, some verifiable proxy of 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: exclusive territories awarded to distributors, so they become more the residual claimants of efforts in the promotion and service associated to the product; imposing certain suppliers of inputs on the distributor, so that cheating in quality is more costly; imposing minimum purchase requirements, etc.

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

---

<sup>44</sup> See, George MATHEWSON and Ralph WINTER, 1985, p. 503; Benjamin KLEIN and Kevin MURPHY, 1988, p. 265; Benjamin KLEIN, 1995, p. 9; Cándido PAZ-ARES, 1997, and 2003, p. 31.

the use of future quasi-rents for the distributor linked to the continuation of the relationship<sup>45</sup>: If the terms of the contract are adjusted so that the distributor (or franchisee, or supplier of inputs, if one thinks of upstream chains from the perspective of the manufacturer) expects to earn quasi-rents on their investments in the distribution chain, if and only if the existing relationship goes on, the party has a powerful motive to stay into the contract, and thus to avoid the kind of negative behavior that can trigger the end of the relationship. One important source of quasi-rents (returns of an investment that are higher than the return in an alternative use) is the existence of relation-specific investments, and a crucial link to the continuation of the contract is a credible threat by the other party to terminate the relationship, if the deviating behavior takes place (although, remember, this fact cannot be shown with sufficient evidentiary force before a Court or other external adjudicator).

Therefore, in order to make this instrument of relation specific quasi-rents to work as an incentive mechanism to achieve cooperation in dimensions outside of what can be verified by a Court, preserving the effectiveness of the termination weapon is crucial. Thus, subjecting termination to a duty to compensate the other party dilutes the force of the threat to terminate. The terminating party would be less inclined to use termination, because, *ceteris paribus*, this has become a more costly option due to the adjoining obligation to compensate, or indemnify, the party (who, remember, at least in this hypothesis, is in breach, however non-verifiable this may be)<sup>46</sup>. The result is likely to be that the manufacturer deploys other –less efficient, arguably, or else he would have used them from the start– means to check the non-cooperative behavior of the distributors, or eventually replacing, in a Coasean fashion, contract –distribution contracts– by authority –vertical integration of distribution inside the firm.

In sum, considerations concerning the essentially incomplete nature of long-term distribution contracts, and the primary relevance of disciplining unverifiable breach of contractual duties by the distributor, would lead then, to a basic position against compensation after termination such contracts.

### 3.2.2. Empirical evidence on the effects of legal restrictions on termination

The previous subsection has summarized the economic arguments advanced by economists and economically oriented legal scholars, in favor of preserving the conditions of unconstrained –by legal requirements concerning good cause, or the imposition of compensation *ex post*– exercise of termination at will in long-term distribution contracts. This view seems to be supported by the available empirical evidence on the effects of legal rules restricting termination<sup>47</sup>. This evidence refers to franchising<sup>48</sup>, but there does not seem to be a powerful reason to doubt that the main

<sup>45</sup> The path-breaking analysis of this instrument is Benjamin KLEIN, and Keith LEFFLER, 1981, p. 615.

<sup>46</sup> Benjamin KLEIN, 1995, p. 28-30, 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.

<sup>47</sup> See, James BRICKLEY, Frederick DARK, and Michael WEISBACH, 1991, p. 101; John BEALES III, and Timothy MURIS, 1995, p. 157; Jonathan KLICK, Bruce KOBAYASHI, and Larry RIBSTEIN, 2006; James BRICKLEY, Sanjog MISRA, and Lawrence VAN HORN, 2006, p. 173.

<sup>48</sup> The reason for this, lies in the fact that the studies are based on the US experience, where State legislation interfering with termination at will has concentrated on franchise contracts. Moreover, it seems that franchise still play a major role in US distribution compared with other areas.

findings would not be applicable to other contractual arrangements in distribution chains sharing similar issues of controlling opportunism by distributor (and, as we will see in a moment, also by manufacturer).

The first and best-known piece is BRICKLEY, DARK, and WEISBACH. They hypothesize that laws restricting franchisor termination rights will lead to less franchising. The reason lying in the fact that if, for instance, cheating franchisees receive compensation after the franchisor terminates, the benefits from cheating increase, and thus the amount of breach. This would lead to less profitable franchising, making other arrangements (such as the franchisor directly running the units) more profitable by comparison. Interestingly, because franchisees are assumed to be able to generate higher revenue in the operation of units than are franchisors themselves, the reduction of franchised units also leads to an aggregate reduction of units: While the franchisor will now –after laws restricting termination– find profitable to run some of the units it would have franchised were it able to commit the franchisee not to cheat, there will be some marginal units that are no longer profitable to run or to franchise.

BRICKLEY, DARK, and WEISBACH also consider that unconstrained termination can be used by the franchisor not only to discipline non-cooperative behaviour by franchisees, but also to exploit and abuse franchisees and try to own (and thus, not share the profit with the franchisee) those units, that through franchisee's sales effort, or through franchisee's market discovery, turn out to be particularly profitable<sup>49</sup>. But if this is the case, and franchisors use their termination rights to expropriate franchisees of their specific investments, and franchisees do not correctly estimate the expected cost of this opportunistic behaviour on the part of franchisors, there will be too much franchising as some franchisees pay –as franchise fee– above their true reservation prices for their units. Laws restricting termination by franchisors would also decrease franchising in this scenario.

BRICKLEY, DARK, and WEISBACH, however, rule out this second possibility by focusing their empirical analysis on differences across industries. Specifically, they argue that if termination primarily serves to discipline franchisee's non-cooperative behaviour, the effect of termination laws on the rate of franchising will be most pronounced in industries with mostly non-repeat business. In areas or sectors with significant repeat business, disciplining franchisees is less important, since the self-enforcement mechanisms induce better behaviour from the franchisee: Otherwise, it will lose its repeat business and suffer a large revenue loss.

In industries without much repeat business, there is less potential for self-enforcement, making termination more important as a policing tool. On the other hand, if termination clauses primarily allow the franchisor to exploit the franchisee, no such cross-industry relation would

---

<sup>49</sup> Both explanations for termination by the franchisor (or by the manufacturer, or principal, more generally), the benevolent –discipline on non-verifiable breach by the other party– and the sinister (expropriation of value from specific investments by the “weaker” party) are consistent with the brute factual observation that it is principals, and not the other parties, who typically terminate the relationship. In a Spanish survey carried out by a business daily newspaper (*Expansión*, 9 December 1996), in 88% of the cases termination is decided by the principal, in 8% by the distributor, and in 4% it is a joint decision (I have taken these figures from Cándido PAZ-ARES, 2003, p. 32). A casual look at litigated cases also strongly points in the direction of the principal or manufacturer being the party behind most disputed cases of termination. In view of this, it seems that we need some more elaborated empirical analysis to test which theoretical explanation is empirically corroborated by facts.

appear, and no systematic difference in the change in franchising across industries would be found.

BRICKLEY, DARK, and WEISBACH data show that the effect of termination restrictive legislation is greater (and statistically significantly so) in the industries they classify as particularly subject to non-repeat customers (restaurants, hotels, and auto rental agencies) as compared to the effect in other sectors.

BEALES and MURIS, in turn, look at whether data on franchise terminations and non-renewals support the efficiency or opportunistic explanation for terminations. What they label an efficient termination is the one in which the franchisor detects breach of quality provision duties by franchisee. They define opportunistic termination as any non-efficient termination, presumably for the exploitative reasons mentioned earlier. They have data on terminations (by both franchisor and franchisee) in thirteen industries over eight years. Their independent variables include: growth in number of outlets (which should increase breach); growth in sales per outlet (should decrease breach); and, proxies for appropriable rent (should increase opportunistic terminations).

Their results do not support or reject the opportunism hypothesis — the estimated coefficients are often of the wrong sign or are statistically insignificant. They do obtain, however, a robust, significant, and negative coefficient on the “growth in outlets” variable. This is likely to imply if opportunism or expropriation by the franchisor is a factor, it is watered down by the franchisor’s interest in keeping a reputation in order to attract additional good franchisees.

KLICK, KOBAYASHI and RIBSTEIN also use data on franchising chains to test the relative importance for termination of the disciplining or the expropriation story. They use state laws limiting franchisor termination rights to identify the effect of termination at will on both the choice to franchise and on franchisor expansion generally.

In the first set of empirical tests, they use firm-level data on franchising in the fast food industry, and their regressions show that constraining termination leads to a reduction in franchising, and to a smaller increase in franchisor-operated units. Their second set tries to connect the changes in laws conditioning termination with state employment in industries characterized by a high degree of franchising. There they find that restrictions on termination at will are correlated with a decrease in the franchised industries’ employment rates relative to employment rates in industries with little franchising.

