Inter-Firm Contracts:
Evidence¹

Francine Lafontaine  
*Stephen M. Ross School of Business*  
*University of Michigan*  
*Ann Arbor, MI*  
*USA*  
*email: laf@umich.edu*

and

Margaret E. Slade  
*Department of Economics*  
*University of British Columbia*  
*Vancouver, BC*  
*Canada*  
*email: m.slade@mac.com*

April 2010

*Journal of Economic Literature* classification numbers: L22, L24

**Keywords:**

¹ This paper will appear as chapter 24 in the *Handbook of Organizational Economics*, Princeton University Press. We would like to thank the editors, Robert Gibbons and John Roberts, as well as several colleagues, including Kenneth Corts, Ricard Gil, Desmond Ho-Fu Lo, Julie Mortimer, Emmanuel Raynaud and participants in workshops at the NBER for helpful comments. We thank our respective institutions for their support. Margaret Slade also would like to acknowledge financial support from the Leverhulme Foundation.
1. INTRODUCTION

1.1 BACKGROUND
1.2 TYPES AND TERMS OF COMMON CONTRACTS

2. THE THEORIES

2.1 AGENCY MODELS
2.2 TRANSACTION COSTS
2.3 PROPERTY RIGHTS
2.4 VERTICAL RESTRAINTS
2.5 SELF ENFORCEMENT

3. METHODS OF ASSESSMENT

3.1 INCIDENCE
3.2 EFFECTS

4. THE EVIDENCE

4.1 PRICING CONTRACTS
4.2 CONTRACT DURATION
4.3 FLEXIBILITY AND ADJUSTMENT CLAUSES
4.4 FIXED PRICE VERSUS COST PLUS
4.5 SHARE CONTRACTS
4.6 VERTICAL RESTRAINTS
4.7 CONTROL RIGHTS ALLOCATION
4.8 RENT, REPUTATION, AND REPETITION
4.9 COMPLEMENTARITIES

5. CONCLUSION
1 Introduction

1.1 Background

Economics is sometimes described as the study of markets. Many market transactions, however, do not take place in arm’s length spot markets but instead are governed by long or short-term contracts. Since those contracts restrict the actions of one or both parties, there must be offsetting benefits. Otherwise parties would not voluntarily enter into agreements that limit their flexibility. In this chapter, we review empirical analyses of inter-firm contracts, paying particular attention to the reasons for entering into contractual relationships and the associated costs and benefits. We consider not only studies of the incidence of such contracts — when different types of contracts and contract terms are chosen — but also studies of the effects of those choices on outcomes such as profits, prices, sales, and firm survival. Since our goal is to review the empirical literature, we rely throughout on authors’ institutional knowledge, and their definition and classification scheme for contracts and contract terms.

Given that inter-firm contracts are found most extensively in the context of procurement and distribution, most of the contracts analyzed herein occur between firms that are vertically related. In particular, one party supplies either an input or a service to the other. Nevertheless, some contracts, such as licensing agreements, arise between firms in the same product market. Yet in the context of a licensing relationship, one firm, the licensor, in fact provides an input, a production technology and associated services, to the other, and in that sense the relationship is also vertical. With few exceptions, we do not discuss alliances and joint ventures, which are more common forms of relationships among firms in the same horizontal market. The interested reader is referred to Azoulay and Lerner, and to Ménard, chapters 13 and 26 in this volume, for a discussion of these types of relationships. Similarly, to keep the scope of the paper manageable, we focus on procurement and distribution contracts rather than contracts that occur between firms and their capital providers, be they banks or individuals. The latter tend to involve a different set of issues, and are discussed further in Gertner and Scharfstein, chapter 19 of this volume.

In our review of the empirical literature on inter-firm contracting, we emphasize analyses of agreements that are framed in the context of transaction costs and contract theory. However, we also consider contributions that predate those developments. Many studies that fall into the latter category are analyses of contract terms known as vertical restraints, where one link in a vertical chain constrains the activities of
another. Historically, such constraints have been viewed either as mechanisms by which firms can achieve, or at least move towards, the vertically integrated outcome without resorting to integration, and as devices that enable them to enhance their market power.\(^2\) In emphasizing the incentive aspects of vertical restraints, our coverage attempts to bring the findings from the vertical–restraints literature into the fold of organizational economics and the empirical analysis of inter–firm contracts.

Inter–firm contracts are often similar in structure to the contracts that are used within firms, including executive and certain types of labor–compensation schemes. Empirical studies of within–firm contracts, however, are also beyond the scope of this chapter. Interested readers should consult Azoulay and Lerner and Waldman, chapters 12 and 13 respectively, on these. A related literature that is beyond the scope of the present survey is that which focuses on the contractual relationships of firms with growers in various agricultural sectors, including the hog and broiler industries. Because growers are usually individual farmers, this literature tends to emphasize agent risk aversion and heterogeneity much in the same way as the within–firm contracting literature does.\(^3\) Similarly, and for similar reasons, we do consider contracts with talent, whether in the context of sports or movie production. Finally, we generally exclude studies of contracts where one party is the government. Both motivations and constraints can be quite different in such contexts.\(^4\) We focus instead on those transactions that parties have chosen to organize across firms.

The fact that we are considering agreements between firms already has implications as to the parties' incentives.\(^5\) In particular, in such contexts, the parties are usually residual claimants, although their profits are likely to depend on each other’s actions. Moreover, risk plays a less salient role in many of the models and empirical applications than it might if the transactions involved individuals.\(^6\) Similarly, many firms are, at least in theory, longer lived than individuals.\(^7\) This in turn implies that

\(^2\) For analyses of vertical integration, see Bresnahan and Levin, chapter 22 in this volume, and Lafontaine and Slade (2007). For a discussion of market–power motives for vertical restraints, see Lafontaine and Slade (2008).

\(^3\) Interested readers should consult e.g. Knoeber (1989), Allen and Lueck, (2002), Dubois and Vukina (2004), and Vukina, and Leegomachai (2006) for more on contracts used in these settings.

\(^4\) For more on government procurement contracts, however, see chapters 27 and 28 of this volume.

\(^5\) See e.g. Gibbons, 2005: 12.

\(^6\) Of course, risk is an important factor in analyses where one of the parties is a small firm, for example an owner-operator or a partnership.

\(^7\) In reality the failure rate of firms is quite high, such that their life expectancy is probably lower than that of individuals. However, what matters is that managers and owners make decisions as if the firm were long lived, either because they do not know when the firm might fail or because they internalize the effects of their decisions on the value of the firm's assets upon failure.
the “shadow of the future” probably looms larger in their relationships with other parties, including, of course, other firms.

1.2 Types and Terms of Common Contracts

In this subsection, we introduce some standard forms of contracts and then go on to discuss some related contracting issues. Contracts can take many forms. Nevertheless, certain standard formats or contract types appear in many industries, time periods, and regions of the world. We devote most of our chapter to those standard forms and practices, of which the following are typical.

Perhaps the simplest contract is a Pricing Contract that just specifies a price at which a good or service can be bought or sold. Such contracts can be very short lived, in which case they are not very different from spot–market interactions, or they can last for many decades. In the latter case, researchers have studied not only the contracts themselves but also their duration and the ways in which they can adapt to changed circumstances. Pricing contracts are common when the product is homogeneous, such as many fuel and non-fuel minerals, and in traditional franchising, such as gasoline and automobile sales.\(^8\) Finally, a special case of pricing contract involves only one item, such as a project, in which case it is referred to as a Fixed–Price Contract.

Cost–Plus Contracts are very different from pricing contracts in that they do not specify a price. Instead the seller is paid her costs plus a fraction of those costs to compensate for the effort involved. This type of contract is often used to procure one item, such as a project or service, that is often unique to a buyer. Moreover, a cost–plus contract is frequently contrasted with a fixed–price contract, which is the other standard method of financing a single project. The choice between the two involves a tradeoff between the flexibility of a cost–plus contract and the incentives for cost control that a fixed–price contract gives the supplier.

A Share Contract is a third standard type. This sort of contract is most often affine in revenues. Specifically, it tends to involve a fixed fee (which can be positive or negative) that is paid by the downstream firm, in addition to each party receiving a share of revenues. Share contracts are commonly found in business–format franchi-

---

\(^8\) Traditional franchising refers to vertical relationships where a product is produced upstream and sold downstream by a separate firm under a franchise contract, e.g., gasoline and autos. It can be distinguished from business–format franchising, which is discussed below under share contracts.
such as fast-food and real estate sales, but also in many other contexts, such as technology licensing and retail leasing. Since each party receives a share of revenues rather than profits, such contracts are distortionary. As a result, it was once thought that they were inefficient and would disappear in modern economies. It is now well understood, however, that share contracts, which are also associated with countervailing benefits, are here to stay.

Vertical Restraints (VR) are not types of contracts. Rather they are restrictions, such as exclusive dealing, tying, and resale price maintenance, that are included in contracts and that limit the activities of one or more parties to the contract. VR are most commonly found in retail settings, where they limit the downstream firm’s activities. Most studies of VR have emphasized their potential for increasing horizontal market power. We do not consider that part of the literature. Instead we survey the studies of VR that emphasize their incentive properties in a vertical context and how they can be used to reallocate decision rights. VR are thus special cases of the ways in which Control Rights are specified.

Most of the studies that we survey involve explicit contracts. This is to be expected since researchers typically analyze data on the terms and clauses that are written into contracts. Nevertheless, many ‘contracts’ are unwritten but well understood by the parties. Implicit Contracts are especially important in areas where the legal system is weak or corrupt, but they are also a common feature of modern economic systems.

Once a contract is agreed upon, Enforcement becomes an issue. Moreover, many of the same mechanisms can be used to enforce explicit and implicit contracts, and those mechanisms tend to be implicit. In particular, most derogations are dealt with informally, and legal breach–of–contract proceedings are relatively rare. At least two mechanisms are commonly used to sustain agreements implicitly. The first is relational, which means that both parties expect to benefit from a continuing one–on–one relationship. This might be accomplished by insuring that both expect to receive a stream of future rents. The second is reputational and involves group enforcement. In this case, failure to conform to accepted behavior will damage a party’s reputational capital more generally.

---

9 One can contrast business–format franchising, where no production takes place upstream, with traditional franchising. With the former, the upstream firm sells a way of doing business and the right to use a trademark, whereas the latter involves upstream production and an associated dealer network.
The chapter is organized as follows. In section 2, we very briefly review some theoretical arguments that have been used to explain the structure of inter-firm contracts. Our coverage of the theories is nontechnical. The interested reader can find a much more complete and rigorous treatment of that literature in Kornhauser and MacLeod, and in Malcomson, chapters 23 and 25 in this volume. In section 3, we review some econometric issues that arise in the empirical assessment of inter-firm contracts. We emphasize some of the pitfalls that researchers face and solutions that have been adopted. To keep the paper tractable, we do not revisit these issues when discussing the evidence. Our hope is that the reader will keep in mind that such issues affect the quality of the evidence in individual papers. We reserve the heart of the chapter for section 4, which contains a detailed review of the empirical literature on the incidence and effects of contract forms and terms, such as fixed fees, revenue shares, contract duration, and price-adjustment clauses. Concluding remarks are found in section 5.

2 The Theories

Given our description of the many forms that contracts between firms can assume, it is not surprising that no single model can capture all of the complexities. Nevertheless, much of the empirical literature aims to assess whether the predictions of just a few theories hold in the specific contexts from which their data are obtained. Since our objective in this section is to provide a framework that can be used to organize the empirical work, our discussion of the theories is short and non-technical. As mentioned previously, readers interested in more complete and advanced treatments are referred to Kornhauser and MacLeod, and to Malcomson, chapters 23 and 25 in this volume respectively.

Throughout our discussion of the theories we consider a vertical chain with at most three links: an upstream supplier, a manufacturer or franchisor, and a downstream retailer. We assume that the manufacturer is the principal who must decide how to interact with her agent, either her supplier or retailer. Explicit inter-firm contracts govern these relationships in all but the extreme cases of vertical integration, in which case there no longer are separate firms in the vertical chain, and pure arms’ length transactions in markets, which only require agreement on spot prices and quantities purchased.

Contracts are normally designed so that, \textit{ex ante}, it is in the parties’ best interest
to accept them. Nevertheless, contingencies can arise *ex post* that can cause one of the parties to want to break the contract. We begin with a discussion of explicit contracts that are legally enforceable, either by the courts or through some agreed upon arbitration scheme. Not all contracts are legally enforceable, however, either because the appropriate institutions do not exist or because reliance on third-party enforcement is prohibitively expensive. We therefore conclude this section with a discussion of contracts that are self enforcing *ex post* as well as *ex ante*.

### 2.1 Agency Models

Agency theory is the theoretical lens that has been most relied upon in the empirical literature to study the existence and the terms of one of the categories of contracts described above, namely share contracts. The reliance on this theoretical framework in the empirical literature on share contracts is due in part to the fact that very early models were developed to explain sharecropping. As such, they focused their attention on contracts that had the property of being linear in, or rather an affine function of, output, as is the case for sharecropping. Stiglitz (1974), for example, derived implications about factors that would affect the choice of share parameter ($\alpha$) in such a contract, implications that could be taken to data directly in a number of contexts where contracts take this form. In that sense, the theory “met the empiricists half way.”

In the pure risk-sharing version of his model, Stiglitz showed that increases in risk would lower the share of output ($\alpha$) that the more risk-averse party, assumed to be the agent, should be awarded under the optimal second–best contract. More typically, however, the agency or moral hazard model of contracting involves a risk and work averse agent who must exert an unobservable effort. In this case, Stiglitz (p. 244) showed that “If workers are risk averse, then $0 < \alpha < 1$, and $\alpha$ is larger the greater is the responsiveness of effort to an increase in the share $\alpha$. [...] If workers are risk neutral, $\alpha = 1$.” Thus early agency–theoretic models yielded testable implications concerning the effect of risk, risk aversion, and the importance of agent effort on the optimal second–best contract terms (i.e., the share parameter).

