

**AN ECONOMIC ANALYSIS OF CONTRACTUAL RELATIONSHIPS
IN FRANCHISING SYSTEMS WITH CASE STUDIES**

ANTONY WILLIAM DNES

VOLUME 1

**PhD
University of Edinburgh
1988**



BEST COPY

AVAILABLE

Declaration

This thesis consists entirely of my own original work and has been composed by myself.

Signed:

Contents of Volume 1

Abstract

Acknowledgements

A Note on Publication

A Note on Franchisee Identification

Glossary of Franchising Terms

Introduction A Fieldwork-Based Study of Franchising

Chapter 1 The Nature of the Firm

**Chapter 2 The Economic Theory of Franchise
Contracts**

Chapter 3 Methodological Aspects of the Study

Chapter 4 The Nature of Franchising

Chapter 5 Fee Schedules: Some Further Analysis

Chapter 6 The Wider Franchise Contract

Chapter 7 Specialised Assets

Chapter 8 Why Franchise?

Chapter 9 Conclusions

References

Contents of Volume 2

Case A	Apollo Blinds
Case B	The Garage Door Company
Case C	Avis
Case D	Mobiletuning
Case E	Young's
Case F	Wimpy
Case G	Pizza Express
Case H	Bally
Case I	Nationwide Investigations
Case J	Budget
Case K	Midas
Case L	Olivers
Case M	Quikframe
Case N	Cure 30
Case O	Barstock
Case P	Computerland
Case Q	Yves Rocher
Case R	Austin Rover
Case S	Ford

Abstract

Fieldwork methods are used to analyse the nature of franchise systems in the UK. These systems are viewed in terms of their contractual relationships following theoretical approaches suggested by the economics of organisation, including agency analysis. In particular, product, brand and specialised-input franchises are identified and fee schedules are placed in the context of the wider franchise contract. The fieldwork covers 19 case studies of UK franchising systems.

The thesis contains a predictive theory of franchising. This identifies initial investments which franchisees undertake as hostages with screening and bonding properties for the franchisor. These hostages influence monitoring costs. The thesis makes an operational use of modern transaction-cost ideas.

Acknowledgements

I wish to thank my PhD Supervisor, Dr. G.C. Reid of the Economics Department of Edinburgh University for his help and support throughout my work for this thesis. I also wish to thank Professor G.K. Shaw of the University of Buckingham for his support during the later stages of my preparation.

Thanks are due to the Business Studies Department of Edinburgh University for bearing the travel costs of my fieldwork. I wish also to thank Buckingham University for help with typing costs.

I am grateful to the total of 77 franchisors, franchisees and managers who granted interviews for the study. I wish them all well in their ventures.

A Note on Publication

Material drawn from Chapters 1 and 6, covering explicit, implicit and relational aspects of contractual relations has been accepted for publication in the Scottish Journal of Political Economy under the title 'The efficiency of central planning: a comment'.

Case C, covering Avis, has been accepted for publication in 'Cases in Marketing Management', edited by L. Moutinho, and due to be published by Addison Wesley in late 1988. This is with the full consent of Avis.

A Note on Franchisee Identification

Throughout the thesis, and particular in the case studies of Volume 2, all franchisees and some managers of company outlets are identified by codes which use the letter of the case and a number. Code B₃, for example, refers to the third franchisee interviewed in Case B (the Garage Door Company). Codes are mainly used so as not to divulge details of franchisees' businesses. To this end, all locational and similar details of sites, which could aid identification have been struck out with four Xs (i.e. XXXX) in the case studies. This step is taken due to the sensitive nature of some information and as an alternative to restricting access to the thesis.

Glossary of Franchising Terms

<u>Bootlegging</u>	This arises when a franchisee sells unauthorised products.
<u>Business-format franchising</u>	This requires the franchisor to supply a full set of sales, operating and administrative systems for the franchisee to use.
<u>Fractional franchise</u>	The franchisee runs the franchised business as a part of his activities on a site.
<u>Initial Fee</u>	The lump-sum payment from franchisee to franchisor at the start of a franchise relationship.
<u>Master franchise</u>	The franchise rights to an entire area are sold to a master franchisee who then sells individual franchises within the area.
<u>Mom and Pop</u>	This is a useful American term describing family-run franchised outlets.
<u>Multiple franchise</u>	A franchisee has one of these if he runs more than one franchised outlet for a franchisor.
<u>Royalty</u>	A regular payment from the franchisee to the franchisor usually calculated as a percentage of the franchisee's sales turnover.
<u>Severability</u>	Franchise agreements, like others, contain severable clauses: each stands independently of the others.
<u>Tie-ins</u>	These are tied-in sales: the franchisor requires the franchisee to buy products from him.