Both sets of tests, thus, tend to give support to the view that the disciplining effect of termination on franchisee’s non-cooperative behaviour seems to outweigh, in their empirical relevance, the opportunities for franchisor abuse and expropriation of value that termination at will may allow.

BRICKLEY, MISRA, and VAN HORN have looked not at the legal provisions concerning termination of franchise contracts, but at whether an “exploitation” theory of franchising (powerful franchisors are able to impose contract terms on weaker franchisees) is confirmed by empirical data. They concentrate on clauses regulating contract duration, as these are typically crucial on

the chances that the franchisees recover the relation-specific investments (those that loose all, or a substantial fraction of value outside the contract<sup>50</sup>) 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 the chances of complete recovery of investment by the franchisee will be.

Using a large sample of franchising firms, they analyze the effects on contract duration clauses of several factors: the number of years the franchisor has been in operation; the number of sites comprised within the franchising network –that is, the size of the franchisor; the average total initial investment of a franchisee entering the franchise network; the number of weeks of off-site training of franchisee’s personnel. The first two factors relate to the power, experience and contractual strength of the franchisor, the second two are good proxies of the level of specific investments made by the franchisee. If the exploited franchisee view were correct, we would expect that the bigger, more sophisticated the franchisor, the more exploitative the contract terms, here, the shorter the contract duration, will be. Again, if the naïve franchisee image were true, the level of specific investments would not raise contract duration, given that the exploitative franchisors would try to appropriate the value of the non-amortized specific investments incurred by the franchisee.

Empirical results show that the four factors are positively and significantly 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, contrary to what the “exploitation” hypothesis would predict. And these results hold with and without taking into account the fixed effects of the particular industry in which the franchisor operates. There seems to be, therefore, evidence indicating that franchisors are responsive to the level of specific investments by franchisees, and they are more responsive the bigger, and better established they are. These results would provide indirect evidence that the threat posed by opportunistic and exploitative behaviour on the part of franchisors does not seem to be a particularly worrying problem in reality<sup>51</sup>, or at least, to be sufficiently secondary not to show in the data.

There is also a striking feature arising from some results in the empirical studies just summarized. It is the observation that legislation restricting termination at will increases, rather than decreases, the number of terminations, that is, that when the law sets some conditions – financial compensation or showing good cause for termination- for terminating a franchise contract, franchisors terminate more often, and not less, as could intuitively be expected.

---

<sup>50</sup> In the next subsection I present a more detailed and rigorous notion of specific investments, and analyze its implications for long term contracting and contract remedies.

<sup>51</sup> It is true, however, that BRICKLEY, MISRA, and VAN HORN also find a positive effect of legal restrictions on franchise termination (in the state where the franchisor has its headquarters) on contract duration clauses: BRICKLEY, MISRA, and VAN HORN, 2006, p. 185. 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.

The explanation that has been forwarded for this counterintuitive empirical finding runs along the following lines<sup>52</sup>: unconstrained termination at will induces franchisors to be more forgiving of minor –even if verifiable– instances of breach by the franchisee. To be forgiving at the beginning is not too costly for the franchisor, given that he always retains the ability to terminate without any restriction, financial or otherwise, as soon as he observes that his benevolence has not been repaid with cooperative behaviour on the part of the franchisee. On the contrary, if the decision to terminate is legally constrained, the franchisor will terminate at the first occasion in which he can legally and costlessly –in terms of severance or compensation payment to the franchisee– do so. He will not forgive a first minor breach if there is sufficient evidence to show that termination would be deemed an acceptable punishment of franchisee’s breach. This would lead, then, to more terminations, rather than less, following legislation that makes termination more difficult and/or costly for the franchisor.

However the soundness of the specifics of this “second chance” theory, the underlying facts would seem to give additional support to the basic position that, on economic grounds, opposes compensation –or other restrictions to its exercise– after termination of long-term distribution contracts. If they actually increase termination levels by manufacturers, and unless compensation is sufficient to restore the fall in welfare of distributors due to the increased terminations (a doubtfully plausible assumption), those legal rules could be aptly characterized as self-defeating: They harm those they intended to protect and benefit.

From this section, thus, there seems to be a *prima facie* powerful case, both on theoretical and on empirical grounds, in favour of scepticism towards rules that impose compensation –in favour of the distributor, in practice, as data shows– after termination of a long-term distribution contracts. Let’s see if there are other considerations that allow a more optimistic view.

### 3.3. The relevance of specific investments

The economic theory of contracts, at its most abstract level, is concerned about designing mechanisms to ensure that in a voluntary economic interaction that protracts into the future (all that is required for an economist to label something as a contract) efficiency is achieved on two dimensions: trade, that is, the actual exchange of goods and services, so that they end up in the hands of whoever values them most; and investment, that is, committing resources in a productive way to enhance the surplus of the transaction, either by decreasing the costs of production, or by increasing the value of the goods and services<sup>53</sup>. The conditions and mechanisms to simultaneously induce efficient trade and efficient investment is the core of modern contract theory<sup>54</sup>.

---

<sup>52</sup> See, John BEALES III and Timothy MURIS, 1995, p. 169; Cándido PAZ-ARES, 2003, p.52.

<sup>53</sup> Good introductions to such topics, in Patrick SCHMITZ, 2001, p. 1; Benjamin HERMALIN, Avery KATZ, and Richard CRASWELL, in A. Mitchell POLINSKY and Steven SHAVELL (eds.), 2006.

<sup>54</sup> The problem is far less trivial that may look like from the outside, at least for a lawyer (who may just simply advise the parties: Hire a good lawyer and sign a detailed contract), given that uncertainty in the future can affect the value of performance, or the cost of production, or both, thus making trade possibly inefficient, and that the levels of investment are typically assumed to be non-contractible.



Contract investments, in the economic literature on incomplete contracting, are typically assumed to be, to a varying degree, but mostly to some extent, relation-specific. An investment is completely relation-specific if the value of the investment is entirely lost if trade between buyer and seller does not occur, that is, outside that particular contractual relationship. The lower the relation-specificity of the investment, the higher the external value of the resources invested. At the extreme, a completely non-specific investment produces the same value inside or outside the relationship.

### 3.3.1. The hold-up problem

The trouble<sup>55</sup> with relation-specific investments is that they are very vulnerable to hold-up by the other party in the relationship<sup>56</sup>. Let's think of the following example:

A supplier of components, S, in a contractual relationship with a manufacturer, M, is deciding whether to invest in modifying its plant to adapt it better to the particular production requirements of the manufacturer. Let's assume that the cost of the investment is 50, and it would reduce the supplier's production cost in 100, so the investment is efficient. The investment is completely relation-specific, which means that in serving another manufacturer the new plant is worth 0. If the contract terms between S and M can be renegotiated<sup>57</sup>, M has an incentive to hold S up by threatening a termination of the relationship if the agreed contract terms are not modified in its favour. For instance, M could force S to accept a reduction in overall price of 90. S would accept it, because it would be obtaining 10 over production cost, which is better than nothing. S would regret having made the investment, because it invested 50 to get a return of just 10 (M is "expropriating" the remaining 90 of the internal -to the relationship- value of the investment). Not surprisingly, the incentives to invest by S are significantly diluted by the likelihood of hold-up. In fact, as long as it cannot capture a fraction of the full value of the investment that allows cost recovery, it will not invest. If S has all the bargaining power in the renegotiation stage, it will efficiently invest, as it would reap the full value of the investment. In fact, as we will briefly see, economic theorists have tried to devise clever ways in which the investing party enjoys bargaining power when entering the renegotiation phase.

With a completely non-specific investment, however, such a hold-up situation would not arise, because the threat point of S in the renegotiation stage would have gone up by 100, because the investment has full value also serving another manufacturer. For intermediate degrees of specificity, the problem exists, albeit in a somewhat attenuated form compared to the complete specificity scenario.

If renegotiation can hardly be ruled out in a spot contract, the opportunities for renegotiation in a long-term contract, distribution or otherwise, are almost infinite. In fact, the number, complexity, and

---

<sup>55</sup> Specificity of investments may not be the result of technological or other constraints, but a conscious -strategic- choice by the contracting parties who, by increasing specificity, reduce the value of outside alternatives to the existing contract, and thus making the existing relationship more attractive in relative terms -at a price, however, because the value of the investment is lost when the relationship, for whatever reasons, comes to an end. This has been explored in the setting of cooperative investments (the following subsection provides a characterization of this kind of investment) as a solution to the shortcomings of monetary instruments to simultaneously reward the investor for making cooperative investment, and punishing him for not investing: See, Matthew ELLMAN, 2006, p. 234.