One drawback of this theoretical framework, however, is that the focus on single principal/agent pairs leads to the conclusion that contracts should be tailored to the characteristics of individual agents and local circumstances. In other words, the theory suggests that differences in agent risk aversion, local exposure to risk, and other characteristics of the agent or transaction should lead the principal to offer different
contract terms. Yet in many contexts, observed contracts are not tailored in these ways. Building on the work of Reid (1977), Rubin (1978), and Eswaran and Kotwal (1985), Bhattacharrya and Lafontaine (1995) argued that in many instances, it is not sufficient to give incentives to the agent. In the case of franchising, for example, the principal must also be given incentives to exert effort on, say, maintaining brand value. When this is true, there is moral hazard on both sides. Nevertheless, assuming that the principal is risk neutral, the second-best contract can still be implemented as an affine function of output. The share parameter, however, is now meant to give incentives to both parties. Bhattacharyya and Lafontaine (1995) show that, although traditional agency models give rise to different contracts for each principal/agent pair, under double-sided moral hazard and risk neutrality the share parameter is much more likely to be the same or very similar across agents and/or sets of circumstances. Moreover, the prediction from the traditional agency model that the share going to the agent should increase with the importance of the agent’s effort remains. This share, however, is now also decreasing in the importance of the principal’s effort.

In reality, of course, contract design is much more than just an issue of choosing a share parameter. In particular, effort is not the only factor that is costly; contract administration, for example, is associated with documentation and enforcement costs. An ideal contract would therefore provide appropriate effort incentives to each party, share the risk in an optimal fashion, and elicit an appropriate quality and quantity of the goods or services that are traded while at the same time keeping administration costs at minimum levels. In addition, unless the contract is very simple and short term, it should also provide efficient methods of adapting to unexpected changes in the economic environment. It is difficult to imagine that any single model could deal with this level of complexity. However, as noted by several authors, there are important complementarities among the various goals that the principal is trying to achieve. This suggests that we are likely to observe fixed packages of contract attributes. In other words, if there are $n$ attributes, each of which can be high or low, we do not observe $2^n$ contract types. Instead, we usually observe only a few contracts, which often can be positioned along a linear scale between two extremes. The choice of the share parameter, $\alpha$, in the theoretical model can then be viewed as a choice of location along such a scale. Table 1, which is adapted from Bajari and Tadelis (2001), illustrates this point. Assume that the agent receives compensation $s(x) = \alpha x + f$. In a situation of procurement, where $x$ is cost, $\alpha < 0$ is the share of costs that the agent pays, and $f > 0$ is a fixed price or payment that he receives.
In a retailing or sales situation, $x$ is output or sales revenue, $\alpha > 0$ is the agent’s or retailer’s share, and $f > 0$ is his fixed wage. Equivalently, in a franchise or licensing contract, $(1 - \alpha)$ and $-f$ are the royalty rate and fixed fee that the agent pays to the principal. With both procurement and retailing then, $|\alpha|$ represents the power of the agent’s incentives. In the procurement case, a higher share of cost borne by the agent gives him incentives to keep costs low. In the retailing or sales context, a higher share of sales revenues going to the agent, or a lower royalty rate paid to the principal, induces the agent to exert more effort towards increasing revenues. However, increasing the power of the agent’s incentives necessarily implies reducing the power of the principal’s incentives.

Table 1 also can be used to illustrate how changes in exogenous factors lead to changes in contract terms (e.g., to changes in $\alpha$). For example, it is usually assumed that the agent is the more risk–averse party, which means that when market or project risk increases, the need to insure the agent also rises and $|\alpha|$ should fall as a consequence. In addition, as the marginal product of one party’s effort rises, that party should be given a higher fraction of residual claims, which will cause $|\alpha|$ to rise or fall, depending on whether that party is the agent or principal. Finally, when the principal is responsible for the costs, it is easier to implement changes in product design, since a fully compensated agent is less likely to object. However, cost–plus contracts are more costly to administer, since the agent must document all expenses.

The simple classification scheme above moreover can encompass a situation in which the agent must perform multiple tasks (as in Holmstrom and Milgrom 1991) in a straightforward way. To illustrate, the agent might have to exert effort on quality as well as quantity production, and the former might be more difficult to measure than the latter. With a share contract of the form that we have been discussing, compensation tends to be based on (measurable) quantity. This means that as quality measurement difficulties increase, $|\alpha|$ should fall. A move towards lower powered

---

10 See e.g. Shepard, 1993, who places the three contract types in gasoline retailing in the U.S. along a scale with one standard contract, company owned stations, representing vertical integration and at the other extreme, open dealers, a form of trade between independent firms. She puts open dealers - which she equates with franchising - somewhere between these two extremes. Similarly, Slade (1998a and b) respectively describe the four standard contracts used in beer retailing in the UK and in gasoline retailing in the US respectively along a similar continuum. Finally, in the labor context, Ichmiowski, Shaw and Premlushi (1997) group human resource practices into four HR systems that they describe as different points along a scale from “most traditional” to “most innovative.”

11 Most of these comparative–static results are derived formally in Lafontaine and Slade (2007) for retailing and Bajari and Tadelis (2001) for procurement.
incentives occurs because high–powered incentives cause the agent to neglect the important quality dimension that is not emphasized by the compensation scheme.

The two extremes of our simple classification scheme, \( \alpha = 0 \) and \( |\alpha| = 1 \), also are worth noting as they represent limiting contracting forms. First, if \( \alpha = 0 \), the agent bears none of the costs of production in the procurement case as, for example, in a cost–plus contract. In a retailing context, the agent is paid a fixed salary only. This case can be equated with vertical integration to the extent that fixed salaries are found mostly within firms. The opposite extreme, with \( |\alpha| = 1 \), is equivalent to spot–market transactions, since the agent is the residual claimant with respect to his product or service. Furthermore, \( \alpha = -1 \) is also a linear or fixed–price contract where the buyer pays either the prevailing market price or one that is specified in the contract \((f)\). All of these situations are encompassed in table 1.

Finally, while our classification scheme focuses on the case of a single principal and agent, in empirical settings the principal may use a single contract with all or most of its agents. One can move from a theory where \( \alpha \in [0, 1] \) for a single agent to a group–level empirical analysis where the many agents of a principal operate under a single contract with \( \alpha \in \{0, 1\} \) if the optimal share parameter for each agent does not differ much. As mentioned above, Bhattacharyya and Lafontaine (1995) show that this is more likely under double than single-sided moral hazard. Moreover, the principal is more likely to use the same contract for all agents if the costs of administering different contract terms for different agents are high, either from an administrative perspective or because agents might be particularly concerned about opportunism when contracts differ. Finally, in many of the settings where contracts are uniform across agents, they are offered on a take–it–or–leave–it basis. Though this is again outside the simple model above, the types of agent that will accept the offered contract are likely to be somewhat similar, in terms of risk aversion and taste for effort for example, such that contract uniformity across agents again need not be costly for the principal.

### 2.2 Transaction Costs

Transaction costs (TC) are the costs of establishing and administering business relationships within and between firms or individuals. TC theories can be traced back to Coase (1937), who focused on the costs of transacting under different organiza-

---

12 Note that this is a linear contract for a single item.
13 See McAfee and Schwartz, 1994, for more on this.
14 See Lo, Ghosh and Lafontaine (2009) for more on this issue.
tional forms, particularly the costs of writing and enforcing contracts. The theories have been developed further by Williamson (1971, 1979, 1983), Klein, Crawford, and Alchian (1978), and others.

The insight from TCE that is most often addressed in the empirical literature is as follows. Parties to a transaction often make investments that have greater value inside than outside the relationship. In other words, the value of the assets in their intended use is higher than their value in alternative uses. Examples include specialized tools that can only be used to produce the products of one manufacturer, training that increases worker productivity exclusively in using those tools, and supplier facilities that have been located in close geographic proximity to purchasers. Specific investments give each party to a relationship a degree of monopoly or monopsony power. Indeed, even when there are many potential trading parties *ex ante*, when investments are specific, parties are locked in *ex post*.

When specific assets are involved, interaction in spot markets is unlikely. Instead, parties are expected to turn to long-term contacts or vertical integration to protect themselves and their assets. If those contracts were complete, specificity would not create problems. The complete contract would specify exactly what will occur and who will control the assets under all possible contingencies. However, writing complete contracts is costly, and not all contingencies can be foreseen. Thus real-world contracts are normally incomplete. Unfortunately, the combination of incompleteness and specificity gives the parties incentives to endeavor to capture the rents associated with the specific assets. This means that they are likely to haggle with one another, thereby increasing the costs of writing and administering the contract. They are also more likely to attempt to renegotiate the contract or, more generally, engage in opportunistic behavior. These possibilities, which are the essence of the hold-up problem, clearly pose problems for long-term contracting. Moreover, those problems are exacerbated in volatile environments, making it more likely that the firms will turn to vertical integration to protect their specific assets.

Here again, forms of contracts that lie between full vertical integration (or hierarchy) and spot market transactions, or hybrid organizational forms as they are often called in this literature (see Williamson, 1985, 1996), are viewed as intermediate solutions to the problem of minimizing haggling, opportunistic behavior, and exploitation. These organizational forms thus are apt to be relied upon when these problems are present yet not too severe. The theory moreover predicts that the contracts will be of longer duration when transaction costs (TC) are larger. However,
as contract duration increases, the probability that the economic environment will change grows. When TC are important and firms still choose to organize their activity via contract, the theory predicts that these longer duration contracts are more apt to incorporate flexibility. This can be done by, for example, including price and quantity adjustment clauses that specify how those variables can be altered, as well as provisions for efficient breach if the relationship becomes disadvantageous.

Finally, the theory provides a number of testable predictions concerning the circumstances under which TC are likely to be important. Specifically, transaction costs are apt to be more problematic when transactions are complex, when they involve specific investments, when those specific assets are more durable, when the quality of those assets is difficult to verify, when the environment is uncertain, and when the quasi rents that are generated by the relationship are large.\textsuperscript{15}

\textbf{2.3 Property Rights}

Property–rights (PR) theories, which are more recent and more formal than transaction–costs arguments, were developed by Grossman and Hart (1986), Hart and Moore (1990), Hart (1995) and others.\textsuperscript{16} Those theories emphasize how asset ownership affects investment incentives. More specifically, they demonstrate how the allocation of property rights, which confer the authority to make decisions concerning the use of assets when unforeseen contingencies arise, changes \textit{ex ante} investment incentives.

Because PR theories deal with relationship–specific assets, incomplete contracts, and \textit{ex post} bargaining,\textsuperscript{17} they are often thought to be closely related to TC models. However, there are important differences. In particular, unlike the TC literature, the PR literature has not focused on \textit{ex post} haggling, renegotiation, and opportunistic behavior. Instead, authors have developed formal models that have shown how costless \textit{ex post} bargaining affects \textit{ex ante} investment in non-contractible assets. Whinston (2003) moreover shows that the predictions from the two sets of theories can be quite distinct. On the other hand, if one equates the probability of vertical integration with the power of incentives in a contract, some of the comparative statics that are listed in table 1 in relation to our discussion of agency theory, emerge under PR models as well. In particular, Grossman and Hart (1986) show that as the importance of the manufacturer’s investment (or alternatively, decisions or effort) grows, manufacturer

\textsuperscript{15} For more on transaction costs, see Tadelis and Williamson, chapter 4 in this volume.
\textsuperscript{16} For an in-depth discussion of property–rights theories, see Segal and Whinston, chapter 3 in this volume.
\textsuperscript{17} PR theories can therefore also be traced back to Coase’s (1937) seminal contribution.
ownership (vertical integration or low–powered incentives for the supplier) becomes more likely, whereas as the importance of the supplier’s investment grows, supplier ownership (vertical separation or high–powered incentives for the supplier) is more apt to dominate. These predictions are consistent with those of double-sided moral hazard models in particular.\textsuperscript{18}

While the theoretical PR literature historically has focused on asset ownership as the mechanism that confers control or decision rights as well as payoff rights, contract clauses can reallocate control rights away from asset owners. And indeed, Baker, Gibbons and Murphy (2009) depart from the assumption of costless bargaining \textit{ex-post} found in the Grossman-Hart-Moore framework to analyze maladaptation during the contract execution phase.\textsuperscript{19} Their approach yields predictions concerning contract design and the allocation of decision rights separately from asset ownership. This theme, that contract clauses can allocate decision rights across fixed firm boundaries, is familiar also from an earlier empirical literature, on vertical restraints, to which we now turn.

\subsection*{2.4 Vertical Restraints}

We have thus far considered how contract design, which determines, for example, the power of the incentives that are given to each party as well as contract duration and flexibility, varies with the economic environment. Clauses that determine who can make important decisions, such as choose product prices and product lines, are also written into contracts. Any restriction that is imposed by one member of a vertical relationship on the other member of that relationship is a vertical restraint. In this subsection, we focus on the possible efficiency aspects of traditional price and non-price vertical restraints. The former refers to resale price maintenance, where, for example, a manufacturer either sets the price or sets a maximum or minimum price that retailers can charge, whereas the latter include exclusive dealing, exclusive territories, quantity forcing, and tying. In all cases, the presence of these clauses in a contract constrains the behavior of some party to the contract while keeping fixed the boundaries of the firms.

There is an important body of literature in economics that considers the effects of such restraints. This is because they have been viewed with some suspicion by the

\textsuperscript{18} For more on the similarities between property–rights and moral–hazard models, see Lafontaine and Slade (2007).

\textsuperscript{19} See also Baker et al. (2006) and Hart and Moore (2008).
antitrust authorities, raising concerns that they might be used to create or increase market power. The market power considerations for such restraints are beyond the scope of the present chapter. From an organizational economics perspective, however, our interest is in considering when firms want to adopt such restraints from an efficiency, not a market power perspective. We thus take the view that such restraints are clauses that principals write into their contracts with agents in order to align up and downstream incentives.

Empirically, vertical restraints most often arise in retail settings, with the upstream firm or manufacturer restricting its downstream retailers’ choices. For example, a manufacturer might limit its retailer’s product line or geographic market, or it might set the retail price. In describing the reasons why these restraints might be used, we therefore focus on retail rather than procurement contracts. The restraints that appear in retail contracts can alleviate many different types of incentive problems. We discuss a representative few that are by no means exhaustive.

First, the typical succession–of–monopoly problem arises when an upstream monopolist sells an input to a downstream firm at a price above marginal cost. If the downstream firm has market power, it is well known that it will choose a price that is higher, and a quantity that is lower, than the price and quantity that would maximize joint profits. The manufacturer can alleviate this problem in a number of ways, including setting the price, or requiring certain minimum quantities.