Introduction

A Fieldwork-Based Study of Franchising

A franchise is created when one party, the franchisor, allows another, the franchisee, to use his trade name in return for a fee which may be charged in a number of ways. Associated franchise services, such as store design, may also be provided. Franchising is usually associated with retailing and accounts for at least 4% of sales turnover in the UK (Stanworth 1985). It is of more importance in the USA where at least 20% of sales in retailing are through franchised outlets. A franchisee may sell a branded service or a branded product. Franchise relationships based on product sales are the oldest and go back to the mid-nineteenth century.

This study examines the contractual relationships created in franchise systems. The analysis has two distinctive features. First, the theoretical background is drawn from a branch of the new institutional economics: this is the economics of organisation as it has been developed by writers like Williamson, Klein, Rubin and others. Secondly great care is taken in the gathering of data on the contractual relationships in franchising: case studies of 19 franchise systems are constructed using fieldwork methods which have been developed for general use within the social sciences. These methods are discussed and developed by Miles and Huberman (1984).

The aim of the study is twofold. First, we wish to understand the nature of franchise relationships. Just what economic purpose is served when franchisors require franchisees to make a lump-sum payment to them upon entering the franchise network? Why are professional-limitation clauses used to prevent franchisees from competing with the franchise network when they leave it?

Why is franchising a profit-sharing relationship? There are a number of questions of this kind which we address throughout this study.

However, we also wish to predict when franchising will be used as an alternative to full forward integration or sales through independent retail outlets. Existing theories of franchising, such as they are, tend not to answer this type of question but concentrate on the nature of franchising. In this study, Chapter 2, and Chapters 4 to 7 are about the nature of franchising. Chapter 8 is predictive. All theoretical insights are gained from analysing case-study material reported in Volume 2 of this study.

This study is of an organisational form lying between markets and hierarchies. Chapter 1 reviews the economics of organisation, which is based on transaction-cost ideas.

Much modern work (for example, Williamson, 1985) attempts to make operational use of transaction-cost notions by discovering and analysing important categories of transaction cost. Transaction costs are costs of running a part of the economy, such as those involved in negotiating and policing contracts. We show that the study of franchising leads to the operationalising of transaction-cost concepts. Franchise contracts are driven by monitoring cost economies that are achieved by some of the contractual devices, like professional limitation, which are used.

This study has three major results. First, franchising is better understood after the analysis of the 19 case studies. Secondly, predictions can be made about when franchising will occur. Finally, transaction-cost concepts are operationalised in the study. The results arise from a fieldwork-based approach which is described in Chapter 3.

CHAPTER 1

The Nature of the Firm

The empirical and theoretical work contained in this study, and the existing theoretical approaches to franchising discussed in the following chapter, make use of ideas that have evolved in the debate over the nature of the firm which began in the 1920s. In this chapter, the work of some of the principal authors in this field is surveyed. These form a comprehensive but not exclusive list of contributors; the main reason for inclusion being the relevance of any work to the subsequent analysis of franchising. The ideas collectively form the basis of a relatively new branch of economics which may be called the economics of organisation.

Risk, entrepreneurship and the firm

Knight (1921) is the first major economist to consider a non-technological explanation for the existence of firms. Prior to this work, it was generally accepted by economists that firms exist to take advantage of scale economies, based on indivisibilities as in Adam Smith's view that the division of labour is limited by the extent of the market (Smith 1776). Knight's work is principally concerned with distinguishing the concepts of risk and uncertainty on the basis that risk is measurable whereas uncertainty is not. Assessing the nature of business risks and uncertainties, leads Knight to a view of the firm as a device for the sharing of risk and bearing of uncertainty.