<sup>56</sup> A good starting reference for the hold-up problem is, Benjamin KLEIN, 1998. The problem was first spotted by Oliver WILLIAMSON, 1985, and shown formally by Oliver HART, and John MOORE, 1988, p. 755.

<sup>57</sup> Contract renegotiation is a mixed blessing in contract theory: it allows *ex post* efficient trade -the Coase Theorem guarantees that if renegotiation faces no transaction costs- but it also allows the possibility of hold-up, and thus reduces

unanticipated nature of many contingencies, almost naturally force the parties to a long-term contract to renegotiate the contract terms in the face of new information and new circumstances almost continuously. With this flexibility and ample room for renegotiation the chances of hold-up for the party making specific investments would rise, and thus the disincentive to invest to the efficient level. And some clever proposals at the interpretive level, such as an interpretation canon favoring the party making relation-specific investments<sup>58</sup>, cannot solve the problem in a general way.

### 3.3.2. Selfish vs. cooperative specific investments

Among specific investments, the literature in Economic theory dealing with incomplete contracts has distinguished two pure types of such investments. Selfish investments are the ones 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 it, it just decreases its production costs, without increasing the value of performance for the buyer.

Cooperative investments<sup>59</sup> are those that confer direct benefits to the other contracting party, and not to the party making the investment: If the buyer makes the investment, it just decreases 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, if they benefit both contracting parties, the investor and his partner, although we will disregard this complication in what follows<sup>60</sup>.

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 the 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 for placing external conditions on the renegotiation phase<sup>61</sup>, others for the use of options and appropriate strike prices for their exercise<sup>62</sup>. Other approaches rely in 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<sup>63</sup>.

---

the incentive to invest. On this issue, see Christine JOLLS, 1997, p. 203; Kevin DAVIES, 2006, p. 487.

<sup>58</sup> Scott BAKER and Kimberly KRAWIEC, 2006, building upon the idea of interpretive rules protective of one of the contracting parties developed by Omri BEN-SHAHAR, 2004, p. 389.

<sup>59</sup> The first treatment of cooperative investments is W. BENTLEY MACLEOD, and James MALCOMSON, 1993, p. 811. The standard treatment of these investments, in Yeon-Koo CHE and Tai-Yeong CHUNG, 1999, p. 84; Yeon-Koo CHE, and Donald HAUSCH, 1999, p. 125.

<sup>60</sup> In fact, few papers explore hybrid investments: Yeon-Koo CHE, and Donald HAUSCH, 1999, p. 125; Ilya SEGAL, and Michael WHINSTON, 2000, p. 603.

<sup>61</sup> Tai-Yeong CHUNG, 1991, p. 1031.

<sup>62</sup> Georg NÖLDEKE, and Klaus SCHMIDT, 1995, p. 163.

<sup>63</sup> Aaron EDLIN, and Stefan REICHELSTEIN, 1996, p. 478.

The second strand of the literature on selfish investments specifically deals with the 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<sup>64</sup>. He addresses two scenarios. The first one is the one in which the investing party is the one that can 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 overreliance lies in the fact that expectation damages fully insures the investing party against the possibility of losing the return on the investment, more than it 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 overinvestment. The reason is that to the full insurance motive to overrely (reliance damages fully insures the investing party because in all possible future states of the world, he obtains at least restitution of the cost of the investment) now it 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<sup>65</sup>.

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 the return on the specific investment less than what is optimal, and therefore the incentives to invest would be too low<sup>66</sup>.

SHAVELL'S analysis in the first scenario (the investing party is not the one 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<sup>67</sup>. 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 arrives. Then, any breach would come eventually from the investing

---

<sup>64</sup> Steven SHAVELL, 1980, p. 466.

<sup>65</sup> Randolph SLOOF, Edwin LEUVEN, Hessel OOSTERBEEK, and Joep SONNEMANS, 2003, p. 205.

<sup>66</sup> Steven SHAVELL, 1980, p. 484.

<sup>67</sup> William ROGERSON, 1984, p. 39.

party, who has efficient breach and investment incentives under the expectation damages remedy<sup>68</sup>. It is true, though, that using up-front payments can have problems of its 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, into one and the same hand.

The question is more complex still for cooperative (i. e. enhancing the other party's pay-off) 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<sup>69</sup>. In fact, the dismal result is that a contract is no better for the parties than no contract at all.

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, but always considering that the investing party is not the one that can take the breach-perform decision<sup>70</sup>. In the first case, with no renegotiation, expectation damages perform very poorly, because as a remedy for breach it induces zero cooperative investment. Reliance damages, in turn, do much better here: although at the price of some distortion –in the direction of excessive breach– in the breach-perform decision, they provide much better incentives for specific investments, and overall improves 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 defence 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<sup>71</sup>. 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 implausible –though not impossible– assumption.

### 3.3.3. Are the rules of the Commercial Agents Directive sound in economic terms?

What lessons can be drawn from the last two subsections on specific investments for the understanding and application of rules on compensation after termination, such as those in art. 17 of the Commercial Agents Directive?

The first and obvious one for those who have struggled with the previous pages is that the treatment of specific investments is very complex, and that the legal rules dealing with them have to take into account the economic intricacies that pervade the entire area, particularly when confronting investment

---

<sup>68</sup> Aaron Edlin, 1996, p. 98.

<sup>69</sup> Yeon-Koo CHE, and Donald HAUSCH, 1999, p. 125; Eric MASKIN, and John MOORE, 1999, p. 39.

<sup>70</sup> Yeon-Koo CHE, and Tai-Yeong CHUNG, 1999, p. 84.

<sup>71</sup> Urs SCHWEIZER, 2004.

in long-term contracting. No simple or all-encompassing rules can function properly across the whole range of factors and circumstances that have appeared in the preceding analysis. We would need a diverse set of rules to deal with the different factual scenarios which, in turn, have different implications in terms of the economic consequences of alternative legal instruments.

The second, and also quite obvious one, is that compensation rules after termination do affect the incentives of the contracting parties to undertake relation-specific investments: If, for instance, the distributor knows he will receive –let’s assume, full- compensation for the specific investments –let’s assume, of selfish nature- in cases in which the contract ends, for reasons such as death, illness or old age, or principal’s insolvency, he will invest excessively, compared to what would be optimal. In fact, when termination cannot be linked to behavior of the principal or manufacturer that could be described as breach of contractual duties –broadly understood, including opportunistic expropriation of value- compensation after termination cannot perform a desirable incentive role, because it cannot restrain breach of those duties by the principal, because there is none of it, whereas it induces excessive specific investment by the agent or distributor. This policy may be justified, perhaps, on grounds of social protection of a disadvantaged group<sup>72</sup>, but certainly not on economic grounds, because nothing is gained in efficiency terms by mandating compensation in those circumstances<sup>73</sup>.

Through the kind and level of compensation, if any, following termination of the contract, we can influence the behavior of the distributor with respect to reliance expenditures. The previous analysis on incentives to make specific investments under alternative damage rules is relevant for capturing the likely consequences of the options we may consider as policy alternatives for the compensation issue. This is clear in case of early termination of a distribution contract for a definite time period, which could be, naturally and easily, even undisputedly, be considered a breach of the distribution contract by the principal or the manufacturer<sup>74</sup>. It is probably also the case of termination by the “stronger” party in a contract of unlimited duration, given that in effect what we have here is a decision not to trade any more in the future, to borrow the terminology of incomplete contracting. It is true that termination of this sort can serve very different purposes, as we have already pointed out several times before, some of those being benign, even commendable, other being exploitative. The problem is that, due to the very nature of the long-term relationship in the distribution context, except in the most egregious cases

---

<sup>72</sup> I strongly doubt, however, that social protection can provide a solid basis for a general policy of compensating distributors in such cases. Distributors vary widely in many respects (size, wealth, future prospects) that are relevant for social protection policy so as to allow a one-size-fits all approach to the matter. Moreover, given that there is a contractual relationship between the parties, imposing social protection on one party, at the cost of the other –and not, the general taxpayer- may well lead to a worsening of contract terms in other respects: An increase in duties or rights that is not efficient, in the sense of increasing the surplus of the interaction, will imply a readjustment of the terms to the detriment of the beneficiary, the distributor in this case. This is the well known argument of passing-on of legal costs: Richard CRASWELL, 1991, p. 361; David WEISBACH, 2002.