Second, manufacturers who invest in improving retail outlets, promoting retail products, or training outlet managers might worry that dealers will free ride on those investments. For example, dealers might encourage customers who visit their store to switch to a competing brand that has a lower price — thereby making the sale easier — or that has a higher retail margin — thereby making the sale privately more profitable. Exclusive dealing resolves this problem by making it impossible for the dealer to propose an alternative brand to customers. In such a context, exclusive dealing is a mechanism that enables manufacturers to protect their investments against potential dealer opportunism. Furthermore, in its absence, potentially profitable investments might not be undertaken.

Third, a dealer–incentive issue arises in situations where the manufacturer wants the dealer to invest \textit{ex ante} in specific facilities or human capital that would allow him to provide better service to consumers. As per the PR models above, unless the dealer can be assured that his investments are fully protected, he will choose to

\footnote{Interested readers are referred to Lafontaine and Slade (2008).}

\footnote{This, of course, is a form of the standard underinvestment problem analyzed in the PR literature.}
underinvest or not invest at all. Combined with a long contract duration, a vertical restraint such as an exclusive territory can provide the type of reassurance that the dealer needs.\footnote{For this solution to work, the upstream firm must be able to verify downstream investment and to terminate the contract if it is unsatisfactory. At the same time, the upstream firm must be able to commit not to terminate opportunistically.}

Finally, dealer services offered at the point of sale during the contract period can enhance the demand for a manufacturer’s or franchisor’s product. Retailers of course also benefit from providing such services. However, when there are multiple outlets in a retail chain, retailers do not fully internalize the benefit that is associated with their own decisions, as some of their satisfied customers will patronize other units of the same chain rather than returning to their unit in the future. In contrast, retailers bear the full cost of supplying the services. As a result, retailers are expected to provide a service level that is too low from the perspective of the upstream firm. In other words, in the context of retail chains, not only do dealers have incentives to free ride on the value of the brand, a vertical externality, they also have incentives to free ride on services offered by other dealers, a horizontal externality. Furthermore, the problem worsens as the fraction of repeat business that retailers face falls.

When this problem takes the form of a franchisee wanting to use lower quality inputs in the production process, it can be resolved with input–purchase requirements (tying) or approved–supplier programs as long as defection from such programs is not too difficult to detect.

When dealer service issues take other forms, Telser (1960) argued that minimum price restraints could solve the dealer service incentive problem by preventing retailers from competing on price and leading them to compete instead on quality or customer service. Klein and Murphy (1988) instead proposed that manufacturers could use vertical restraints such as minimum resale prices or exclusive territories to ensure that their dealers earn above normal returns, thereby creating rents that the dealers would lose if their contracts were terminated. Such rent, in combination with ongoing quality or service monitoring and the threat of termination, could entice dealers to provide desired levels of quality or service. In either case, since the quality and service levels in question are valued by customers — if it were otherwise manufacturers would not value them — quantities sold and hence consumer satisfaction should be enhanced.

This last argument, which states that upstream firms can use vertical restraints in their retail contracts as part of a mechanism that creates rent for retailers, rent that retailers can lose if they do not abide by the terms of their contract and are thus
terminated, was one of the early applications of the notion of self-enforcing contracts in the context of inter-firm contracts, a topic to which we now turn.

2.5 Self Enforcement

In many situations, parties to a contract cannot rely on formal enforcement of contractual terms. This can occur because formal institutions do not exist or, when they do, they are not effective, or because reliance on third-party enforcement is prohibitively costly. In such cases, parties must rely on informal enforcement. The notion of self-enforcing contracts, which was developed by Klein and Leffler (1981), Bull (1987), Klein and Murphy (1988), MacLeod and Malcomsom (1989) and others, builds upon the legal concept of relational contracting (see Macneil, 1978, and, more recently, Baker, Gibbons, and Murphy, (2002) and Levin, (2003)). Very generally, the idea is that the breaching party must face some future loss when breach occurs. This could occur, for example, when a contract that is advantageous to the breaching party is terminated or when that party’s reputational capital is damaged. Furthermore, the punishment can be inflicted by the party who is harmed in a bilateral relationship, in which case the contract is said to be relational, or enforcement can rely on group punishment, in which case other parties participate in sanctioning unacceptable behavior.

A relational contract is one that is sustained on the value of future interaction between the parties. In this literature, interaction is usually modeled as a repeated game with imperfect monitoring of the agent’s effort. Since the agent’s actions are unobservable, contracts cannot be written on effort. Moreover, since outcomes, while observed by the parties, are not verifiable, contracts on outcomes cannot be enforced by third parties. However, it is assumed that effort shifts the distribution of outcomes. In particular, high effort causes good outcomes to become more likely. In that setting, if the reward to good behavior is sufficiently high, the agent will eschew cheating in favor of high effort. In the context of inter-firm contracting, this implies that the agent must earn higher than competitive rewards. In other words, he is not taken down to his reservation value. Relational-contracting models are therefore similar to efficiency-wage models.

Although an important mechanism, relational considerations are not the only

---

23 For a recent survey of formal and informal enforcement, see MacLeod (2007).
24 See Telser (1980) for an early model of self enforcement in a repeated game context.
25 Akerlof and Yellen (1986) provide a collection of important contributions on efficiency wages.
means of sustaining contracts informally. Indeed, the notion of reputation as an asset that can be damaged by failure to live up to others’ expectations is very general. Furthermore, many reputational models rely on group punishments rather than one–on–one interactions as a mechanism to sustain cooperation (see, e.g., Klein and Leffler 1981). If group punishment is to be effective, however, others must be able to attribute blame when a relationship dissolves. This can be accomplished by invoking social norms, as in Okuno–Fugiwara and Postlewaite (1995), or by word–of–mouth communication, as in Grief (1989). Furthermore, published scores that are based on past performance, such as credit ratings, can play a useful communication role.

As we noted earlier, the notion of self enforcement has been used to explain a number of empirical regularities in inter–firm contracting. In particular, Klein (1980 and 1995) notes that explicit constraints on franchisee behavior are sustained by the value of the bilateral relationship, whereas implicit constraints on franchisor opportunism are constrained by group punishment. Similarly, Klein and Murphy (1988) argue that vertical restraints can play an important self–enforcement role. Specifically, a manufacturer who offers minimum resale prices or exclusive territories to dealers might do so to ensure that the latter earn rent. The presence of such rent in a manufacturer–dealer relationship creates something that the dealer will lose if he is caught misbehaving and is terminated as a consequence. The restraint thus ensures that the dealer will not go against the manufacturer’s stated service requirements.

3 Methods of Assessment

Our focus in this chapter is on empirical assessment of models of inter–firm contracts. Unfortunately, there are numerous pitfalls that researchers must face when attempting to tease out causal relationships rather than merely uncovering simple correlations, and some of the studies that we discuss rely on simple econometric methods that are not adequate for the task. We highlight some of the pitfalls before turning to a discussion of the empirical studies, organizing our discussion of potential problems around incidence and effects.

3.1 Incidence

Most of the literature on incidence — that is the literature where authors have tried to identify circumstances under which parties to a contract will choose to rely on one contract form or another, or on one contract term or another — uses what
amounts to a comparative-institutions approach. In other words, the papers consider how firms choose among a specific set of contractual alternatives. Examples include studies of franchising where the alternative for a franchisor is to integrate vertically (e.g., Brickley and Dark, 1987, Lafontaine, 1992), while the alternatives for a franchisee might be to operate as an independent business person (e.g. Williams, 1999, Mazzeo, 2004) or work as an employee. In other cases, the authors examine whether a particular contracting practice is relied upon — for example, one might consider whether contracts include take-or-pay provisions, as in Masten and Crocker (1985), or whether upstream firms grant exclusive territories, as in Brickley (1999). In these cases, the alternative is the absence of the contract clause of interest. Finally, in a few cases, contracting practices can be captured better by a continuous variable. Such is the case, for example, for contract duration (Joskow, 1987, Brickley et al., 2003), the share parameter in franchise contracts (Lafontaine, 1992, Lafontaine and Shaw, 1999), and for the proportion of franchised outlets in franchised chains (Brickley and Dark 1987, Lafontaine 1992, Lafontaine and Shaw, 2005).

When the dependent variable is continuous, authors have used OLS, or in some cases a limited-dependent-variable estimator such as a Tobit, to examine how the characteristics of the contracting parties and the transaction affect the continuous choice. In this literature, most authors ignore the potential endogeneity issue. Nevertheless, it can be a problem. To illustrate, certain types of investments, which can become characteristics in the regression model, might be undertaken only if the contract is of long enough duration, which could be the dependent variable. Thus a franchisee might not invest as much in specific assets, e.g. retrofitting a building to accommodate a particular style of restaurant, if her franchise contract is of short duration. This reverse causality would bias estimates in regressions of contract duration on investment levels (see e.g. Brickley et al., 2003 for more on this). The standard solution to the endogeneity problem is to use an instrumental-variables technique. Of course, this method is only viable if valid instruments can be found. Unfortunately, the problem of finding good instruments is particularly acute in these studies, as in most empirical studies in organizational economics, because the factors that lead firms to choose particular characteristics are also likely to affect desired contracting practices or terms.

When the set of contractual alternatives is limited to just two (or a few), the standard empirical study uses a discrete-choice model to relate the decision to use a

---

26 See notably Williamson (1991) on this.
contract term to the characteristics of the transaction and of the contracting parties. Methods for dealing with discrete dependent variables are well known. There are, however, a number of problems that are apt to surface in discrete-choice studies of contractual practices, problems whose solutions are more complex than when the dependent variable is continuous.

First, the ubiquitous endogeneity problem surfaces again. To illustrate, outlet characteristics are usually included among the explanatory variables that determine the method of transacting between manufacturer and retailer (e.g. Shepard, 1993, Slade, 1996, and Pinkse and Slade, 1998 for gasoline retailing). When an upstream firm decides to change the nature of the contract with the retailer, however, it might well decide to also change some of the outlet’s characteristics and vice versa. This could be the case with gasoline retailing, for example, where stations that are changed from full to self service also often increase their number of pumps and may be changed from independent to lessee dealer at the same time. In such circumstances, the direction of causality is particularly unclear. A standard method for overcoming the endogeneity problem in this context is to use two-stage least squares, where the probability that a transaction is organized in a given way is assessed under a linear-probability model. However, the linear-probability model has other undesirable features, including the fact that it is usually not possible to constrain the predicted probabilities to lie between 0 and 1. Other solutions to the problem of endogenous explanatory variables in the presence of limited dependent variables normally require strong assumptions (see, e.g., Wooldridge, 2002, pp. 472–477).

A related empirical problem arises from the fact that contract terms play a role in attracting particular contracting parties — that is, we see endogenous matching as well as selection. For example, if agents are heterogeneous with respect to risk aversion, some might simultaneously choose risky (safe) projects and contractual packages that are high (low) powered. If this selection problem is ignored, it can lead to the conclusion that when risk increases, agents are offered less insurance. In other words, the estimated coefficients will not only be biased but can also have the wrong sign. The bias can be eliminated if we include all relevant characteristics of contracting parties in the regression equation. Unfortunately, it is rarely possible to measure all relevant variables. In particular, risk aversion presents an often insurmountable problem. The solution that is suggested by Ackerberg and Botticini (2002) requires instruments that affect the matching process but not the contract choice.

\[27\text{ Of course, the problem of finding valid instruments is just as acute here as with continuous choice.}\]
Second, the errors in a discrete-choice model are likely to be spatially correlated in the sense that the off-diagonal entries in the variance/covariance matrix at a point in time are nonzero. For example, retail outlets that are located in a city center might experience common shocks that are not experienced by the retail outlets in the suburbs; or outlets that sell brands of a common manufacturer might have common private information. One possible remedy is to use the correction for spatial and time-series correlation of an unknown form developed in Pinkse, Slade, and Shen (2006) in a discrete-choice context.

Finally, a much broader problem with empirical studies of contract terms is that authors typically focus on just one component of the contract at a time. Unfortunately, contracts may be better described as sets of contract terms, where each choice interacts with the others. The issue of complementarity among contract terms has been mentioned in both the theoretical and empirical literatures, but the data requirements and empirical difficulties associated with correctly addressing those concerns have meant that only limited progress has been achieved on this front. We come back to this issue in section 4.9.

3.2 Effects

The empirical literature on contracting has focused more on incidence than on the consequences of contracting decisions. This may seem surprising given the interest in establishing the value of various contractual alternatives. Indeed, what matters at the end of the day is performance: is it beneficial for firms to rely on a given contract type, or do those that include a particular term in their contracts with their suppliers or retailers do better than those that do not? If so, should we suggest that the transaction be organized in this way? If not, why not? In other words, are there normative conclusions that can be drawn from analyses of inter-firm contracts?

Unfortunately, studies of the effects of contract terms on firm performance or other outcome variables, such as prices, sales, profits, growth, and survival, are relatively rare for a reason. First, studies of profitability or cost differences require detailed data that are typically proprietary. For that reason, much of the literature focuses on firm growth or survival, which may not be as related to performance as one would want. But more importantly, the endogeneity issue is particularly problematic in these studies. Simply put, the effects of various contractual decisions are difficult to identify empirically given that firms do not make contractual choices randomly. Instead, parties to a contract choose certain options based on what they expect will
give the best outcome in a given situation. This, of course, is exactly what the literature on incidence relies upon and tries to capture. Unfortunately, this also raises important issues when it comes to assessing the effects of contractual practices.

There are well-established techniques, such as Heckman’s (1978, 1979) selection and endogenous dummy variable models, as well as treatment-effects models, that can be used to deal with endogenous organizational–form decisions. But those techniques require valid instruments, and such instruments are particularly difficult to come by in the settings that we are concerned with. After all, firm and transaction characteristics are all to some extent the result of decisions made by managers, and in that sense all are endogenous. Authors thus often rely on various arguments to justify treating certain firm characteristics as exogenous, or at least predetermined, even in the incidence literature. When it comes to consequences, it is particularly difficult to argue that a variable that is postulated to affect the likelihood that a contract term is used will not also affect performance directly. To make matters worse, in the absence of non-suspect instruments, it is impossible to perform a formal assessment of the validity of any instrument. Still, we see promise in the increasing availability of various forms of spatial information about competitor and own outlet characteristics in the same or other markets as potential sources of instruments.

At the same time, due to the concern with endogenous selection, many studies of effects rely on external sources of variation in organizational form for identification. Specifically, they look for mandated changes in contracting practices — i.e. legislation — to assess effects. Some studies of this type use techniques that were developed in the natural–experiments literature, whereas others use an event–study approach.