In particular, Knight agrees that whilst businesses will be organised to promote unity of purpose and specialisation of individuals' functions in order to turn uncertainty into measurable risk as far as possible,

uncertainty will necessarily remain a part of business enterprise. A special social class then emerges comprising businessmen, or entrepreneurs, who enter into a double contract with their employees. In exchange for guaranteeing a certain wage to them the businessman is permitted to direct the actions of his employees. In this view, nobody would be prepared to assume responsibility for the incomes generated by others unless given control over their actions. Furthermore, the world divides into more or less risk-averse types who become employees or employers respectively.

It is a standard criticism of Knight's approach, originally put forward by Coase (1937), that such a specialisation of function could occur without the need to form a firm. It is often noted that the businessman could specialise in uncertainty bearing simply by selling these services in the market place. Knight can be defended by noting the importance that he attaches to the problem of moral hazard which arises in insurance contracts.

'The fact which limits the application of the insurance principle to business risks ... is ... the peculiarly obstinate connection of a moral hazard with this sort of risk.' (Knight, 1921, 251)

Moral hazard means that shifting of risk causes a weakening of incentives to efficiency (Arrow, 1962, 614). This means that uncertainty bearing and control of employees cannot be separated: nobody will guarantee another's income if he is then vulnerable to reductions in the other's productivity brought about because the insurance reduces effort. McManus (1975) argues that the firm arises because other means of enforcing behavioural constraints upon the insured are too costly to arrange. For McManus, Knight may be seen as concentrating on particular transactions costs: those of organising

incentives that would cause managers to maximise the welfare of the uncertainty-bearing entrepreneur.

Knight defined the firm as the kind of business organisation that he saw around him. This is not in itself a problem. Whilst the argument is most easily stated in terms of some kind of sole proprietorship, Knight does consider the joint-stock company. It is a matter of detail, particularly concerning the difficulty of enforcing behavioural constraints on workers and managers, to extend the view to encompass such things as worker's cooperatives.

It may be noted that some of Knight's view of the firm is revived in the modern guise of principal-agent analysis. In this, it is standard practice to regard the principal as entering into a risk-sharing contract with his agent (Ross 1973, Shavell, 1979). The problem is then to find a contractual structure which overcomes the problem of moral hazard. Casson's (1982, 373) theory of entrepreneurship also draws on Knight's view of the entrepreneur as bearer of uncertainty.

Knight's analysis of the firm was an early contribution to the economics of organisation. Following McManus, this may be seen as less divergent from the transactions-cost analysis which followed it and which has given more obvious direction to modern thought. In retrospect, Knight gave an important reason for the formation of firms which may be included within more general frames of analysis.

Transaction costs and the firm

Coase (1937) attempts to discover why firms should emerge at all in exchange economies. The answer which he provides is that it becomes profitable to establish a firm

when doing so economises on the transaction costs of using market contracts to achieve the same objectives. Transaction costs refer to the costs of operating economic institutions like markets or firms: the costs of running the system as Arrow (1969) puts it. The transaction-cost terminology has really arisen from the work of Commons (1934) who argues that the 'transaction' should be the elementary unit of analysis in economics.

Coase examines the main categories of transactions costs which he believes are relevant to the emergence of firms. There are three of these. First, there is a cost attached to discovering the relevant prices ruling in some market. Secondly, there is the cost of negotiating separate contacts for every transaction. Finally, over a long-term economic relationship, it may prove difficult to predict exactly what is required of the contracting parties: contracts may then be left incomplete, being specified in general terms to allow easy adaptation of detailed performance to changing circumstances. It should be noted that Coase sees the firm as existing in an environment of uncertainty. Coase summarises his argument as follows:

'When the direction of resources becomes dependent on the buyer, the relationship that I term a "firm" may be obtained. ... the operation of a market costs something and by ... allowing some authority (an "entrepreneur") to direct the resources, certain marketing costs are saved.' (Coase, 1937, 392-3)

This dependency of the direction of resources on the buyer is of importance in factor markets where services are being exchanged. Markets on the whole work well for commodity exchange where it is easier to state the requirements of any contract, according to Coase.

Coase gives an analytical definition of the firm.