<sup>73</sup> Later in the text I briefly examine whether risk-aversion by the agents and the preference for protection against undesirable events such as death or sickness can provide a rationale explaining the compensation rules that art. 17 Commercial Agents Directive envisions for such a set of contingencies.

<sup>74</sup> I am disregarding here, and in fact, in the entire treatment of compensation after termination, the possibility of termination by the distributor and possible compensation to the principal or manufacturer. This seems to be a much rarer occurrence, at least according to reported cases and the Spanish survey cited above, note 47. Moreover, although there well may be specific investments on the part of manufacturer and principal, and the opportunities for hold-up by the other party, legal systems seem to have addressed this scenario through general rules of contract Law, and not special compensation provisions on termination. Probably, overall it makes sense to handle them through the rules against duress and on contract modification. See, Oren BAR-GILL, and Omri BEN-SHAHAR, 2004, p. 391; Steven SHAVELL, 2005.

(of previous breach by the distributor, on the one side, or of hold-up by the manufacturer<sup>75</sup> on the other) Courts would be unable to tell one kind of termination from the other, so they are likely to end up being subject to the same legal solution in terms of compensation following the decision by the manufacturer to terminate.

What results from art. 17 of the Commercial Agents Directive under the light of the economic theory of specific investment and damage rules is, to a large extent, the impression of a confused picture. The two schemes, indemnity and compensation, make no mention of the distinction between selfish and cooperative investments, which is very natural (legal rules do not need to use terms of art of disciplines such as economics; moreover, the notions in themselves have neatly emerged in Economic theory only in the past ten years or so). The problem is that they point, confusingly, at the same time in similar and in opposite directions. They mention, in different manner, though, factors that point at cooperative aspects of the investment, and at selfish ones, to elements related to the features of expectation damages (loss of commission or remuneration), and to others related to reliance damages (unamortized costs and expenses).

Moreover, the indemnity scheme hints at the use, instead of a measure of harm (be it at the reliance or at the expectation measure) of some sort of disgorgement of benefits acquired by the principal as a result of the new customers –or the increased volume of existing customers- brought by the efforts of the agent<sup>76</sup>. To have compensation after termination based on restitution or disgorgement of gains to the principal, instead of on liability for harm caused is by no means a matter of indifference for the functioning of the compensation scheme. The alternative between gain-based and harm-based compensation is not exclusive to this setting: Violation of intellectual property rights, such as copyright or patent, violation of the right of publicity, provide also good examples of scenarios in which legal rules and Courts have to choose between both approaches to estimate the adequate remedy.

The choice has also been analyzed from an economic perspective, and in the economic theory of liability it is a well-established result that, if the legal system wants to give proper incentives through the imposition of liability, it is better to base it on harm to the other party –the victim- than to the gain accruing to the liable party. The reason is twofold. First, it is more efficient that the party who is the decision maker (who chooses between performance and breach, or between termination or continuation of the contract) faces the relevant curve –cost or harm function, in this case- of the other party, and not its own –the benefit or profit function, in this case<sup>77</sup>- because otherwise he would exploit to his own advantage the superior information –compared to the other party, or the external adjudicator- on his own function that he would typically possess. Given that in our setting it is the principal who takes the decision that may cause harm –termination- he should be facing the cost of the agent resulting from the decision, and not his own profit function.

---

<sup>75</sup> And for the former we can resort to ordinary law of contract breach (art. 18 a) of the Commercial Agents Directive excludes this case from the compensation schemes of art. 17), and for the latter, to the law of duress or the good faith requirements.

<sup>76</sup> See, for instance, Joanna GOYDER, 2005, p. 186, interpreting the provision as a rule of restitution of goodwill attributable to the agent's efforts.

<sup>77</sup> See, Donald WITTMAN, 1985, p. 173.

Second, gain-based liability, or disgorgement of profits, is much more vulnerable to mistakes made by Courts or adjudicators in the calculation of the relevant amounts. Even very minor mistakes can produce extremely undesirable outcomes, whereas with harm the effects are much more attenuated<sup>78</sup>.

In addition to these two general advantages of liability for harm vs. disgorgement of profit as instruments to induce efficient behaviour in a given target population, there are other reasons disfavouring gain-based compensation in the particular circumstances of termination of a commercial agency contract (reasons that apply to a long-term distribution contract as well).

First, the gain in new clients and increased volume of orders resulting from the effort invested by the agent –in marketing, sales and client management- is, to a large extent, uncertain and stochastic at the point in time of investing the effort. The agent, when making the specific investment, does not know for sure if the effort will pay or not: Marketing and sales effort may be very successful, or may fail, to a larger or smaller extent, both outcomes being influenced by factors beyond the control of the agent – and the principal too. A given effort may bring a very large number of new clients, or none at all, as a result of random or non-controllable events, such as general economic conditions, market situation, evolution of clients' tastes, and so on. The level of effort, on the contrary, is not subject to this uncertainty.

The former implies that in our setting, gain-based compensation is inherently more risky than harm-based compensation. So, if the agent is risk-averse, for an identical expected amount of compensation, he would prefer the more certain remedy, in this case, the one based on the certain harm (the invested effort, for instance). Moreover, the realization of the random events may be such that very few new clients, and thus few gains, are observed; or the opposite, that many new clients and gains are observed. If the litigated cases are biased, for whatever reasons, in favour of one or the other set, the level of compensation would be then too low, or too high compared to the ideal one.

Allow me now a brief excursus arising from this idea of the uncertainty associated with gain-based schemes. It is, in fact, the inherent riskiness of a gain-based compensation system what also makes it an unattractive safety net mechanism against circumstances that the agent may fear as threatening the continuation of the relationship, but are entirely independent of a voluntary decision by the principal. In case of termination due to death (art. 17.4 Commercial Agents Directive) or other circumstance –age, infirmity or illness- that reasonably prevents the agent from performing the activities (art. 18 (b) Commercial Agents Directive) indemnity or compensation is payable to the agent. It is true that the agent, if risk-averse, would like to be covered financially against the risk of those contingencies, which seriously and negatively affect the stream of income he receives. In short, he would like to be insured against such events. But awarding compensation based on the accrued clientele to the principal is an unattractive way to cover the risk of the agent. It is bad insurance, due to its intrinsic risk of the gains derived from the new clientele. And even the compensation option contained in art. 17 Commercial Agents Directive, as it also requires that the principal has obtained substantial benefits linked to the commercial agent's activities, entails some –although more limited- element of risk in the compensation for the agent, again an imperfection from the point of view of insurance. For such eventualities (death, sickness) agents would be better-off if instead of being compensated *ex post* as the Directive provides, they could get the exact expected amount of the compensation during the life of the contract, and

---

<sup>78</sup> See, A. Mitchell POLINSKY, and Steven SHAVELL, 1994, p. 427.

buying an insurance policy in the insurance market with that amount. This means that the “insurance” mandated by the Directive is inefficient, and parties would prefer to have a different insurance arrangement, a preference that the mandatory nature of the rules in the Directive makes impossible.

The second reason why a compensation system based on restitution of the gains from new clientele is an undesirable way to induce efficient specific investments is more subtle, and it is also independent of its uncertain nature. It has to do with the fact that we are dealing here with investments which are, at the same time, relation-specific and cooperative in the sense of benefiting the party not making the investment. It is clear that the increased stock of clients substantially benefits the principal<sup>79</sup>, thus making the investment to enhance this set a cooperative investment in nature. Moreover, if the investment is truly specific this means that it has no value at all outside the particular relationship in which the investment takes place. In the jargon of incomplete contract theory, the value of the investment is zero if there is no trade between the contracting parties.

In such a setting, a remedy in case of termination -such as restitution of profit from the new clientele- which is based upon the value of the investment after termination - a “no trade” situation- makes no sense, because the specific nature of the investment means that this outside value is zero. One would need to interpret, for it to make sense in a setting of cooperative specific investment, the gain-based indemnity not as the actual value -after termination- of the investment to enhance clientele, but as the *ex ante* expected value -in terms of revenue from increased clientele- of a given level of investment, let’s say, to put the gain-based remedy under the best possible light, the optimal level of investment. But this expected value of the optimal investment in increasing clientele is, for the purposes of the optimization problem of the agent -the one deciding on the investment- functionally equivalent to a form of liquidated damages clause, that is, a constant amount determined before the decision to invest is made and which is payable by the breaching -terminating- party. And in the setting of cooperative specific investments, CHE and CHUNG<sup>80</sup> have demonstrated that reliance damages outperform the best -for the parties- available liquidated damages clause agreed by the parties beforehand. So if the gain-based remedy in our setting works as a liquidated damages term, and it can be shown that reliance damages is preferable to the best liquidated damages clause, it must be the case that reliance damages are superior to gain-based restitution in an scenario of cooperative specific investment.