In the former category, for example, researchers have used data on firms that operate in regions where legislators ban a particular practice or restraint (the treatment group), as well as data on firms that operate in regions where the practice is not banned (the control group). Since the ban comes from outside of the relationship, it is often assumed to be exogenous. But while the endogeneity problem is lessened in this situation compared to cases where the firms choose the contract terms, it is not eliminated. To illustrate, suppose that franchised chains do better when they do not face state termination restrictions, that is when they can terminate their franchise

---

28 See Wooldridge (2002, chapters 17 and 18) for a comprehensive discussion of econometric methods used for policy evaluation.

29 For a general discussion of how spatial data can be used to generate instruments in parametric and nonparametric contexts, see Pinske, Slade, and Brett (2001) and Pinkse and Slade (2010). For contracting applications in a parametric context, see Schneider (2009) and Kosova, Lafontaine and Perrigot (2009).
relationships at will. One might be tempted to conclude that the imposition of termination restrictions causes lower franchisor performance. However, it is also possible that termination restrictions are imposed exactly in those states where franchisors do poorly and thus have particularly strained relationships with their franchisees, leading franchisees to lobby for protection. Under those circumstances, the causality runs from poor performance to legislation.

If the underlying omitted factors that affect both the use of a particular form of contract and the performance variable of interest are time invariant — as might be the case for example for a manager’s innate ability or the difficulty of monitoring a particular task — the endogeneity problem can be overcome through the use of panel data. In particular, with panel data, one can use a fixed-effects estimator to remove the influence of time-invariant unobserved regional, brand, or outlet characteristics that cause the endogeneity problem. With this procedure, however, the effect of a contracting practice is identified solely through time-series variation. In other words, one is essentially assessing how changes in the use of a contracting practice lead to changes in performance. Unfortunately, there is often little time-series variation in organizational form. Furthermore, with panel data it is tempting to use lagged endogenous variables as instruments in the hopes that they are predetermined. This hope will be thwarted, however, if the errors are serially correlated, as is highly likely.

An alternative approach is to use a before-and-after estimation strategy — a time-series model. This approach requires access to data on firm performance that include periods before and after a legally mandated change such as the banning of a practice. The problem with this approach is that many things change over time, and, although it is tempting to attribute any significant performance change to the new legal requirement, this attribution might not be valid.

Some studies use a before-and-after approach on a cross section of firms when the data are not a panel. Those studies typically rely on data on stock prices to measure firm performance (i.e., event studies). Unfortunately, this reduces the set of firms whose actions can be examined empirically to publicly traded firms. This is problematic to the extent that the contracting practices of interest are used by relatively small firms, as in the case of franchise contracts.

Finally, recent work on within-firm incentives addresses the issues of selection and endogeneity by relying on data from field experiments (e.g. Lazear, 2000, Schearer, 2004, and Bandiera et al. 2007). Unfortunately, it is very unlikely that researchers can obtain experimental data that would be suited to the study of many questions.
involving how firms interact with one another. As best we can ascertain, no study of inter-firm contracts relies on such data.

Perhaps in part because experimental data are hard to come by, in the empirical literature in industrial economics authors increasingly use assumptions about utility functions, costs functions, and market equilibrium to generate estimating equations that allow them to identify underlying taste and technology parameters. These parameters are then used to perform counterfactual or “what if” analyses. In recent years, this methodology has been applied to the study of effects of contracting practices, in particular by Asker (2005) on exclusive dealing, Brenkers and Verboven (2006) on exclusive dealing and exclusive territories, Villas-Boas (2007) and Bonnet and Dubois (2009) on non-linear wholesale pricing, Mortimer (2008) on share contracts, Crawford and Yurukoglu (2009) on bundling, and Ho, Ho and Mortimer (2009) on full-line forcing. As with all structural analyses, the requirements of the models in terms of data and industry knowledge are very stringent, even more so in these vertical settings as the researcher needs to model horizontal games among upstream and among downstream firms as well as bargaining games among links in the vertical chain. Other problems with this approach include the fact that costs savings arising from better incentive alignment are difficult to capture in such models, and that evaluation of effects obtained from such models are valid only to the extent that the assumptions embedded in the model are valid. Despite these problems, we view this type of modelling as a fruitful area for future research on the effects of vertical contracting. Still, given the strong assumptions required to make these models tractable, we expect that reduced form analyses with strong identification strategies will continue to be an important source of insights.

4 The Evidence

In this section, we discuss the empirical evidence on inter-firm contracting, organizing this discussion around a series of themes related to contract types and contract terms or practices. We begin each subsection with a discussion of evidence concerning the incidence of a choice or practice, followed by a discussion of evidence on the consequences of the same choice or practice. In the latter case, we focus on consequences from the firms’ perspective, rather than from a public–policy or total–welfare perspective, as efficiency for the firm is the typical concern in the organizational economics literature. This is not to say that public–policy concerns are absent from
the empirical literature on inter–firm contracts, but simply that the goal of organizational economics is to uncover the benefits that the firm derives from using different contracting options.

4.1 Pricing Contracts

Many simple contracts specify a price at which a variable quantity of an input can be purchased. Perhaps because they are so common, or because they do not differ much from spot market transactions, linear pricing contracts tend to be studied less often than, for example, share contracts, at least relative to their occurrence. In this section, we discuss some work on incidence, much of which comes from traditional franchising. Note that in this context, the alternative to contracting that the franchisor considers tends to be vertical integration (rather than spot transactions). In that sense, much of the vast literature on the make–or–buy decision in procurement includes tests of simple pricing contracts when the input under the “buy” option is supplied under contract. The manufacturer’s decision, however, is treated simply as a buy in that literature, and thus the form and terms of the contract that ties the supplier to the manufacturer are typically not emphasized or even described in any detail. Exceptions include work on shipping and contracting in mineral markets, both of which we discuss below.

i) Incidence

A number of authors have examined the choice of contractual form in gasoline retailing. Often this involves the choice between transacting at a wholesale price (a pricing contract), which gives the station operator high–powered incentives, and vertical integration (VI), which normally involves salaried employment. Most studies estimate discrete–choice equations for organizational form (pricing contract v. VI). The object and focus of each study, however, are different. To illustrate, Shepard (1993) finds that higher–powered incentives (pricing contracts) are more apt to be relied upon when monitoring is difficult, as is the case, for example, for stations with repair facilities. Slade (1996), in contrast, casts her analysis in a multitask setting and finds that higher–powered contracts are less likely to be offered when tasks are complementary, where complementarities can occur in demand (as measured by cross–price elasticities) or in uncertainty (as measured by covariation in sources of risk).

See Bresnahan and Levin, chapter 22 in this volume, and Lafontaine and Slade, 2007, for recent reviews of the empirical literature on vertical integration.
Both of these studies lend support to moral–hazard models that emphasize incentive issues in choosing efficient contracts.

There is also evidence, however, that contract choice in gasoline retailing is influenced by competitive concerns. For example, Pinkse and Slade (1998) find that contract types (pricing contract v. VI) tend to cluster in geographic space, which is consistent with a model in which market–share motives, which are associated with similarity of offerings,\footnote{Since there is a strong relationship between contract type and station offerings, contract clustering implies attribute clustering.} outweigh market–power motives, which are associated with differentiation. Furthermore, Slade (1998b) finds evidence that supports strategic delegation of the pricing decision.\footnote{With company operation, prices are chosen by the principal or company, whereas with franchising, they are chosen by the agent or station operator.} In particular, she finds that delegation (in other words, a pricing contract) is chosen more often in contexts where competition among operators is less aggressive, a situation under which delegation is predicted to yield a greater increase in profits. It is important to recognize, however, that these findings, which imply that contracting practices are affected by the intensity of local competition, are more likely to surface in industries such as gasoline retailing, which tend to be less competitive than those in which business–format franchising is the norm.

Shipping is another industry in which linear pricing contracts are common. Unlike gasoline retailing, however, with shipping the choice is often contracting v. market transaction. In addition to specifying rates (or formulas for rates), shipping contracts can contain incentive provisions, such as penalties for delays and take–or–pay clauses. Researchers who have studied such contracts have paid particular attention to the role of market thickness in determining contract choice, under the hypothesis that, when markets are thin the advantages of contracting increase relative to spot–market transactions. Pirrong (1993), who assesses ocean bulk–shipping contracts, and Hubbard (2001), who examines long and short–haul trucking contracts, find support for this hypothesis. Furthermore, Hubbard finds that the thickness effect is strong for long but weak for short hauls, and explains this regularity by the fact that, when hauls are short, the costs of writing contracts exceed the benefits.

\textit{ii) Effects}

Several studies of consequences consider how retail prices vary with contract choice (e.g., company ownership versus franchising with linear or affine prices versus spot–market transactions). For example, Shepard (1993) compares gasoline prices in...
leasee–dealer (contract) and company units and finds evidence that, for some products, prices charged at leasee dealerships are higher. On the other hand, Hastings (2004) makes a similar comparison and finds no difference in price levels.

The above studies consider the effect of contract choice on retail prices, which are relatively easy to observe. Normally, it is more difficult to assess prices in long–term procurement contracts, since data on those prices are not usually available to the public. A comparison of contract and spot prices is possible, however, for nonferrous metals such as copper, lead, and zinc, since those commodities were traded in North America under two price systems. Indeed, spot and a long–term contract prices, which coexisted between the end of World War II and the late 1970s, are available from public data sources. The price upon which long–term contracts were based was known as the US producer price, whereas the spot price was the cash–settlement price from the London Metal Exchange. Both prices are transactions prices, since there was very little discounting off published prices. Nevertheless, although both prices were highly visible, differences between the two were often sizable and persistent. Not surprisingly, economists have compared the behavior of these two prices. Slade (1991), for example, has shown that, on average, there was little difference in the levels of these two prices but substantially greater variability in the spot price. Hubbard and Weiner (1989), for their part, find that the increased reliance on the spot price that occurred prior to the demise of the producer price led to faster adjustment of prices to supply and demand shocks. Note that prices in these industries are notoriously volatile. It therefore seems that firms entered into long–term supply contracts partially to reduce the amplitude of price fluctuations and to facilitate planning. However, as geographic markets became more integrated, and competition increased as a consequence, it became more difficult, and less desirable, to maintain a two–price system for such homogeneous commodities.

Wolak (1996) explores similar issues in electric–utility steam–coal markets and finds that, in spite of the fact that contract prices were systematically higher in that market, buyers entered into long–term contracts to insure against unforeseen supply interruptions and unwanted input–price variability.

iii) Government Intervention and Effects

In the above studies, the choice of contract type (VI v. linear or affine pricing contract v. arm’s length transaction) was made by the firms involved in the transaction. There have been many instances, however, where local governments have intervened
and prohibited certain types of contracts. Perhaps the most famous is the case of gasoline divorcement in the US, where divorcement means prohibition of company operation (VI) but not prohibition of ownership. In other words, firms were forced to transact using wholesale prices (to use a pricing contract) even with the branded stations that they owned. Divorcement laws, which have been passed by a number of US state legislatures, usually result from lobbying on the part of franchised dealers who claim that, when a company acts as both supplier and horizontal competitor, its behavior is influenced by considerations of foreclosure. The empirical literature (e.g., Barron and Umbeck (1984), Vita (2000), and Blass and Carlton (2001)), in contrast, shows that prices and costs rose and hours became shorter after oil companies were prevented from operating stations directly.

A different sort of divorcement is examined in Slade (1998a) — the forced move that occurred in the UK beer industry from franchising with two-part tariffs (contracting) to market interaction under linear prices. In the UK beer industry, similar to the US gasoline industry, tenanted pubs are owned by the brewer but operated by the publican, and the publican sets the price. Slade finds that draft beer prices rose after divestiture and attributes the rise to double marginalization — successive oligopoly markups — that is apt to occur after the removal of fixed fees.

The evidence thus indicates that government intervention that prohibits contractual arrangements involving certain modes of operation or ownership structures, including some that replace vertical integration with pricing contracts and contracts involving two-part tariffs by arm’s length transactions, is inefficient. This is so not only for the firms involved, who would have voluntarily chosen the new arrangements had they considered them profitable, but also for consumers.33

### 4.2 Contract Duration

The period over which a contract is binding can be very short or it can span many decades. It is therefore natural for researchers to be interested in uncovering the determinants and effects of the choice of duration. Transaction-cost theory is the main framework that has been relied upon to generate predictions concerning contract length. Specifically, the theory implies that contracts will be longer when firms have more specific investments at stake, since the need to protect those investments is greater. They will be shorter, in contrast, when environments are more uncertain, since flexibility assumes greater importance in that case. To our knowledge, however,

---

33 See also Lafontaine and Slade (2008) for more on this.
while this framework has generated studies of incidence, or put differently duration decisions, there are no empirical analyses of the consequences of such decisions for prices, quantities, or other performance measures.

In his seminal papers, Joskow (1985, 1987) relates the duration of contracts between US electric utilities and coal companies, which can last as little as one or as long as 50 years, to various proxies that capture the amount of relationship-specific investment, and thus quasi rent, involved. He finds that mine-mouth plants, namely plants that choose to locate next to specific coal mines with the expectation that they will obtain their coal from those mines — a classical case of site specificity — operate under much longer contracts than do other plants. Specifically, his baseline specification shows that mine-mouth contracts are on average 12 to 16 years longer. In addition, he finds that plants that use more coal, and those that operate in the East in contrast to the West or Midwest, use longer-term contracts. He argues that the former reflects the increased difficulty in finding alternative buyers or sellers for large quantities, whereas the latter reflects differences in the types of coal produced and in the production and transportation options available in the three regions. In sum, Joskow finds strong support for the hypothesis that differences in relationship-specific investments determine the duration of electric-utility/coal contracts.

Like Joskow, in their study of natural-gas sales contract, Crocker and Masten (1988) find that firms use longer-term contracts when they face a greater likelihood of hold up, for example, when they have fewer buyer, seller, or transportation options. However, the authors consider not only the benefit of long-term contracts in protecting specific assets, but also the cost of using longer-term contracts, which relates to the loss of flexibility in dealing with unforeseen events. Consistent with their hypothesis, they find that contract duration was reduced substantially by the 1973 oil embargo, which increased the amount of uncertainty in the market for natural gas. Saussier (2000) documents similar effects in his sample of coal procurement contracts for Électricité de France, even after endogenizing the degree of asset specificity in the transaction. Finally, Pirrong (1993) finds that contracts used in bulk shipping are of longer duration when markets are thin and carriers are specialized. All these findings support transaction-cost determinants of procurement-contract duration.