'A firm, therefore, consists of the system of relationships which comes into existence when the direction of resources is dependent on an entrepreneur.' (Coase, 1937, 393)

This definition gives scientific meaning to what is meant by stating that a firm gets larger or smaller. In particular, Coase argues that Knight (1921) cannot deal properly with this question on the basis of an uncertainty-bearing view of the firm. For Coase, a firm become larger as more transactions are organised by the entrepreneur and become internalised. However, this process is expected to be subject to eventually diminishing returns to management so that firms will reach determinate sizes. Robinson (1934) shows that diminishing returns may result from attempts to expand hierarchies within firms which is an idea developed by Williamson (1967). A point will be reached where the cost of organising another transaction within the firm equals the costs of using the market or of carrying out the transaction within another firm.

Coase also criticises the idea that division of labour automatically implies that firms will exist (Dobb 1925 and 1928). Coase argues that the integration of large-scale enterprise could in principle be achieved by the price mechanism, if transaction-cost differences between firms and markets were not significant.

'The "integrating force in a differential economy" [Dobb, 1928, 10] already exists in the form of the price mechanism.' (Coase, 1937, 398)

Thus, scale economies only have significance in economics due to transaction-cost barriers to market-based sharing of equipment; this point is repeated by Kay (1984).

It is from Coase's work on the nature of the firm that modern treatments of the economics of organisation

have developed. According to correspondence between Chueng and Coase, the paper was written by 1934 and was conceived in 1931-32 when Coase was 21 years old and before he had received his bachelor's degree from the London School of Economics (Chueng 1983). It did not have an initial impact which would have suggested the future importance of the work: it was 'much cited and little used' (Coase 1972, 63).

Chueng (1983) has criticised Coase on the grounds that the transactions-cost thesis is difficult to understand as it is never possible to give an unambiguous definition of the firm. This is a problem which affects everybody's work in this area. As Jensen and Meckling point out:

'...[The] firm is simply one form of legal fiction which serves as a nexus for contracting relationships'. (Jensen and Meckling 1976, 311)

Chueng implies that this is not a serious problem for Coase as the paper on the firm really refers to the economics of contractual relationships (Chueng, 1983). Chueng is then at pains to emphasise the value of this approach in studying particular forms of economic organisation. Chueng accepts that Coase appreciates the ambiguity (Coase, 1937, note 1, 392, quoted in Chueng, 1983, 18).

There is no doubt that a careful inspection of Coase's paper on the firm reveals that it is essentially an analysis of contracting. The terminology refers to 'contracting costs'. The concept of an incomplete employment contract is used, where today we would use the notion of explicit and implicit aspects of a contract (Holmstrom, 1981). In the final section of the paper, Coase compares his definition of the firm with the legal relationship between employer and employee and concludes

that they match and that therefore the definition fits the real world. Because of this contractual basis the paper is extendable, as later work has shown, to the analysis of virtually any specific organisational form when this is treated as a nexus of contracts.

Williamson (1975) argues that Coase's paper fails to operationalise the concept of transaction costs sufficiently to give the basis for assessing the transactions efficiency of alternative institutions. This is true in that Coase defines the firm as what results from transaction-cost minimisation, whilst just giving three examples of such costs. Much recent work has consisted of writers trying to fill in the details of Coase's insight into transaction costs.

Cohen (1979) points out that transferring a transaction from the market to the firm does not remove the need to identify relevant prices. The true private cost of a resource is its private opportunity cost however it is organised. This is not a strong criticism of Coase as it may be argued that the spirit of later transaction-cost work that the internal organisation of firms helps managerial hierarchies to devise incentive packages which may satisfy factor owners and safeguard efficient production. Thus, the existence of the firm may make it easier to discover certain prices. Indeed opportunities and opportunity costs might change as an organisation comes into being.

Coase's paper on the nature of the firm represents a tremendous insight into the nature of economic life which is one of the major developments in twentieth-century economic thought. The insight is into the contractual nature of organisations and is developed by later writers.

Knowledge, entrepreneurship and the firm

Hayek (1945) argues that the economic problem facing any society is not the one defined by Robbins of how to allocate given resources to alternative ends.