It is arguable, however, that the set of compensation schemes included in art. 17 Commercial Agents Directive are designed not for a setting of specific investment, but for one of general investment, that is, an investment whose value is portable, by the party benefiting from it, outside the particular relationship or contract in which the investment was made. One can think that this portability assumption is clear behind the gain-based indemnity in art. 17.2, provided that the Directive speaks of the fact that “...the principal continues to derive substantial benefits from the business with such customers,” - those brought by the agent- as a prerequisite for the indemnity. We are here squarely in a situation of portable or general cooperative investment. And even the compensation option in art. 17.3 may also implicitly be assuming this general character in the investment, as it also speaks -more neutrally and less explicitly, in temporal terms- of “...providing the principal with substantial benefits linked to the

---

<sup>79</sup> It is true that, depending on the way in which remuneration is determined -and art. 6 Commercial Agents Directive allows for multiple remuneration systems- the agent may also benefit -through increased commissions- from the new set of clients, but I disregard this complicating factor which would not substantially alter the qualitative nature of the reasoning.



*commercial agents activities,”* as one of the circumstances in which termination occurs and damage to the agent is presumed.

In other words, the principle lying behind the remedies in art. 17.2 Commercial Agents Directive may be linked more closely to a rationale of avoiding the more extreme consequences of unjust enrichment by the principal when the investment by the agent is of generic nature, than to a logic of reducing incentives for hold-up in detriment of the agent.

The fact that the principal can benefit from the new clients produced by the agent's investment does not automatically provide a sound basis for compensation in favour of the agent, however. If a cooperative investment (benefiting the non-investing party) is entirely generic, the investing party will not invest at all if he is not compensated for the cost of the investment, given that, contrary to what happens with the case of the cooperative specific investment, here the investment does not increase the likelihood of trade between the parties, that is, the chances of continuation of the existing relationship: if the investment is specific, the value of the contract, compared to outside alternatives, increases for the principal, and thus the likelihood that the latter will not terminate the relationship; if the investment is generic, its value is perfectly portable and thus the likelihood of termination is not positively affected by the investment. This means that, absent compensation, the agent will invest zero if the investment is cooperative and generic. And given that the principal anticipates this outcome, he will know that he needs to compensate the agent in order to obtain the positive value of the investment.

There is, then, no need to mandate compensation for such investments, given that it is in the interest of both parties that investment gets compensated. This is, for example, what happens in the employment relationship with covenants not to compete, which serve as implicit compensation –guarantee may be more precise here- for the employer to induce him to provide generic training to the employees, a theme that permeates the treatment of covenants not to compete that I present in the next section. In fact, if transaction costs are sufficiently close to zero, as one could imagine is the case between the principal and the agent during the relationship, the Coase Theorem ensures that parties would achieve the optimal level of cooperative generic investment, using the compensation scheme (explicit or implicit, prior or posterior to the investment, during or at the expiration of the relationship) that is more convenient for the parties. The legally mandated system would only come as a less efficient –less preferred- surrogate to the one that the parties would have chosen absent the legal imposition –because if there is mandatory compensation *ex post*, this would crowd-out the other mechanisms that the parties may have crafted: The principal is not going to pay twice for the cost of the cooperative generic investment.

The need for an *ex post* compensation system legally imposed upon the parties may arise if the agent is boundedly rational to the extent that he may misunderstand the nature or actual cost of the investment, or the reality of the agreed compensation system for the investment. But this is hard to assume as a good description of actual condition of agents, specially to assume it at the broad and general level of an abstract rule of private Law. A legal remedy may also be required in cases in which the principal reneges upon an implicit promise of compensating the agent for the investment. For instance, the parties may have a tacit understanding on a longer duration of the relationship, and increased commissions in that period, which may serve the agent to recover the cooperative investment made,

---

<sup>80</sup> Yeon-Koo CHE, and Tai-Yeong CHUNG, 1999, p. 96.

and then the principal abruptly terminates the relationship without honouring the implicit compensation scheme. Here, a legal remedy to curtail such opportunistic behaviour by the principal is *de rigueur*, but such a remedy does not coincide with the capped indemnity of art. 17.2 Commercial Agents Directive, nor with the complex and convoluted compensation provision of art. 17.3 Commercial Agents Directive. In such circumstances, the principal should simply compensate the agent for the unrecovered cost of the cooperative and generic investment.

All the previous observations strongly point in the direction that, when seen through the lenses of the economic perspective of incomplete contract theory, the compensation rules in the Commercial Agents Directive look seriously flawed, more in need of reform or reinterpretation than allowing application as general rules in the framework of long-term distribution chains.

This does not mean that it is never sensible to contemplate two –or more– compensation schemes as optional or default rules for parties to choose from, but the menu can hardly be the suspicious outcome of the EU law-maker paying respects to two divergent national legal traditions on the matter (the German and the French). One could refer to selfish specific (selfish generic investments pose no incentive problems) investments, and for these, expectation damages (that is, loss of profit, but for the agent, not the principal) perform better than reliance damages, particularly if coupled with a cap (like the one in art. 17.2 b), although the figure in this provision is arbitrary, which does not mean it is not sensible) to avoid the over-reliance effect of expectation damages over specific investments, when the investing party and the breaching party are different. The other could refer to cooperative investments, and for these a rule of compensating unamortized expenditures that do increase benefits to the principal –the notion of expenses incurred by the agent at the principal’s advice is just a proxy for this, not always precise: the principal can give advice the agent, or a distributor generally, on how to reduce own cost, due to more experience in the business) would seem a sensible solution. For cooperative generic investments, may be simply a rule that ordains compensation of full investment costs when the implicit –evidenced in any form satisfying the preponderance of the evidence burden of proof standard– understanding for recovery of such investments has been broken by the principal or the manufacturer.

As they now appear in the Directive, and in the various implementations in Member States’ legal systems, the two schemes significantly diverge from this model, inspired in the lessons from the economic theory of damages and specific investments. It is not only that the provisions in the Directive are confused and confusing, as I have argued before. A big divide between how the system looks like in theory, and how it works in practice, seems to exist. For instance, if one looks at the French system, in which the compensation scheme is based, in practice it amounts at payment of two years gross commission<sup>81</sup>, and has nothing to do either with substantial benefits for the principal, nor with amortized costs. In fact, two years commissions amount simply to average or standardized expectation damages, which are particularly counterproductive for cooperative relation-specific investments.

Given these shortcomings, one should be extremely hesitant to extend the application of the regime on compensation following termination contained in the Commercial Agents Directive, to other long-term contracts in distribution chains, even as a default regime, and not mandatory, as in the Directive is

---

<sup>81</sup> See, Report on the Application of Article 17 of Council Directive on the Co-ordination of the Laws of the Member States Relating to Self-Employed Commercial Agents (86/653/EEC), of 23.7.1996, p. 16.

foreseen for commercial agency contracts. Even if one wanted to pursue some explicit redistributionist policy in the field of distribution chains, one should be cautious about this extended application. Due to the contractual nature of the relationship between the parties, most if not all of the redistribution would be moot, due to readjustment in other terms that are not mandatory, while the incentives of the parties would have been inefficiently distorted. In sum, even if, as is the case, it is true that one cannot –and should not– radically rule out, on economic grounds, all compensation after termination in the area of distribution contracts, one would probably need to have a closer look at the relevant factors, which would not lead us to mimic the compensation schemes we now have in the *Acquis*.

### 3.4. Covenants not to compete

Non-compete agreements, or covenants not to compete, after the termination of the relationship, are contractual instruments that immediately affect the consequences of termination, and through them, can influence the contractual behavior of the contracting parties. Covenants not to compete are common not only in distribution contracts, but also in other types of contracts, namely in employment and in acquisition of companies.

Perhaps due to the deeply ingrained modern attitude towards any reminiscence of indentured servitude or slavery, and the corresponding fundamental principle of freedom to work, legislations and Courts all over Europe and all over the world are intuitively suspicious of covenants not to compete. We have already summarized the (relatively) restrictive regime, contained in art. 20 of the Commercial Agents Directive. The block exemption regulation also deals with non-compete obligations after the termination of the agreement. In order for the entire agreement to benefit of the block exemption, the non-compete clauses in the agreement are subject also to restrictive conditions: they should be limited to goods or services that compete with the contract goods and services; they should be limited to the land and premises from which the distributor has operated during the relationship; is indispensable to protect know-how transferred from the manufacturer to the distributor; does not exceed one year in duration.