With business-format franchising contracts, which last about 15 years on average, duration varies substantially across firms both within and between sectors (see e.g. Lafontaine, 1992, Blair and Lafontaine, 2005). Brickley, et al. (2003) analyze the factors that affect the duration of franchise contracts and find that better-established
franchisors rely on longer-term contracts, as do those franchisors that require greater investment levels from their franchisees and those who face higher recontracting costs. They explain the result on franchisee investment using asset-specificity arguments, and interpret the franchisor-experience effect in terms of reduced uncertainty. In other words, their analyses yield results that are consistent with those obtained in the procurement literature.

4.3 Flexibility and Adjustment Clauses

With long-term contracts, it is crucial to incorporate flexibility. The sort of flexibility that can be built into contracts includes, among other things, adjustment clauses for price or quantity and clauses that make breach easier. Moreover, flexibility is related to duration in the sense that shorter contracts are in essence more flexible. Not surprisingly then, like empirical analyses of duration, studies of flexibility are mainly cast in a transaction-cost framework.

i) Incidence

Provisions for price adjustment in contracts can take many forms. Most schemes, however, can be classified as either redetermination or renegotiation mechanisms, where the former specifies a formula and the latter specifies a process. The choice between the two must take into account the tradeoff between flexibility, which favors renegotiation, and freedom from opportunism, which favors redetermination. As conditions are apt to change more during the life of a longer-term contract, transaction-cost theory predicts that such contracts will include terms that yield more flexibility. Crocker and Masten (1991) assess that choice in natural gas contracts and find that flexible adjustment (renegotiation) is indeed more apt to be chosen in longer duration contracts. However, conditional on contract length, they find no evidence that increases in quasi rents or market volatility affect the choice of adjustment mechanism.

In their study of contracts between producers and consumers of petroleum coke, Goldberg and Ericson (1987) found that over 90% of the contracts contained some form of adjustment mechanism. Moreover, those mechanisms ranged from price indexing based on crude-oil prices, to renegotiation when that price was above or below some limits, to negotiation at fixed periods. After 1973, however, when the volatility

---

34 For an early discussion, see Goldberg (1976).
35 This distinction is due to Crocker and Masten (1991).
of the market increased markedly, indexing clauses that were meant to be in force for the duration of the contract became less common as they were replaced by renegotiation clauses. These authors also found that, after 1973, the period between price changes fell substantially and termination became easier. This evidence also can be rationalized in terms of transaction costs.

Adjustment clauses are but one form of flexibility that can be built into contracts. Contracts also differ in the degree of detail, penalties, legal sanctions, and other specifics that they include, making some documents long and complex while others are short and simple. Murrell and Paun (2008) study contract complexity in agreements among Romanian firms and find that complexity increases (decreases) with seller (buyer) relationship–specific investment and with the quality of the legal system. Similarly, Lyons (1994) assesses the probability that UK engineering sub-contractors will formalize their relationships with customers rather than relying on more informal agreements. He finds evidence that pricing contracts tend to be more formal in cases where subcontractors are more vulnerable to customer opportunism, namely when customers account for larger shares of the engineering firm’s output and when output tends to be more specifically designed to the customer’s requirements or requires significant specific investment on the part of the engineering firm. These findings and those of Murrell and Paun (2008) are interpreted as supportive of modified versions of transaction–cost economics.

**ii) Effects**

A large segment of the literature on adjustment clauses attempts to distinguish between efficiency and market–power–enhancing effects of contract flexibility. The most favored nation (MFN) clause, which guarantees buyers (sellers) the lowest (highest) price that is offered to others in a region, is perhaps the most studied. Given the antitrust authorities’ stance on the anticompetitive nature of such contracts (see, e.g., Salop 1986 on the Ethyl case), it is not surprising that most authors have modeled the use of MFN clauses as practices that facilitate oligopolistic coordination. In particular, MFN clauses are expected to eliminate the possibility of selective price discounts and thus enhance cartel stability. Crocker and Lyon (1994), however, argue that MFN provisions facilitate efficient price adjustment in long–term contracts. In contrast to most research in this area, which is theoretical, they use data from natural gas contracts to distinguish empirically between competing explanations.\(^{36}\) After

\(^{36}\) See also Arbatskaya, Hvid, and Shaffer (2004), who, in a non-contractual context, use informal techniques and data obtained from newspaper advertisements to argue that low–price guarantees do
noting that market power and thus collusion opportunities reside with buyers in this market, Crocker and Lyon claim that the evidence in favor of efficiency rather than collusion is twofold. First, they find that the use of MFN becomes much more likely as the number of buyers increases, which they argue is inconsistent with the notion that MFN clauses facilitate buyer collusion. Second, they show that the nondiscrimination regions over which MFN clauses are defined are small and correspond more closely to sellers’ alternative market opportunities than to buyers’. Yet to be effective as practices that facilitate buyer collusion, MFN clauses would need to be applied to the set of competitors that buyers face rather than the set of competitors that sellers face. In addition, they note that MFN adoption patterns parallel those of clauses indexing gas prices to those of other fuels, evidence they argue further supports their efficiency argument.

The effect of take–or–pay provisions, which obligate buyers to pay for a contractually specified minimum quantity, called a take percentage, even when delivery is not taken, has also been studied. With these provisions, flexibility increases as take percentages fall, but protection of specific investments also declines, creating a tension between these two goals. Various explanations for the existence of take–or–pay provisions (e.g., risk sharing) have been proposed. Masten and Crocker (1985), however, argue that they provide an efficient means for contract breach. They test this hypothesis in natural–gas markets. Specifically, they explain take percentages as functions of buyer and seller numbers and find that take percentages fall (flexibility increases) when sellers are few and buyers are many. These results are consistent with their efficiency rationale for the provision as both of those factors raise the alternative value of gas reserves and make breach more desirable. Mulherin (1986) also argues that take–or–pay and MFN clauses are efficiency rather than market–power enhancing in natural–gas markets, and provides some empirical evidence consistent with the idea that the use of these clauses is related to bilateral contracting hazards.

A different, but related, question concerns just how flexible stipulated contract prices really are. To answer this question, Joskow (1988 and 1990) compared realized coal contract prices to market prices. He found that, since most coal contracts were indexed to cost factors, in periods of stable or predictable growth in demand, contract prices were relatively flexible to changing cost conditions, and thus contractual relationships did not break down. When demand turned down, however, the market price for coal was reduced and substantial deviations between market and contract prices

not facilitate collusion.
arose. In other words, the contract prices and associated pricing rules did not track changes in market conditions well. Nevertheless, in spite of unfavorable conditions for buyers, most long–term contracts remained in effect. In some cases, parties were able to renegotiate their contracts, relying either on scheduled re–opener provisions or changed quantity commitments. But this occurred because the contract or specific conditions permitted it. In general, the formal contract terms remained binding, that is, with clear contractual promises, litigation and breach were the exception, not the rule.

4.4 Fixed Price versus Cost Plus

Recall that, with a cost–plus contract, the seller is paid a percentage of costs, where the percentage exceeds 100. The fixed–price contract, in contrast, pays the seller a fixed amount that is determined \textit{ex ante}. Most empirical studies of cost–plus contracts involve neither a choice between contracting and vertical integration nor between contracting and spot–market transactions. Instead the choice is between two forms of contracts. Specifically, taking the existence of a contract for granted, it is the decision to rely on a particular type of contract — cost plus v. fixed price — that is assessed.

\textit{i) Incidence}

Since cost–plus contracts adjust automatically to changed circumstances affecting costs, they are more flexible, and one would expect them to prevail when i) projects are highly uncertain, ii) the technology is complex or untested, iii) quality is important but difficult to verify, and iv) trading parties trust each other. In contrast, since fixed–price contracts are associated with better incentives for cost control, they are more apt to be chosen when i) the project is fairly standard, ii) the measurement of costs is problematic, and iii) important changes in specification are not anticipated.

A number of researchers have examined whether these theoretical predictions hold up in practice. Leffler and Rucker (1991) assess private timber harvesting contracts, Banerjee and Duflo (2000) look at contracts for Indian customized software, Kalnins and Mayer (2004) and Shi and Susarla (2008) assess contracts for the provision of IT services and computer–related hardware, Corts and Singh (2004) look at contracts between oil exploration and production companies on the one hand, and independent drilling contractors on the other, and Bajari, McMillan and Tadelis (2008) consider private sector construction contracts. To summarize, they find evidence that cost–plus
contracts are preferred when presale measurement costs are high and monitoring is relatively cheap (Leffler and Rucker), when firms are older, where age is equated with reputation (Banerjee and Duflo), when costs are uncertain \textit{ex ante}, when the cost of measuring quality \textit{ex post} is high, and when the project does not involve the buyer’s hardware or proprietary technology (Kalnins and Mayer), and when development rather than exploratory wells are involved, since drilling activities for the former are more complex (Corts and Singh). Bajari \textit{et al.} (2008), moreover, find that more complex projects are more likely to be awarded via negotiation than through an auction mechanism, which they note amounts to saying that complex projects will be cost-plus rather than fixed price. Finally, Shi and Susarla (2008) find that vendors that can be trusted to negotiate fairly or to keep costs low are more likely to be awarded fixed-price or cost-plus contracts respectively. All of these predictions are supportive of the theoretical predictions.

Theoretical predictions concerning other effects, however, are more ambiguous. In particular, the issue of how previous experience with the same trading partner affects the choice of contract cannot be signed \textit{a priori}. Indeed, as Corts and Singh (2004) argue, whether repeated interaction makes fixed–price or cost–plus contracts more attractive depends on how such interaction affects incentive provision relative to contracting costs. On the one hand, experience with a partner can lower the need for high–powered incentives, and thus favor reliance on cost–plus contracts. On the other hand, it can lower contracting/recontracting costs, and thus favor fixed–price contracting. Like the \textit{a priori} predictions, the empirical findings concerning previous interaction are also somewhat mixed. Specifically, Banerjee and Duflo (2000) find that the choice of contract for software procurement is unaffected by whether the software firm has previously worked for a client, whereas Corts and Singh (2004), Kalnins and Mayer (2004) and Shi and Susarla (2008) find that frequent interaction leads firms to rely more on cost–plus contracting in offshore drilling and IT service procurement. The results of these studies thus suggest that, at least in these industries, repetition reduces the need for high-powered incentives more than contracting costs. We return to these issues in section 4.8.

\textsuperscript{37} See also Crocker and Reynolds (1993), who examine government procurement contracts, on this issue.
4.5 Share Contracts

With a share contract, which may or may not involve fixed fees, each party receives a portion of some output variable, usually revenues. Relative to their occurrence, share contracts have received a large amount of attention in the empirical literature. Much of this literature, however, has been about contracts that arise within firms (e.g. executive compensation) or contexts where at least one of the two parties is an individual (e.g. sharecropping or compensation of talent in the legal, real estate, or movie production industries). Still, other analyses have focused on inter-firm share contracts such as those used in business–format franchising and technology licensing. The main question addressed in the empirical work is the “why” of these contracts, or the question of incidence. However, interest in assessing consequences has grown more recently.

i) Incidence

Table 2 summarizes several studies that examine how firms choose to interact with other firms using a share contract versus some other option(s). The table shows the sector of the economy in which the firms operate, the author(s) of the study, the date of its publication, what the parties share under the share contract, the type of data or empirical technique used, the dependent variable for the study and, finally, the principal conclusions that the author(s) draw.

The research in table 2 first illustrates the different settings in which sharing among firms arises in the economy. In addition, the results from the studies support three main conclusions.

First, share contracting is often used in contexts in which incentive issues are important. Specifically, parties choose to share the outcome of their efforts when both of them need to cooperate in a form of team production and neither contribution is easily assessed by the other. This is one of the conclusions that authors who have examined franchising, licensing, and real–estate–leasing contracts in particular have reached.

Second, we find another type of explanation arising in settings where the value of the good being exchanged is unknown to both parties at the time of contracting, and depends on factors outside their control.38 This is the case, for example, for movie distributors selling copies of movies to video rental stores (Cachon and Larivi`ere, 2005, Mortimer, 2008). How popular the movie will turn out to be, and thus the level

38 See Goldberg 1976) for an early statement of this argument.
of revenues that the video rental store can earn from stocking copies of the video, is unknown at the time of contracting. If the video rental store must purchase all copies of the video at a fixed price ex ante, it will be wary of buying too many copies. Just as auction participants underbid if they are concerned about the value of the good that they are buying (i.e. underbidding is the best response to the “winner’s curse”), the downstream firm, worried about over-evaluating demand for a video will “underbid,” but in this case underbidding will take the form of choosing a lower quantity, i.e. purchasing fewer videos. From the distributor’s perspective, this increases the chances of stock-outs downstream, and thus hurts the revenue stream of the movie distributor as well as that of the video rental store. Specifying the price of the video as a function of the revenues it brings in ex post, which is exactly what revenue sharing achieves, can thus serve as a flexible pricing device that prevents “under buying”. Moreover, it reduces the need for ex ante search and for renegotiation ex post (Leffler and Rucker, 1991, Lafontaine and Masten, 2002, Gil and Lafontaine, 2010.)

Third, and finally, the findings of these studies are important in what they do not support, which is the notion that risk sharing is an important factor explaining the use of share contracting. In fact, authors have tended to find a positive relationship between sharing and risk in contexts where alternative contracts would insure the risk averse party much better. This relationship is inconsistent with risk sharing but consistent with the notion that uncertainty exacerbates monitoring problems and induces different levels of delegation, which in turn lead to more sharing (Lafontaine, 1992, Lafontaine and Bhattacharyya, 1995, Prendergast, 2002).

A much smaller group of studies considers how share parameters vary across contracts. These include Lafontaine (1992, 1993), Lafontaine and Shaw (1999), and Brickley (2002) on franchising, and Wheaton (2000) and Gould, Pashigian and Prendergast (2005) on retail–lease contract terms. In general, those authors find results that are consistent with the findings from the empirical literature on the choice of organizational form (i.e. the decision to use sharing or not). Lafontaine (1992) for example finds that the right–hand–side variables used to capture double–sided moral hazard issues play the same role in the setting of royalty rates (share parameters) and franchise fees as they do in the choice of the extent of franchising in a chain. However, the variables in question explain much less of the variance in royalty rates and franchise fees than they do in the proportion of outlets that are franchised.40

---

39 The argument thus also relates to the measurement–cost argument of Barzel (1982), and to the self–enforcement argument of Kenney and Klein (1983, 2000) for the use of block booking.