From a theoretical perspective, covenants not to compete are typically seen as instruments in relation to investments in training and know-how that one party to the contract makes, and that the other party enjoys. In this respect, they are analyzed under the paradigm of what we labelled as cooperative investments in the past section.

In labor economics it is customary to distinguish between general and specific human capital, and accordingly, between general and specific on-the-job training. General on-the-job training is an investment in human capital that raises productivity in the current firm and in other firms equally. Completely specific on-the-job training is an investment in human capital that raises productivity only in the current firm. This distinction carries over neatly to the setting of long-term distribution contracts. Manufacturers can provide training, or know-how transfers, that can be equally beneficial to the distributor if in a relationship with a different manufacturer. This would be general training. Manufacturers, however, can provide training or know-how which is valuable only inside the existing distribution chain. It is not portable. This would be specific, and the investments to finance this acquisition of skills and know-how are also specific. How do covenants not to compete enter into the picture, and what is their role

### 3.4.1. Different forms of training and the role of covenants not to compete

In front of general training, the first question to be asked is: Who shall foot the bill of general training? From the economic theory of the employment contract, the answer is straightforward. Given its complete portability, and thus that the employee can enjoy the full return from the increase in productivity resulting from the investment, in the current employment or at other firms, it is the employee (the distributor, in our setting) who should bear the cost of general training. So the employee should pay the employer for such training, either directly –something that is seldom the case nowadays– or, more probably, indirectly, through a lower wage during a period of time allowing full repayment of training costs<sup>82</sup>.

The problem is that the implicit agreement on which the initial provision of training by the employer (manufacturer, in our setting) should be paid back, can collapse due to the employee's opportunism: Once the training period has ended, and given the perfect portability of the skills and know-how acquired, the employee would terminate the contract, move to another firm, and obtain a pay raise in the face of the increased productivity.

Covenants not to compete could then be seen as contractual means to enforce this implicit agreement to pay back investments in general training and know-how. Moreover, it may be a more effective enforcement mechanism than, say, a liquidated damages clause, due to problems of limited assets or personal bankruptcy of the employee<sup>83</sup>.

With perfectly specific training, the question is different. The employee cannot take with him the increased productivity resulting from the training. In addition, there may be a double hold-up situation, depending on who pays the bill of the training initially. If the employer pays for it, and expects the full return, he would like to pay the employer the opportunity cost of remaining in the firm (the best outside salary he could obtain). But then, the employee would be indifferent between remaining and leaving, and so he would credibly threaten to quit unless paid part of the value of the investment. If the employee pays for the training through lower wages (or directly), he expects to get the full value of the productivity increase. It is now, however, the employer, who has an incentive to engage in hold-up: because the employer is paying to the employee the full value of his services, he is indifferent between that –more productive– employee, and any other one, without the training, so he can credibly threaten with termination. In sum, no matter who makes the investment, there is always one party –the non-investor– who would credibly threaten to renege on the implicit promise sustaining the investment in specific training. The typical solution proposed by the labor economics literature to alleviate –not to entirely solve it– is that parties share the costs of this kind of 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<sup>84</sup>.

---

<sup>82</sup> See, Edward LAZEAR, 1998, p. 143.

<sup>83</sup> See, Paul RUBIN, and Peter SHEDD, 1981, p. 93.

<sup>84</sup> See, Edward LAZEAR, 1998, p. 152.

What happens in this setting of investment in training when there are chances that a new contracting partner may appear having a higher value for the services, or the performance of the employee (or the distributor)<sup>85</sup>?

The answer depends on the possibility of renegotiating the contract, and eventually, the covenant not to compete, when the third party bids for the services or performance of the employee. Without renegotiation<sup>86</sup>, if the investment in training the employee has been paid substantially by the employer, the incentive to invest in this case (by the employer, financing the training of the employee) is not distorted by a covenant not to compete, compared with other solutions –damages, no remedy at all. Moreover, in this case, there are no incentives to incur excessive investments to extract value and prevent entry by potential third parties, an issue that will be discussed in subsection 3.4.4.

The only problem with a covenant not to compete in this respect is that it may distort the decision to stay with the current partner, or to move, but other remedies have distortions as well, and it is inconclusive what the comparative advantages of the different alternatives would be<sup>87</sup>.

When the training is general, absent renegotiation, there would be a problem of underinvestment in such training under covenants not to compete, as long as the covenant does not cover all firms that may benefit from the general training. True, one could also extend the scope of the covenant, making it universal (covering the universe of all firms) but this would worsen the situation and welfare of the parties with respect to the decision to stay or to change partner.

### 3.4.2. The role of renegotiation in covenants not to compete

As in other aspects of incomplete contract theory, renegotiation also matters for the outcome of the parties' interaction when specific and general training are an issue. In the case of investments in specific training, there is a clear incentive to overinvest on the part of the employer (and the employee too, as we will see). Here, as long as there is a likelihood that the services or performance of the employee may be more valuable to a different contract partner, the parties to the original contract have an incentive to engage in excessive investment, if the contract provisions can be renegotiated *ex post*, in the presence of the contractual bid of the third party. The reason is straightforward: a higher investment level induces higher value extracted from the third party to obtain the "release" of the valuable employee, even if this implies a waste of resources because, by definition, the training is not worth when the employee moves to another partner. The incentive to overinvest is strategic: higher specific investments lead to a higher price that the third party has to pay to entice the employee to perform for him, and this additional revenue coming from the third party can be distributed by the parties to the original contract so that both are better-off. In sum, the employer and the employee have inefficient investment incentives in order to benefit at the expense of third parties. This feature, however, is not exclusive of covenants not to compete, although they are one of the most natural ways to obtain this result in termination of long-term contracts, such as employment or distribution. The same problem, it has to be noticed, would

<sup>85</sup> Typically, as in the incomplete contracts literature, this would be modeled as uncertainty regarding the valuation of the current firm and/or other firms.

<sup>86</sup> The same irrelevance result of exclusivity clauses (a covenant not to compete is part of the family of such clauses) can be proven even with renegotiation but when the third party gets no surplus from the interaction (maybe he is in a perfectly competitive industry): Ilya SEGAL, and Michael WHINSTON, 2000, p. 603.

<sup>87</sup> See, Eric POSNER, and George TRIANTIS, 2001.

affect other remedies protecting the investor, such as expectation and reliance damages, specific performance –in fact, covenants not to compete are an imperfect form of this, in this context- and liquidated damages or contractual penalties<sup>88</sup>.

In fact, this dismal result regarding investment in specific training and know how appears even without renegotiation, if the party who can terminate the contract is also the investor, when there is a potential bid from a third party: The higher the specific investments that the potential terminating (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 of any switching costs, in order to induce him to terminate<sup>89</sup>. So, specific investments, even without renegotiation, tend to be excessive when decided by the party who can also decide on termination<sup>90</sup>.

Things are not very different when dealing with investments in general training. Again the incentive to externalize the cost of the investment and extract value from third parties who may have a higher valuation of services or performance is present, both under covenants not to compete and with other remedies (damages, specific performance). It must be said, however, that covenants not to compete may be 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 benefit from the general training, and excluding from the enforceable scope other outside options for which such training is worthless. For instance, training may be industry- or sector-specific, so third parties belonging to the same sector would benefit from it, but not firms outside it. Reducing the scope of the covenant not to compete to just the industry, is efficient *ex post*, because guarantees enough incentives for investing in the training, but eliminates the incentive of excessive incentives made with the purpose of extracting rent even from those who value the employee but not the training<sup>91</sup>.

### 3.4.3. Empirical evidence: Silicon Valley vs. Route 128

Although there is no comparable empirical evidence with respect to the use of covenants not to compete, to the evidence available concerning legislation restricting termination of long-term distribution contracts, there is some empirical features which may have some bearing on this matter.

The major source of this suggestive evidence comes from the comparison between two relatively similar industrial districts: Silicon Valley in California, and Route 128 in Massachusetts.

---

<sup>88</sup> See, Eric POSNER, and George TRIANTIS, 2001.

<sup>89</sup> See, Tai-Yeong CHUNG, 1998, p. 1057.

<sup>90</sup> 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 the manufacturer, more realistically, in a distribution contract) who is more likely to terminate, the opposite result would be desirable, however.

<sup>91</sup> See, Eric POSNER, and George 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.

Different accounts of the relative performance of the two competing industrial districts underlined the advantage of Silicon Valley in terms of a culture of knowledge transfer, mobility across employers and start-ups, less vertical integration, and more dynamic environment, over Route 128<sup>92</sup>.

An explanation for these differences, and the relative superiority of the Californian district, has been found in the different attitudes of the relevant legal systems towards covenants not to compete. Whereas California Law prohibits post-contract covenants not to compete, the Law in Massachusetts consider them perfectly enforceable<sup>93</sup>.

Whatever the empirical robustness of this explanation, one implication seems reasonable from this suggestive evidence: Covenants not to compete can influence the competitive environment in a given economic sector. To this idea I devote the next subsection.

### 3.4.4. Covenants not to compete and other clauses as barriers to entry

It is well known in economic theory since the pioneering paper by AGHION and BOLTON, that contractual clauses can produce negative externalities concerning entry by potential new participants in a market, or potential new bidders for the services of existing firms. The logic of the argument is the same I have already mentioned with respect to covenants not to compete: An incumbent facing potential entry may try to deter such entry by writing long-term contracts with some of the existing buyers, and these may agree to entering into those -inefficiently- long relationships, because the contracts can be designed to transfer surplus from the entrants that finally penetrate the market, in favor of the incumbent, who then would share that surplus with the buyer<sup>94</sup>.

This same logic has been shown to apply, under certain circumstances, to other contract clauses, such as liquidated damages and penalty payments<sup>95</sup>, and covenants not to compete<sup>96</sup>. What lessons can be drawn from these for the treatment of covenants not to compete in our setting of long-term distribution contracts?

First of all, that the issue may be not as relevant in the distribution context as it is in the employment relationship and in contracts for the acquisition of enterprises. The reason for this relative unimportance lies in the fact that in our setting, it is typically the manufacturer who pays for the training of the distributor<sup>97</sup>, and at the same time is the one who typically exercises the termination decision. In this case, as has already been shown, the party possessing the choice concerning trade (termination or

---

<sup>92</sup> See, Anna Lee SAXENIAN, 1996.

<sup>93</sup> See, Ronald GILSON, 1999, p. 575.

<sup>94</sup> Philippe AGHION, and Patrick BOLTON, 1987, p. 388.

<sup>95</sup> Tai-Yeong CHUNG, 1992, p. 280; Kathryn SPIER, and Michael WHINSTON, 1995, p. 180.

<sup>96</sup> Eric POSNER, and George TRIANTIS, 2001. This does not preclude other functions, of course. In addition to what was said in subsection III.4.1 above, they can also serve as instruments to reduce outside options to one party, and thus increase the relative attractiveness of the existing relationship, and improving the incentives within the relationship: See, Matthew ELLMAN, 2006, p. 251. Janet BERCOVITZ, 2000, finds, for franchise contracts, a significant empirical correlation between covenants not to compete and opportunities for franchisee shirking in quality, thus giving some support to this view of the role of such covenants in long-term contracts.

<sup>97</sup> If the distributor had initially paid for the training, which benefits himself, we would outside the cooperative investment notion, and back in the general scenario of selfish specific investments in front of the possibility of termination (breach) by the other party.

continuation), with or without covenants not to compete, or the other clauses, has an incentive to overinvest, and thus to restrict entry and raise price to potential contracting parties for himself.

Essentially, then, the relevance of covenants not to compete arises (in the relatively small subset of cases) when it is the distributor who exercises termination to move to another distribution chain. In this scenario, covenants not to compete pose no particular problems when renegotiation *ex post* decision is, for whatever reason, infeasible or very costly. Here, as has already been mentioned, covenants not to compete cannot perform the function of restricting entry by potential third parties, given that there are no extra payments that can be extracted from the latter. It is true, however, that there can be distortions, not with respect to anti-competitive measures against entrants, but to the decision to terminate or not. In this respect, given the cooperative nature of the investment, reliance damages may perform well, although this remedy would require that the cost of training footed by the manufacturer is verifiable before the Court.

It is more natural to think, however, that in a manufacturer-distributor-third party interaction, transaction costs would be low enough to allow significant amounts of renegotiation. If this is the case, covenants not to compete pose the same anti-competitive threats that long-term agreements, or excessive penalties for termination, and should be carefully subject to scrutiny by Courts and Antitrust authorities, as potential sources of negative externalities to third parties. Particularly relevant for covenants not to compete, under this light, is the issue of the scope of enforceability of the covenant. It seems that a covenant that extends beyond the sector in which the manufacturer operates is not likely to respond to an efficient incentive purpose for investing in general training, but, *prima facie*, to a desire to extract surplus from new firms in other sectors willing to contract with the (former) distributor.

#### 4. Conclusions

As I have tried to argue in this paper, the issue of compensation following termination of a long-term distribution contract, is one of the crucial issues for the general legal understanding of such relationships, and is tightly interrelated with other contractual aspects of distribution chains. Unfortunately, from the point of view of European Law, the existing basis in the *Acquis* (essentially, art. 17 thru 20 of the Commercial Agents Directive, but also some isolated provisions in the Consumer Credit and Package Travel Directives) does not provide a solid and coherent starting point for a satisfactory legal treatment of this issue with respect to the entire field of distribution contracts. The contrast of those existing rules with the theoretical findings of the Law and Economics literature dealing with the matter does not offer a promising view of their likely consequences for the contractual behaviour of the parties.

Moreover, the available empirical evidence (which is now noticeable, if still not abundant), specifically concerning franchising, seems to cast another shadow of doubt on the merits of legislation imposing – and as a mandatory rule: Art. 19 Commercial Agents Directive- compensation after termination. This does not imply that the optimal legal regime is necessarily one in which no compensation at all is ever imposed as a remedy. The economic literature, and in particular the branch dealing with incomplete contracting, has identified several dimensions of long-term contracting, also in distribution chains, and may be with special relevance here, which play a role in this context. The nature and timing of the parties' interaction, the role of reputation to control opportunities for non-cooperative behaviour, the



incentives to make investments in the relationship, particularly relation-specific investments, all have a bearing on the desirability of the existence, scope, and nature of a compensation scheme favouring the distributor when the manufacturer has terminated the relationship. Translating this set of factors into abstract rules of Contract Law is, at least at this stage, too daunting a task to be undertaken with reasonable likelihood of success.

## 5. References

Philippe AGHION, and Patrick BOLTON (1987), "Contracts as Barriers to Entry", 77 *American Economic Review*.

Scott BAKER, and Kimberly KRAWIEC (2006), "Incomplete Contracts in a Complete Contracts World", Working Paper, University of North Carolina School of Law.

Oren BAR-GILL, and Omri BEN-SHAHAR (2004), "The Law of Duress and the Economics of Credible Threats", 33 *Journal of Legal Studies*.

BAUMBACH and HOPT (2006), *Handelsgesetzbuch Kommentar*, 32<sup>nd</sup> ed., Beck, Munich.

John BEALES III, and Timothy MURIS (1995), "The Foundations of Franchise Regulation: Issues and Evidence", 2 *Journal of Corporate Finance* (1995).

Omri BEN-SHAHAR (2004), "Agreeing to Disagree: Filling Gaps in Deliberately Incomplete Contracts", *Wisconsin Law Review*.

Janet BERCOVITZ (2000), "An Analysis of the Contract Provisions in Business-Format Franchise Agreements", Working Paper, Fuqua School of Business, Duke University.

Massimo C. BIANCA, and GRUNDMANN (eds.) (2002), *EU Sales Directive Commentary*, Intersentia, Antwerp-Oxford-New York.

Geert BOGAERT and Ulrich LOHMANN eds. (1993), *Commercial Agency and Distribution Agreements. Law and Practice in the Member States of the EC and the EFTA*, 2<sup>nd</sup> edition, Graham&Trotman/Martinus Nijhoff, London.

James BRICKLEY, Frederick DARK, and Michael WEISBACH (1991), "The Economic Effects of Franchise Termination Laws", 34 *Journal of Law & Economics*.

James BRICKLEY, Sanjog MISRA, and Lawrence VAN HORN (2006), "Contract Duration: Evidence from Franchising", 49 *Journal of Law & Economics*.

Martin BROWN, Armin FALK, and Ernest FEHR (2002), "Contractual Incompleteness and the Nature of Market Interactions", CEPR Discussion Paper.

Colin CAMERER (2003), *Behavioral Game Theory. Experimental Studies of Strategic Interaction*, Princeton University Press, Princeton (NJ).

Yeon-Koo CHE (1999), and Tai-Yeong CHUNG, "Contract Damages and Cooperative Investments", 30 *Rand Journal of Economics*.

Yeon-Koo CHE (1999), and Donald HAUSCH, "Cooperative Investments and the Value of Contracting", 89 *American Economic Review*.

Tai-Yeong CHUNG (1991), "Incomplete Contracts, Specific Investment, and Risk Sharing", 58 *Review of Economic Studies*.