40 One factor that explains some of the variance in royalty rates is the franchisor’s reliance, in
Similarly, for licensing contracts, Taylor and Silbertson (1973), Contractor (1981), and Caves, Crookell, and Killing (1983) find that the variation in fixed fees and royalty rates across industries is difficult to explain. Gil and Lafontaine (2010), however, show that, in the context of movie distribution, the shares to distributors vary in predictable ways when viewed as a mechanism to extract downstream value. Specifically, distributor shares are higher (and decline more slowly) for movies that are expected to do better at the box office, for movies that face less competition from other movies released simultaneously, and for movies shown in theaters whose owners have little local market power.

Finally, and not surprisingly, state laws that protect franchisees in business-format franchising have been found to affect the use of franchising, as well as the terms of these contracts. Specifically, Brickley, Dark, and Weisbach (1991) find that franchising is used less relative to company ownership in states that have enacted franchise-termination laws, and that the value of franchised companies operating in a state is negatively affected by the passage of such a law. Moreover Brickley (2002) shows that franchisors headquartered in states with termination laws — namely laws that make franchisee termination more costly — charge higher royalty rates and lower franchise fees, such that the prices paid for franchises by franchisees are higher in those states. These results are all consistent with the notion that franchisors value termination rights, a result that suggests that franchisee behavior in business-format franchising is controlled at least in part via the threat of losing future benefits from the relationship. We come back to these issues briefly below.

In sum, the literature on incidence has shown that share contracting between firms tends to occur most often when incentive problems arise for both parties to a contract. In addition, the empirical literature suggests that in some contexts sharing is a response to the difficulties of “setting the right price” and minimizing renegotiation costs in markets where transactions are heterogeneous and values are uncertain a priori.

\textit{\textbf{ii) Effects}}

A typical study of consequences considers the effects of sharing on firm outcomes relative to what would have happened under another organizational form, for example, some sectors, on input sales as an alternative form of profit extraction. Lafontaine (1992) finds that royalty rates are lower, for example, in chains that sell more to their franchisees. Rao and Srinivasan (1995) provide related evidence that royalty rates are lower for franchisors in retail compared to those that sell services.
a comparison of firm profitability, service quality, or survival between company and franchised units of a chain, or for chains that franchise versus those that do not. Shelton’s (1967) analysis is a classic in this respect. He uses data on costs, revenues, and profits for outlets in a single chain to examine the effect of switching from franchising to company ownership and from company ownership back to franchising. He finds no difference in revenues across the two governance regimes. However, under company ownership, costs are higher, and thus profits are lower, than under franchising.

The main advantage of Shelton’s study is that its within-outlet design holds most things constant as the mode of organization changes. Its main drawback, however, is that units in this chain are operated under company ownership only during periods of transition. In other words, franchising is the preferred mode, and company ownership only a transitory phase. Consequently, company ownership is likely to be inefficient, and inefficiently implemented, in this particular chain. This, then, might explain Shelton’s findings.

A number of other authors have looked for price or cost differences between franchised and company units of chains when the firm chooses which outlets are franchised and which are operated by the company. In particular, Krueger (1991) found that company employees were paid slightly more and faced somewhat steeper earnings profiles than employees in franchised units. He argued that the lower–powered incentive contracts of the managers of company restaurants make it necessary to offer greater incentives to employees in the form of efficiency wages and steeper earnings profiles. Kosova, Lafontaine and Perrigot (2010) compare revenues, occupancy rates and prices among the hotels of a large multi-chain company. They find significant differences between these outcomes for franchised and non-franchised hotels in aggregate data patterns and in analyses that treat organizational form as exogenous. However, the differences are small even in aggregate data, and when they use information about the company’s other local operations to instrument for organizational form at a given hotel, the differences between franchised and company operations become statistically and economically insignificant.41

Some studies have looked also for quality differences between franchised and company units of the same chain. Bradach (1998: 109), in particular, interviewed managers in five fast-food chains, and concluded that the two arrangements exhibited

41 Arruñada, Vásquez and Zanarone (2008) in contrast find that company-owned car dealerships in Spain are much less productive than franchised car dealerships. However, they explain these performance differences not based on organizational form per se, but rather based on the pro-labor legal environment, which affected the terms of labor contracts more in vertically integrated dealerships, making them particularly inefficient.
similar levels of (standard adherence) uniformity. For the two firms in his sample that used third-party evaluators to assess quality, the average score was 94.6 (out of 100 points) for the franchised units and 93.9 for the company units in the first chain, and 89.7 and 90.6, respectively, for the other. He concluded that there was no quality difference between franchised and company-owned restaurants within these chains. Using data on quality ratings published by Consumer Reports, Michael (2000) found that quality was negatively associated with franchising in both the restaurant and hotel industries, and concluded that free-riding was a problem for franchised chains. Jin and Leslie (2008) also found evidence that hygiene scores (a measure of quality) were higher among company-owned restaurants than among the franchised units of the same chains in their data. A new policy requiring that restaurants post their hygiene scores, however, eliminated this difference.

Whether franchising affects firm survival is a question that has received some attention in the literature as well. Most of the studies have been concerned with whether affiliation with a franchised chain augments the likelihood of survival for an entrepreneur. For example, Bates (1995a, 1995b) used the Characteristics of Business Owners (CBO) database produced in 1992 by the U.S. Census Bureau to assess the rate of failure among a representative sample of small businesses, both franchised and non-franchised. He found that failure rates of franchised small businesses were greater than those of independent businesses, though not significantly so. Specifically, he observed that over a five-year period, 34.7 percent of franchised businesses failed as opposed to 28.0 percent for independent. Bates (1998) then further distinguished units sold to new franchisees and those sold to existing franchisees. He found that the vast majority (84 percent) of new franchised units are opened by existing multi-unit operators, and that these units were very likely to survive, much more so than independent businesses. New units opened by new franchisees, however, were less likely to survive than independents.

Early empirical evidence, however, had suggested that franchising attracts people who would not have chosen to open a business by themselves (Hunt, 1972 and Stanworth, 1977). Consistent with these findings, Williams (1999) used the same CBO database to document differences in human capital, such as formal training, business experience, and so on, between individuals who choose to purchase a franchise and those who go in business for themselves. He found that those who opt for franchising tend to have higher education and work experience, but lower levels of business experience. He then showed evidence that those who choose franchising are substantially
better off as franchisees than they would have been if they had tried to start their business on their own.

As for franchisor, as opposed to franchisee, survival, a number of studies have documented a high rate of failure for these (e.g., Shane (1996), Stanworth (1996) and Lafontaine and Shaw (1998)). To our knowledge, no study compares the exit rates of franchised and non-franchised companies explicitly. Instead, authors examine factors that affect the likelihood of survival of franchised chains. Lafontaine and Shaw (1998), for example, show that most of the chain characteristics at the time it begins franchising, including contract terms such as royalty rates and franchise fees, have little predictive power relative to success or failure. The number of years in business prior to the start of franchising was one of the very few factors that increased the likelihood of success in franchising. Silvester, Stanworth, Purdy and Hatcliffe (1996) found that a franchisor’s initial financial investment in the business and his or her strategy of choosing franchisees with prior experience also affect the survival and growth of their chain, while Azoulay and Shane (2001) suggest that offering exclusive territories to franchisees increases the likelihood of survival for young franchised chains. Finally, using a unique unbalanced panel dataset covering about 1000 franchise chains annually from 1980 to 2001, Kosova and Lafontaine (2010) show that the usual variables from the industry dynamics literature, namely age and size, affect franchised chain growth and survival most, even after controlling for chain fixed effects and other characteristics such as contract terms. They also confirm the result in Lafontaine and Shaw (1998) for the effect of the number of years that the franchisor spends developing the franchise concept before starting to franchise. As for the chains contracting practices, they find that a larger proportion of company units in the chain is positively related to growth and survival in franchising. Moreover, though these effects were not always statistically significant, they find that higher ongoing revenue streams for franchisors, in the form of royalties or other such fees, and lower upfront requirements for franchisees, are positively associated with both franchise chain growth and continued involvement or survival in franchising.

Very few studies have considered the consequences of share contracting outside of franchising. One such study, by Mortimer (2008), analyzes the move from linear pricing to revenue sharing in video rentals, a change that was caused by a fall in the cost of monitoring transactions on a per movie basis. Not surprisingly, since the change was voluntarily undertaken by the up and downstream firms — studios and retailers — she finds that it enhanced profits for both. More interestingly, using a
structural econometric model of firms’ contracting choices, she is able to quantify the benefit for both upstream and downstream firms. She finds that both up and downstream firm profits increased by 10% for popular, and even more so for less popular, titles. She also shows that small retailers benefit more from revenue sharing than larger retailers do, and that consumer welfare increased as a result of the adoption of revenue sharing in this industry. Both of these effects would have been hard to predict a priori. A second study, by Gil (2009b), shows that for movies whose revenues are most difficult to predict, namely non-blockbuster movies, reliance on revenue sharing in movie distribution contracts does not fully align distributor and exhibitor incentives. Consequently, movies of vertically integrated distributors are shown longer in distributors’ own theaters than in vertically separated theaters despite the systematic reliance on revenue sharing in this industry.

4.6 Vertical Restraints

Vertical restraints (VR) are restrictive contract clauses that one link in a vertical chain imposes on another. Usually, the manufacturer or upstream firm restricts the retailer or downstream firm in some way. VR are therefore instruments of control. There are many other ways in which one level in the vertical chain can exercise control over another or transfer certain rights across firm boundaries. Those methods are discussed in the next subsection. Here we limit discussion to the traditional VR that the antitrust literature has focused on.

In contrast to the share–contracting literature, the bulk of the VR literature evaluates the consequences of employing various forms of restraints, not the incidence. This is likely due to its focus on competition policy. Still, although the literature has often just mentioned where these occur, and then examined their effects, a few studies have tried to arrive at some conclusions as to why these are used by analyzing where they occur most. We begin with a few examples of the use of specific VR and then discuss some of the reasons for employing them.

i) Incidence

Exclusive dealing — where a manufacturer requires that a retailer sell only her products — is perhaps the most common form of VR. Indeed, all franchising involves some form of exclusive–dealing arrangement. However, one also finds exclusive dealing outside of franchise relationships. For example, the GM - Fisher Body contract involved exclusive dealing (Klein et al., 1978), as did the contracts between boat cap-
tains and tuna processors studied by Gallick (1984). Exclusive dealing also is used by manufacturers with their distributors: Heide et al. (1998) found that 46 of the 147 manufacturers they surveyed – all of them from the industrial machinery and equipment or the electronic and electric equipment sectors – used exclusive dealing clauses in their contracts with distributors.

An exclusive territory is granted when a manufacturer assures a downstream firm that he will be the exclusive reseller of a brand in a geographic market. As noted by Marvel (1982), exclusive territories often accompany exclusive dealing clauses. Not surprisingly then, they are commonly granted to industrial sales forces and wholesale distributors. It is also customary, for example, for cleaning–service franchises to grant exclusive territories. In fact, about 3 out of every 4 franchised chains grant some form of exclusive territory to their franchisees (Blair and Lafontaine, 2005, ch. 8). In their studies of manufacturing firms in the industrial machinery and equipment and the electronic and electric equipment sectors, Dutta et al. (1999) found that 69 of the 147 firms in their final sample used territorial restrictions.

Tying refers to situations in which a manufacturer requires its customers to purchase product B as a condition for obtaining what they really want, namely product A. Well known examples include IBM, which required that purchasers of computers also buy punch cards, and movie distributors who practiced block booking in the early days of the industry. Block booking is a form of bundling that requires that exhibition houses rent packages of, rather than individual, films. Perhaps the most famous tying case, however, involved Microsoft’s attempt to tie the use of Internet Explorer to its Windows operation system (see Whinston (2001)).

With resale price maintenance (RPM), the upstream firm exerts control over the price that the downstream firm can charge. RPM takes many forms including setting a specific price or a price floor or ceiling. Due to the fact that RPM is or has been illegal in most countries, examples often come from antitrust challenges, which have included cases involving firms in sectors such as gasoline distribution, recreational equipment, and brewing and distilling among others. Franchisors in particular have been known to exert downward pressure on the prices charged by their franchisees (i.e. maximum RPM; see Blair and Lafontaine, 2005), while in other contexts manufacturers, including high-end electronics and fashion firms, have successfully implemented minimum pricing requirements.

Going beyond mere examples, Ippolito (1991), in particular, examines the population of all 203 reported cases of resale price maintenance in the US between 1975
and 1982, a period during which a fairly broad interpretation of what constitutes RPM was adopted by the courts, and during which she argues the courts adhered quite strictly to the *per se* standard. She shows first that vertical restraints are often used together. Firms simultaneously relied on other vertical restraints in 122 of the RPM cases, most frequently territorial, tying, or customer restrictions (49, 31 and 32 of the cases respectively). Cases of RPM also often involved other charges, in particular horizontal price fixing in 30, and refusal to deal in 40 of the cases. In addition, Ippolito finds evidence that a non-trivial portion of RPM cases, namely 65% of all private, and 68% of all public cases in her data, arise in contexts where products can be classified as complex, new, or infrequently purchased, which are the types of products where the special services theory for RPM is most likely to hold. She also finds another largely overlapping segment of both private and public cases arising in contexts where dealers can influence the quality of the final good or the customer’s experience in important ways. Here again, manufacturer controlled pricing can alleviate the fundamental principal-agent problem that efficiency motives and organizational economics emphasize. Yet another set of (mostly franchising) cases seems well explained by concerns over vertical sales-effort externality problems. She concludes that collusion is not the primary explanation for the RPM practices that were prosecuted during this period.

Heide *et al.* (1998), for their part, focus on exclusive dealing, which historically has not been treated as harshly as RPM by the U.S. antitrust authorities. As a result, they were able to obtain survey data which they used to examine what leads manufacturers to use exclusive dealing in their contracts with distributors. They found that manufacturers who were more concerned that their promotional efforts, training, or general support of distributors might benefit their competitors were much more likely to adopt exclusive dealing arrangements. On the other hand, when it was difficult for manufacturers to assess whether their dealers sold other manufacturers’ products (i.e. when monitoring the behavior of dealers was difficult *ex post*), or when manufacturers perceived that their customers had a preference for multi–product distribution, they were less likely to rely on exclusive dealing. Again, these results are consistent with the type of efficiency or principal-agent arguments one finds in the organization economics literature.