--- (1992), "On the Social Optimality of Liquidated Damage Clauses: An Economic Analysis", 8 *Journal of Law, Economics & Organization*.

--- (1998), "Commitment through Specific Investment in Contractual Relationship", 31 *Canadian Journal of Economics*.

Richard CRASWELL (1991), 'Passing-On the Costs of Legal Rules: Efficiency and Distribution in Buyer-Seller Relationships' 43 *Stanford Law Review*.

Kevin DAVIES (2006), "The Demand for Immutable Contracts: Another Look at the Law and Economics of Contract Modifications", 81 *New York University Law Review* (2006).

Aaron EDLIN (1996), "Cadillac Contracts and Up-Front Payments: Efficient Investment Under Expectation Damages", 12 *Journal of Law, Economics & Organization*.

Aaron EDLIN, and Stefan REICHELSTEIN (1996), "Holdups, Standard Breach Remedies, and Optimal Investment", 86 *American Economic Review*.

Matthew ELLMAN (2006), "Specificity Revisited: The Role of Cross-Investments", 22 *Journal of Law, Economics and Organization*.

Robert GIBBONS (2001), "Trust in Social Structure: Coase and Hobbes Meet Repeated Games", in COOK (Ed.), *Trust in Society*, Russell Sage Foundation, New York.

Ronald GILSON (1999), "The Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants not to Compete", 74 *New York University Law Review*.

Joanna GOYDER (2005), *EU Distribution Law*, 4<sup>th</sup> edition, Hart Publishing, Oxford and Portland (OR)

Oliver HART, and John MOORE (1988), "Incomplete Contracts and Renegotiation", 56 *Econometrica*.

Benjamin HERMALIN, Avery KATZ, and Richard CRASWELL (2006), "Law and Economics of Contracts", in A. Mitchell POLINSKY and Steven SHAVELL (eds.), *The Handbook of Law and Economics*, North Holland (2006).

Christine JOLLS (1997), "Contracts as Bilateral Commitments: A New Perspective on Contract Modification", 26 *Journal of Legal Studies*.

Jonathan KLICK, Bruce KOBAYASHI, and Larry RIBSTEIN (2006), "Incomplete Contracts and Opportunism in Franchising Arrangements: The Role of Termination Clauses", Working Paper, George Mason University School of Law

Benjamin KLEIN (1995), "The Economics of Franchise Contracts", 2 *Journal of Corporate Finance*).

--- (1998), "Hold-Up Problem", in Peter NEWMAN (ed.) *The New Palgrave Dictionary of Economics and the Law*, Vol. II, MacMillan, New York-London.

Benjamin KLEIN, and Keith LEFFLER (1981), "The Role of Market Forces in Assuring Contractual Performance", 89 *Journal of Political Economy*.

Benjamin KLEIN and Kevin MURPHY (1988), "Vertical Restraints as Contract Enforcement Mechanisms", 31 *Journal of Law & Economics* (1988).

David KREPS (1990), "Corporate Culture and Economic Theory", in ALT and SHEPSLE (Eds.), *Perspectives in Positive Political Economy*, Cambridge University Press.

Edward LAZEAR (1998), *Personnel Economics for Managers*, John Wiley & Sons, New York.

W. Bentley MACLEOD, and James MALCOMSON (1993), "Investments, Holdup, and the Form of Market Contracts", 83 *American Economic Review*.

Fernando MARTÍNEZ, Montiano MONTEAGUDO, and Felipe PALAU (2000), *Comentario a la ley sobre contrato de agencia*, Cívitas, Madrid

Eric MASKIN, and John MOORE (1999), "Implementation and Renegotiation", 66 *Review of Economic Studies*.

George MATHEWSON and Ralph WINTER (1985), "The Economics of Franchise Contracts", 28 *Journal of Law & Economics*.

Georg NÖLDEKE, and Klaus SCHMIDT (1995), "Option Contracts and Renegotiation: A Solution to the Hold-Up Problem" 26 *Rand Journal of Economics*.

Cándido PAZ-ARES (1997), "La indemnización por clientela en el contrato de concesión", *La Ley* (1997), D-105 T. 2

--- (2003), "La terminación de los contratos de distribución", 8 *Advocatus*.

A. Mitchell POLINSKY, and Steven SHAVELL (1994), "Should Liability be Based on the Harm to the Victim or the Gain to the Injurer?", 10 *Journal of Law, Economics & Organization*.

Eric POSNER, and George TRIANTIS (2001), "Covenants not to Compete From an Incomplete Contract Perspective", University of Chicago Law School, John M. Olin Law & Economics Working Paper No. 137.

William ROGERSON (1984), "Efficient Reliance and Damage Measures for Breach of Contract", 15 *Rand Journal of Economics*.

Paul RUBIN, and Peter SHEDD (1981), "Human Capital and Covenants Not to Compete", 10 *Journal of Legal Studies*.

Paul SEABRIGHT (2004), *In the Company of Strangers. A Natural History of Economic Life*, Princeton University Press, Princeton (NJ).

Ilya SEGAL, and Michael WHINSTON (2000), "Exclusive Contracts and Protection of Investments", 31 *Rand Journal of Economics* (2000).

Anna Lee SAXENIAN (1996), *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*, Harvard University Press, Cambridge (MA).

Patrick SCHMITZ (2001), "The Hold-Up Problem and Incomplete Contracts: A Survey of Recent Topics in Contract Theory", 53 *Bulletin of Economic Research*.

Urs SCHWEIZER (2004), "Cooperative Investments Induced by Contract Law", *German Working Papers in Law & Economics*, Bepress.

Steven SHAVELL (1980), "Damage Measures for Breach of Contract", 11 *Bell Journal of Economics*.

--- (2005), "Contracts, Holdup and Legal Intervention" Working Paper, John M. Olin Center for Law, Economics and Business, Harvard Law School.

Randolph SLOOF, Edwin LEUVEN, Hessel OOSTERBEEK, and Joep SONNEMANS (2003), "An Experimental Comparison of Reliance Levels Under Alternative Breach Remedies", 34 *Rand Journal of Economics*.

Kathryn SPIER, and Michael WHINSTON (1995), "On the Efficiency of Privately Stipulated Damages for Breach of Contract: Entry Barriers, Reliance, and Renegotiation", 26 *Rand Journal of Economics*.

Robert TRIVERS (1971), "The Evolution of Reciprocal Altruism", 46 *Quarterly Review of Biology*

Joel WATSON (2002), *Introduction to Game Theory*, Norton.

Stephen WEATHERILL (2005), *EU Consumer Law and Policy*, Edward Elgar, Cheltenham.

David WEISBACH (2002), 'Taxes and Torts in the Redistribution of Income', University of Chicago Law School, John M. Olin Program in Law & Economics, Working Paper No. 148.

Oliver WILLIAMSON (1985), *The Economic Institutions of Capitalism*, The Free Press, New York

Donald WITTMAN (1985), "Should Compensation Be Based on Costs or Benefits?", 5 *International Review of Law & Economics*.

## 6. Decisions

*Supreme Court of Spain*

<i>Court Room and date</i>	<i>Ar.</i>	<i>Judge</i>	<i>Parties</i>
1ª, 27.10.2005	8043	José Almagro Nosete	"Genieser, SA" vs. "Adimed Andaluza, SA"
1ª, 21.11.2005	7677	Antonio Gullón Ballesteros	"Carratalá, SL, vs. "Egurko, S. COOP" and "Ortza, S. COOP"
4ª, 5.4.2004	3437	Gonzalo Moliner Tamborero	Mariano vs. "Ferraz Shawmut Ibérica, SA"
4ª, 21.1.2004	1727	Luis Ramón Martínez Garrido	Carlos Francisco vs. "Leche Celta, SA"
4ª, 2.7.2003	2004\18	Gonzalo Moliner Tamborero	Rafael vs. "Grosfillex Española, SA"
4ª, 21.3.2001	4106	Gonzalo Moliner Tamborero	José Manuel G.F. vs. "Banco Español de Crédito, SA"
4ª, 18.5.1998	4654	Arturo Fernández López	Confederación de Trabajadores Independientes de "Telefónica" vs. "Telefónica de España, SA"

*European Court of Justice*

<i>Date</i>	<i>Reference</i>	<i>Parties</i>
30.4.1998	C-215/97	Barbara Bellone v. Yokohama SpA
9.11.2000	C-381/98	Ingmar GB Ltd v. Eaton Leonard Technologies Inc
23.3.2006	C-465/04	Honyvem Informazioni Commerciali Srl v. Mariella De Zotti