Finally, Zanarone (2009) compares contracts used by nineteen car manufacturers with their dealers in Italy before and after the European Commission prohibited the use of location clauses in car distribution in 2002. These clauses prevented dealers
from selling cars outside of their territories. Zanarone shows that once exclusive ter-
ritories became illegal, the number of car manufacturers who imposed price ceilings,
required dealers to abide by a variety of explicit standards, and required dealers to
contribute to an advertising fund that the manufacturer controlled, went up signifi-
cantly. He explains the latter two changes as direct responses to the reduced dealer
incentives to advertise and provide pre-sales services. As for price ceilings, he sug-
gests that they may have become necessary to prevent dealers from circumventing
their quantity floors, something they could do he argues by selling aggressively out-
side their territories while maintaining supra-normal prices in their own, perhaps
isolated, markets.

\textit{ii) Effects}

The results of studies that have examined the effects of vertical restraints on firm
performance as well as downstream prices and quantities have been summarized in two
recent papers, Cooper et al. (2005) and Lafontaine and Slade (2008). In both cases,
the authors find that, on the whole, vertical restraints imposed by manufacturers on
their resellers tend to be associated with lower costs, greater consumption, higher
stock returns, and better chances of upstream firm survival. In other words, they
are efficient devices for aligning incentives, eliminating free riding, and controlling
opportunistic behavior, as the studies on the incidence of these practices summarized
above also suggest. Moreover, there is little evidence of foreclosure.

The evidence also suggests strongly that mandated restraints, such as the exclusive
territories that car manufacturers are required to provide to their dealers in some
states, lead to higher prices, higher costs, shorter hours of operation, lower consump-
tion, and fewer dealerships. For example, in his study of the effect of state laws
protecting the territories of car retailers, all of whom operate under exclusive–dealing
contracts, Smith II (1982) found that car prices and dealership values rose, while
hours of operation fell, after the state laws were enacted. In line with the results
in Zanarone (2009) mentioned above, however, Brickley (2002) finds that franchisors
adjust the terms of their contracts after the passage of laws protecting franchisees
against termination such that, when all is said and done, franchisees are no better
off.\footnote{Similarly, the benefits of increased state regulation governing car manufacturer/dealer relations
accrue to incumbent dealers at the time they are enacted, in that these dealers either derive greater
profits or other benefits directly, or can resell their dealerships at higher prices. Thus the state
regulation encourages potentially excessive investment levels by dealers. See Lafontaine and Scott
Morton (2010).} So although the focus of the latter is not traditional VR, the conclusions
reached are the same as in that literature.

The authors of both surveys above conclude that manufacturer and consumer welfare tend to be aligned when it comes to vertical restraints, as manufacturers have every incentive to develop lean distribution systems that get their products to their customers at the lowest possible cost. Retailers and distributors, on the other hand, do better, or at least anticipate doing better, in the more protected, or less competitive, environments that government mandated restraints sometimes afford them.

While the results from the reviews of the literature on effects of vertical restraints are striking, they cannot be definitive if for no other reason than that the number of studies from which these conclusions are drawn is quite small, especially when one considers the number of different restraints and industries covered. Tests of the effects of vertical restraints are rare despite the intense interest of antitrust authorities in these issues, however, because of the challenges associated with endogeneity and the difficulty in obtaining the right type of data. In particular, efficiency motives for these restraints imply changes in costs, yet cost data are often unavailable. The recent work by Asker (2005), on exclusive dealing, Crawford and Yurukoglu (2009), on bundling, and Ho, Ho and Mortimer (2009), on full-line forcing, much of which post-date the surveys mentioned above, rely on structural approaches, along with detailed data and institutional knowledge, to estimate demand and costs, and evaluate the effects of the restraints. Consistent with the literature reviewed in the surveys above, Asker, Brenkers and Verboven, and Ho, Ho and Mortimer find that the restraints cannot be explained by anti-competitive motives - respectively foreclosure, strategic delegation, and leveraging of monopoly power. Also consistent with prior literature on the cable industry, Crawford and Yurukoglu (2009), however, find that bundling of services is costly to consumers in that industry.

Clearly, much more work is needed in this area. We hope that our discussion of the literature on VRs will encourage empirical researchers to consider those restraints as well as other types of contract clauses in their analyses of governance, and to bring vertical restraints more squarely into the realm of organization economics and contract theory.

4.7 Control Rights Allocation

All studies of contracting practices in some sense are about the allocation of control rights. Indeed, except for the very simplest fixed-price contracts, virtually all
contracts specify who can make certain decisions and/or who owns and controls the use of certain assets. For example, the choice between company operation and franchising, whether traditional or business–format, also largely determines who has the right to hire employees and set prices. Similarly, the vertical restraints discussed above allocate decision rights to one party or the other. Some empirical analyses of contracts, however, focus more directly on the issue of how different control rights are allocated. Moreover, the more recent studies of this type often test aspects of the theory of incomplete contracts, which emphasizes the importance of allocating decision or control rights to those parties whose decisions, investments, and/or efforts will have the greatest impact on final outcomes. These studies then explain the presence of various clauses based on the potential impact of allocating decision rights to the parties involved. To our knowledge, there has been no attempt to empirically assess the effect of this allocation.\textsuperscript{43} As a result, this section focuses on incidence only.

Some of the empirical work on contractual agreements that predates the development of incomplete contract theory has included detailed descriptions of the various components of contracts, often in the form of frequency tables for various clauses from a cross-section of contractual agreements. For example, Udell (1972) analyzes a total of 172 contracts for 167 distinct contract provisions in business–format franchise contracts, many of which entail the allocation of control rights. He finds, among other things, that 59% of the contracts include exclusive–territory clauses, 58% specify the days (hours) of operation for the franchised unit, thereby allocating to the franchisor the right to make such decisions, and 60% include clauses stipulating that the franchisor controls the products, services, or menus offered by the franchisee. Similarly, Contractor (1981) provides information on the frequency of various restrictions imposed on licensees in international license contracts.

Table 3 summarizes information about the main contract clauses described in these and in more recent studies such as Lerner and Merges (1998) and Ryall and Sampson (2009), who analyze control–rights allocation in technology alliances in the biotech industry and the telecommunications and microelectronic industries respectively, as well as Arrunada et al. (2001) who focus on the contracting practices of car manu-

\textsuperscript{43} Ciliberto (2006)’s analysis of joint ventures, which he treats similarly and for which he finds similar results as for vertical integration, could be viewed as an exception. He notes that in the relationship between hospital and physicians, the asset is the physician’s patients. Under a joint venture, he argues that the hospital gains some control over this asset. He then shows that joint ventures between hospitals and doctors positively affect hospital investment in new technologies, just like vertical integration does.
facturers with their dealers in Spain. Authors of such studies often comment that contracting practices are not easily labeled or classified. Not surprisingly then, categorization schemes vary across studies, making it particularly difficult to summarize. In addition, often due to data constraints, different studies focus on certain aspects of agreements while ignoring others. Finally, some of the studies are concerned with franchising while others are about licensing and yet others about procurement. In spite of this heterogeneity, a few themes emerge from Table 3.

First, and not surprisingly, we find that there is much variety in contracting practices, even within contract types, that is when one examines only franchising arrangements, only licensing agreements, and so on. This variety reflects the fact that these contracts are used by numerous firms engaged in a variety of business activities.

Second, despite the variety, we find that the same issues are addressed repeatedly: typically, the contracts specify the terms of the exchange, including prices and sometimes quantities, and then limit the rights of the agent (licensee, franchisee, or supplier) explicitly to a time and place. They also give the principal the capacity to monitor the behavior of the agent and to terminate the contract at will or under certain conditions. Finally, they spell out what happens after termination. In other words, these contracts resemble leasing arrangements, with the principal providing an asset—whether it be a brand and business format, or a technology, or simply some desirable business as in the case of IT procurement—and then spelling out the limits within which the agent can use and profit from this asset. These limits, in turn, protect the principal’s ongoing interest in the value of the asset.

Third, the data in the table reflect the fact that the number or extent of restrictions imposed on the agent typically increases with the value of the asset involved in the transaction (see also table 2). Thus when a firm licenses its core technology, Caves et al. (1983) find that it imposes more restrictions on its licensees than when the technology is more peripheral to the licensor’s business. Brickley (1999) analyzes the tendency of a large sample of franchisors to rely on area-development agreements, which he equates with the provision of an exclusive territory—though in reality these also entail some expansion rights—, or their decision to allow passive ownership, which dictates whether the franchisee can pursue other work or business activities while a franchisee, and finally, their decision to mandate advertising levels. He argues and finds that franchisors are more likely to impose restrictions such as these when there

44 See also Elfenbein and Lerner (2003) on contracts for portal alliances. This paper is not included in the table because the paper does not provide simple frequencies on the contract clauses they focus on.
are significant externalities across outlets in a chain — that is when the brand is high value. Similarly, Arruña\textadddash, Garicano, and Vásquez (2001) find that contracts allocate more rights to manufacturers when the potential cost of dealer moral hazard is higher and when manufacturers’ opportunism is better controlled by reputation. The evidence in Lafontaine and Shaw (2005) goes even further as it establishes that the likelihood that there will be a contract, i.e. the probability that a principal will be willing to “lease the asset” to an agent, decreases with the value of the asset, which in their study is the value of a business format and franchise brand.

4.8 Rent, Reputation, and Repetition

As with any repeat\–business situation, the rent that parties earn within their relationship, or the difference between the profits that they can expect if they remain in their relationship compared to what they could earn outside of it, can play a crucial role in the maintenance and day\–to\–day functioning of the relationship. Firms can rely on this difference, along with some monitoring, to ensure that their contractual partners behave as requested or expected. Alternatively, firms can rely on the reputation of their contracting partner, and the cost the partner would bear if this reputation were damaged, to ensure performance.

i) Incidence

In a series of interviews, and subsequent analyses, Contractor (1981) found that licensor executives did not attempt to maximize their own profits by extracting all rent from licensees. Caves \textit{et al.} (1983) similarly note that contracts generally fail to capture for licensors the full rent that licensees can be expected to obtain. These authors conclude that leaving rent with licensees can mitigate moral hazard and attending control issues. Specifically, they note that licensees are expected to provide inputs that raise the value of the technology, inputs whose provision cannot be specified or priced easily in the contract. They write “With the licensee gaining no specific recompense for these inputs, letting him share the general profits benefits the licensor by mitigating their underprovision.” (p. 264) Similarly, in his series of interviews of railroad carriers and shippers, Palay (1984) found that those involved in the transport of goods such as cars, where the transport itself requires specific investments, were much more willing to make necessary adjustments to their agreement than parties involved in the transport of goods requiring no such equipment. He notes that shippers, in particular, indicated a desire to keep carriers “viable.”
The franchising literature also provides empirical support for the idea that franchisees earn rent. In particular, Brickley, Dark and Weisbach (1991), Dnes (1992), Kaufmann and Lafontaine (1994), and Michael and Moore (1995) find evidence suggesting that rent, combined with the threat of termination, plays an important role in franchising.

To illustrate, Brickley, Dark and Weisbach exploit variation in the existence and timing of state laws requiring good cause for termination of franchise contracts to show that franchisors use franchising less in states where they face laws that restrict their ability to terminate franchisees. This is true especially for franchisors operating in industries with mostly transient customers as the consumption decisions of such customers do not directly discipline franchisees that free ride. Those authors also find evidence that the passage of a law requiring good cause for termination in California in 1980 was associated with relatively large losses for the shareholders of publicly traded franchisors with operations in that state. Based on these results the authors conclude that those laws, by making it more difficult to use termination, increase the cost of controlling quality in chains.

Kaufmann and Lafontaine (1994), for their part, examine the profit and loss statements from typical McDonalds outlets and the resale prices of a small set of such outlets. They conclude that McDonalds indeed leaves rent with its franchisees, rent that the franchisees lose if they are found in violation of firm policy and are terminated. Moreover, McDonalds does not require franchisees to pay upfront for the full amount of ex–post rent. In other words, it allows franchisees to earn rent from an ex–ante perspective. The authors explain this in part based on the type of franchisees that McDonalds desires, namely individuals who will devote themselves to their role as owner operators of their restaurant(s). The authors also note that the company reserves the right to decide which franchisees are granted expansion rights, that is which franchisees get to own additional McDonalds restaurants, based on a franchisee’s degree of compliance with firm policies. The prospect of additional ex–ante rents thus also creates incentives for current franchisees at McDonalds, encouraging those who aspire to grow their business to continue to function within the bounds set by their franchise contract. The option of adding restaurants is made especially valuable moreover as the contract requires franchisees to work full time at their business and “keep free from conflicting enterprises or any other activities which would be detrimental to or interfere with the business of the Restaurant.”

---

45 Michael and Moore show that other franchise systems also leave rents with their franchisees.
46 McDonald’s Franchise Agreement, 2003.
Finally, Dnes dedicates a whole chapter in his book to what he calls the “wider” franchise contract, where he discusses what he calls the relational aspects of the contract, namely a number of business understandings that have evolved over time within franchise relationships. These include a tendency in some franchise networks to let franchisees sell non-authorized goods, implicit territorial protection, an understanding that the franchisor might support a distressed franchisee directly (e.g. by reducing royalty or lease payments or providing discounts), and an understanding that the franchisor will repurchase franchised outlets if needed. These aspects of the franchise relationship are implicit and are sustained only to the extent that the value of the relationship to both parties is sufficient to prevent them from altering their “understandings.”

ii) Effects

A different perspective on the issues of repetition and reputation is obtained from studies of industries where relationships are shorter term than in the franchising context, but where firms choose to repeatedly do business with the same partner. Anand and Khanna (2000), for example, find that 30% of licensing deals in their data involve firms with prior relationships or otherwise related firms, whereas 34% of international deals are concluded by firms with prior relationships. In Banerjee and Duflo’s (2000) study of contracts for customized software, 41% of external projects involve firms that have worked with the same clients before. Repetition is not ubiquitous in inter-firm contracting, however. For example, only 6% of the technology alliances in Lerner and Merges (1998) involve firms that have been alliance partners before.

Most relational and reputational contracting models rely on repeated interaction to sustain cooperation, a factor that lends itself to empirical assessment. As discussed earlier, Banerjee and Duflo (2000), Corts and Singh (2004), Kalnins and Mayer (2004) and Shi and Susarla (2008) examine the effect of repeated interaction on the choice of fixed-price or cost-plus contracts. The last three find that repetition reduces the need for high-powered incentives, but Banerjee and Duflo find no significant effect of repetition on contract choice. On the other hand, they find a strong effect for reputation, measured by the age of the software firm producing the customized product for the client. Specifically, young firms, which they argue have less reputational capital, are much more likely to be hired under a fixed-price contract than are older firms. This result suggests that firms use high-powered incentives for agents that have no reputation at stake, but are more willing to rely on more flexible lower-powered cost-
plus contracts when the supplier comes to the relationship with more reputational capital. Shi and Susarla find that larger suppliers, who also presumably have more reputational capital, are more often hired under cost-plus contracts. Robinson and Stuart (2007) find that pharmaceutical companies take a smaller equity position in their biotech partners when partners are central or well embedded in the network of prior alliances in the sector. They interpret this result in terms of reputational capital as well, which they argue allows firms to rely less on explicit control mechanisms such as equity.

The issues that underlie relational and reputational contracting assume particular importance in environments in which the legal system is not well developed, such as developing economies. McMillan and Woodruff (1999a and 1999b) examine relational contracting in Vietnam and find that prior experience with a trading partner and prior information gathering on that partner (presumably when the information gathered is good) are associated with increased provision of credit, which they equate with more trust. Furthermore, since they find that customers who are located through business networks also receive more credit, they conclude that networks are used in group sanctions or punishments. However, they also find that retaliation is not as forceful as one would expect in a standard repeated-game framework, as renegotiation often follows breach.47

Whereas the studies above examine the effects of repetition or reputation on the mode of contracting, Gil and Marion (2009) consider effects on costs directly. Specifically, they examine how the stock of prior interactions between contractors and subcontractors in highway construction affect the bidding behavior of contractors. They find that prior interactions with the subcontractors included in a bid allow contractors to bid more aggressively, and participate in more auctions. They interpret these results as evidence that prior interactions improve relationship-specific productivity. They note, however, that past interactions may improve learning or lower coordination costs, rather than provide the type of incentives through repeated business effects that authors are trying to capture in this literature. They then show that the effect of prior relationships on bids and auction participation depends crucially on the number and dollar value of contracts expected to be up for bid in the relevant region in the coming year. They conclude that past interactions affect costs only when firms also have expectations of future business opportunities. Along similar lines, Lyons (2002)

47 See also Gil (2009a) and Gil and Lafontaine (2010) on contracting and renegotiation, and e.g. Wilson, Lafontaine and Perrigot (2009) for an analysis of the effect of regulatory uncertainty on contracting in the hotel industry.
finds evidence that firms can support specific investments via reputation and rent through the establishment of partnership agreements, whereby firms agree to work together in the future, or preferred supplier agreements, which also imply longer-term relationships.

Viewed as a whole, this empirical literature suggests that explicit incentives become less important when either reputation or repeated interaction, and associated rents, are present and thus can be used to support transactions. Conversely, explicit incentives become especially important in the absence of reputation or expectations of future interactions as long as the legal system is capable of enforcing those explicit provisions. Ryall and Sampson (2009), however, find that repeat interactions are associated with an increase in contract detail in their sample of technology development agreements, suggesting a type of complementarity between formal and informal contracting. Similarly, Lafontaine and Raynaud (2002) argue that self-enforcement mechanisms complement residual claims allocations in providing incentives to franchisees. This issue of complementarity among contract terms or contracting practices has taken many forms in the literature. We turn to it in the next, and last section of this chapter.

4.9 Complementarities

Most of the literature on contract terms focuses only on some aspects of contractual relationships, e.g., financial terms, quantity clauses, contract duration, or control rights. Though authors emphasize individual contract characteristics, individual terms are often but one component of complex sets of contracting practices that interact and work together as a group. This interaction is similar to that which characterizes human–resource practices as described by Ichniowski, Shaw and Prennushi (1997). Indeed, in his study of the practices of five fast-food restaurant chains, Bradach (1997) argues that the mechanisms and systems that franchisors rely on to govern their relationships with their franchisees interact with one another, as do the mechanisms employed within the managerial employment contract on the company–owned side of these firms. He further suggests that the two separate forms of governance complement one another so that franchisors who use both company and franchised units can better address what he describes as the main managerial challenges that retail chains face.

Formally, complementarities occur when the marginal profitability of one action (e.g., practice or contract clause) increases with the level of another. In other words,
there are synergies among the choices. Although the idea that one needs to consider complementarities in analyses of the choice of mode of governance has gained momentum since the mid 1990’s, in reality, the idea did not go unnoticed by early authors in this literature. For example, Goldberg and Erickson (1987) note that because many contractual provisions and organizational–form decisions are made simultaneously, they can interact, and thus empirical studies should strive to estimate decisions concerning the set of contractual provisions and organizational decisions together.

Overall, the empirical evidence supports these claims. In particular, Lafontaine and Raynaud (2002), who approach the problem descriptively, review the agency and self-enforcement arguments for franchise contracts and describe in some detail the set of clauses that support each. Echoing Klein (1995), they conclude that explicit and implicit incentive–provision mechanisms are themselves complements rather than substitutes. Specifically, they note that while the ownership stake of a franchisee and his residual claims give him reasons to work hard on the day–to–day operations of the business, these same residual claims can also lead him to free ride. This can take the form, for example, of cutting costs in a way that can harm the brand and thus the chain, or of catering to local customers to a degree that is excessive from the chain’s point of view. The self–enforcing aspects of the contract give the franchisor an opportunity to control exactly those behaviors that can arise as side effects from the allocation of residual claims and ownership rights to franchisees. Hueth, Ligon and Melkonyan (2008) show a similar tendency for explicit and implicit contracting to be positively correlated in their study of California fruit and vegetable intermediaries’ contracts. Finally, as mentioned above, Ryall and Sampson (2009) find that contracts are more detailed when a firm engages in many deals, with the same or different partners. They interpret this finding to mean that formal contracting complements relational aspects of contractual relationship in high-tech industries.

If one can characterize the relationship among the contracting parties using just a few “important” characteristics, it becomes possible to examine all possibilities. Brickley (1999), for example, examines three specific non–financial contract terms, namely restrictions on passive ownership, area development plans, and mandatory advertising requirements. He finds that the occurrence of these provisions is positively correlated, and thus concludes that they are apt to be complementary instruments of control.

In most real world contexts, however, there are numerous contract terms to consider. This, in turn, leads to the problem of dimensionality: if there are \( n \) factors to
assess, there are $n(n - 1)/2$ possible interactions, which in many cases is intractable. One must therefore put some structure on the problem. Unfortunately, this is not always done in a satisfactory manner. To illustrate, a standard practice in estimating a linear equation for the choice of organizational form, for example, is to include the set of other contractual provisions in the choice equation. While this allows each provision to affect the firm’s choice of organizational form, it does not allow for interactions among factors leading to that decision. Moreover, although most discrete-choice models that include other contracting practices among the regressors imply that the marginal effect of one clause depends on the others (since the model is nonlinear), that dependency is not flexible. Instead, it is mostly determined by the assumed distribution of the error term.$^{48}$

An alternative, and we believe promising, way to address the problem is to group contracts among a smaller set of types, focusing on major groups rather than small variants within groups. This, of course, is the comparative institutions approach emphasized by Williamson (1991, 1996) applied to contracts and contract terms, where the latter often lend themselves quite well to grouping.$^{49}$

To illustrate, Bessy and Brousseau (1998) use cluster analysis to classify their sample of technology licensing contracts among five categories from simple, transactional contracts to complex multidimensional and very incomplete contracts. They argue that the characteristics of industries and intellectual property regimes largely explain the choice that firms make among these different types of contracts. Arruñada et al. (2001) also “group” sets of rights found in the car dealership contracts that they study, and find positive correlations among their categories of rights (completion, monitoring, and termination), and between these and the financial terms of the contract, namely the level of discounts offered to dealers. They also find that these correlations are much lower after controlling for characteristics of the transaction, however, suggesting that these characteristics, and thus the same underlying factors, affect the use of different contract terms. Still, they find a significant positive relationship between completion and termination rights, and between monitoring rights and discount levels (incentives), even after controlling for common sources of variation. Similarly, Kaplan and Strömberg (2003) show that their 213 venture capital contracts

---

$^{48}$ To illustrate, with a logit, if there are $n$ practices or clauses, once the pattern of complementarity/substitutability between an arbitrary clause, $x_i$, and the remaining $n - 1$ clauses is determined, all other relationships are known.

$^{49}$ This approach has been useful in other contexts as well. For example, clustering or grouping is the approach used by Ichniowski, Shaw and Prennushi (1997) in their study of the productivity effects of human resource management practices.
can easily be classified among five separate groups, and that the clusters thus generated are easily ordered in terms of either cash flow rights that are relinquished by the founder, or board and voting rights allocated to the venture capitalist. Thus the first (fifth) cluster is one where the venture capitalist (founder) obtains most of the cash flow and control rights. In other words, the rights “move” together to create the desired set of incentives and venture capitalist control.

In essence, grouping contract clauses is also what one does when contrasting say franchising and company ownership, or cost–plus versus fixed–price contracts, or independents versus affiliated businesses. Of course, looking at these as discrete alternatives is not a perfect solution. In particular, there may remain much variation among the members of each group. For example, franchise contracts can differ importantly across chains, and this variety is interesting in its own right. Still, franchising is also fundamentally different from company operation, such that contrasting these governance modes facilitates assessment of incidence and effects. The fact that empirical studies of franchising have found that factors suggested by theory explain the choice of how much to franchise better than they do financial contract terms further supports the idea that contractual relationships may best be understood as combinations of contractual practices rather than as phenomena that require that every component be analyzed separately.\footnote{The fact that franchise contract terms do not vary across franchisees that join at a given time, and that these terms also do not change much over time, suggests that firms adjust to different circumstances more by the choice among the two types of contracts than by designing new forms of the same contract type.}

Consistent with this approach, Bajari and Tadelis (2001) suggest that, in spite of the fact that they are limit cases, fixed price and cost plus contracts are fundamentally different from linear contracts where the supplier is reimbursed for a portion of his costs, i.e., the case where $1 < \alpha < 0$ in the notation in table 1. Lafontaine and Slade’s (2007) approach to relating the decision to vertically integrate or contract with an independent agent to the agency–theoretic model predicting the optimal share parameter also relies on the introduction of some non-convexity, in this case a cost of contracting, when the share parameter is not zero.

Finally, the theoretical and empirical literature on multitasking (e.g., Holmstrom and Milgrom (1991) and Slade (1996), respectively) is also an examination of complementarity and substitutability, since it shows how the characteristics of one task affect compensation for another. This in turn implies that models that assess compensation for one activity in isolation are incapable of capturing the full picture. Moreover, the
theory makes it clear that the problem worsens when the outputs of some tasks are difficult to measure and thus not compensated directly. Under those circumstances, low–powered incentives for all tasks may be preferred, since such incentives are not associated with a diversion of effort from hard to easy–to–measure tasks.

One should perhaps not be surprised that contract types can often be classified as described above. After all, in most industries, contracting parties and their lawyers have devised a set of templates or standard forms that many industry members rely on. As Bajari and Tadelis (2001) note, the central clauses found in such standard forms have the advantage that they are well understood in the industry and their interpretation has been clarified through a substantial body of case law. Moreover, those templates can still provide much flexibility when it comes to specifics. Perhaps what is more surprising is that the use of such templates seems to offer at least a partial solution to the dimensionality problem inherent to the study of contracts in the presence of complementarities.

5 CONCLUSION

In this chapter, we have reviewed a large body of empirical literature that deals with inter-firm contracts. We have made a conscious effort to include within the purview of the chapter some literature that predates the development of many of the theoretical models that now guide much of this work. By doing so, we hope to have brought some focus back onto these earlier, often more descriptive, contributions to our understanding of inter-firm contracts. We also hope to have convinced researchers interested in questions of organizational theory that the literature on vertical restraints is an important component of the larger literature on how firms organize their relationships.

While our chapter and references cover a lot of ground, we have also chosen to exclude certain areas of research. In particular, we have focused on firms’ relationships with other firms in their supply chain rather than their relationships with the providers of capital, be they banks or venture capitalists. Readers interested in organizational issues related to firm finance should consult Gertner and Scharfstein, chapter 19 in this volume. When faced with constraints on length or a need to keep things manageable, we have chosen to emphasize mostly early and recent work on

51 See MacLeod (2007) and Kornhauser and MacLeod, chapter 23 in this volume for more on this.
52 Note that although many theoretical analyses of vertical restraints rely on considerations of incentive provision, this body of literature is not well integrated into the more orthodox study of organizational economics.
particular issues, rather than review the entirety of the literature on a topic. We did this to give readers an appreciation for where the literature started and where it now stands, in the hope that this would provide the most relevant information for those interested in contributing further to this literature. Finally, again to keep things manageable, we have focused on the literature in economics, although we have tried to include some of the more recent or relevant contributions from the growing literature on contracts that appears in management journals.\footnote{See also Argyres, Bercovitz and Mayer (2007) and the references therein for more on the management literature on contracts.}

In addition to taking stock of the literature, our goal, ultimately, was to generate interest in the area of inter-firm contracts. We believe that this chapter has highlighted the need for much more empirical research on a variety of topics. Indeed, our understanding of numerous issues is still quite primitive. Those issues include how firms choose the many terms of their contracts, when and why contract terms tend to vary significantly across firms but little over time within firms, how contracting practices interact with one another, and how explicit and implicit contracting are affected by learning, reputation, and the value of future trade. We wish to stress, in particular, that in contrast to the theory, the entire area of dynamics (e.g., relational and reputational contracting) is under explored from an empirical point of view. In addition, there is a need to develop empirical models that can capture flexible patterns of complementary interactions while remaining empirically tractable.\footnote{One could perhaps borrow from the empirical literature on differentiated products to structure such models. For example, the distance–metric approach of Pinkse, Slade, and Brett (2001) could be used to organize the numerous possible interactions. To illustrate, if it were possible to assign characteristics (or clauses or attributes) to groups (e.g., factors that measure the importance of the agent’s effort or the riskiness of the market, or clauses related to “control”), one could model a common (presumably) substitution pattern within each group but allow a flexible pattern of complementarity/substitutability across groups.} Given the importance of inter-firm contracts in the economy, our summary of the evidence hopefully will encourage researchers to develop further theory and empirical work in the area, and a stronger link between the two.
References