'It is rather a problem of how to secure the best use of resources known to any of the members of society, for ends whose relative importance only these individuals know.' (Hayek, 1945, 519)

Hayek's analysis is conventionally taken to lead to an understanding of the nature of markets. We learn from Hayek, following Smith, that a price mechanism coordinates economic activity in an uncertain world without the need for any one mind to comprehend all the details of economic life. Coordination becomes an unintended consequence of individual economic action. Prices are regarded as indices of market information to be used by economic agents without the need for a full statement of some supply or demand change.

The idea that information has characteristics of time and place and is dispersed in bits throughout the economy has implications for the study of organisations like firms. Individuals will have idiosyncratic knowledge and skills which may be integrated by the price mechanism but which may sometimes be beneficially internalised within a firm. Williamson (1975, 31) has developed the idea of 'information impactedness' essentially from Hayek's (1945, 524) notion of the 'particular knowledge of time and place'. Information impactedness refers to a condition where information is localised and is not shared by all parties to a transaction. Williamson uses observations on the specialised nature of knowledge, together with individuals' different and limited capacities for handling information (their 'bounded rationality') to identify

transaction costs which may be reduced by internal organisation.

Problems of team production

Alchian and Demsetz (1972) argue that Coase (1937) fails to account for important features of the firm. This is due to his concentration on limited aspects of the costs of using markets or internal organisation for transactions. Coase's transaction costs encompass the discovery of relevant prices, the achievement of flexible working arrangements and the negotiations needed to support some economic relationship. Alchian and Demsetz provide an explanation of the nature of the firm which they claim is consistent with Coase's but which emphasises different transaction costs. The advantage of this is that they can explain features of the firm not addressed by Coase such as the residual-claimant status of the firm's owners,

They state that any theory of the organisation of firms must be able

' ... to explain the conditions that determine whether the gains from specialisation and cooperative production can better be obtained within an organisation like the firm, or across markets, and to explain the structure of the organisation.' (Alchian and Demsetz 1972, 777)

Alchian and Demsetz make predictions of the circumstances likely to generate particular types of firm and then offer explanations of the structure of these. The types of firm include the modern corporation, partnerships, socialist firms and mutual organisations. They claim that these different types of firm are difficult to distinguish on the basis of Coase's market transaction costs and that the additional transaction-cost elements which they introduce are crucial to a proper understanding of them. The firm

is defined as a contractual device which includes the examples just mentioned. They are well aware of the problem of ambiguity which is attached to definitions of the firm.

Alchian and Demsetz argue that firms are formed to organise team production efficiently. Team production arises whenever several resources are used together and their outputs are not separable.¹ Separability arises if the marginal product of a factor is unaffected by the output of another factor. When this is not true, individual cooperative inputs do not yield identifiable outputs which can be summed to a total. The example given by Alchian and Demsetz of a non-separable production function is where two men jointly lift heavy cargo into trucks (Alchian and Demsetz, 1972, 780). The task is impossible for each man taken separately which makes it difficult to measure the contribution of each to output.

Team production will be used whenever a non-separable production function gives an output advantage over any separable functions which may exist. This may formally be stated as follows:

$$Z(x_1, x_2, \dots, x_n) > g(x_1) + h(x_2) + \dots + f(x_n) \\ Z_{x_i x_j} \neq 0$$

Where Z is a single-output production function and x_i is an input and $Z_{x_i x_j}$ is a cross partial derivative. The first condition tells us that there are gains from specialisation and cooperation which may be enjoyed if an organisational structure can be found for the team at a sufficiently low cost. According to Alchian and Demsetz the second condition implies that the measurement of a

¹ There is a parallel here with the literature on separable and non-separable externality. See Davis and Whinston (1962, 1966).

factor's marginal product will be much more difficult than in cases where it is not affected by the input from other factors. Superadditivity obtains, with the whole being more than the sum of the parts, yet individual contributions are harder to identify due to their jointness. The basic idea is similar to that of synergy in the corporate strategy literature (Ansoff, 1965)

Alchian and Demsetz argue that non-separability of inputs will cause a monitoring problem as individual factor owners have an incentive to free ride on the efforts of other team members. Free riding refers to a condition where an individual can hide his contribution of a minimal effort to group output whilst benefitting disproportionately from the reward to the group. It is this which brings forth the firm as an alternative to market transactions as a means of forming and monitoring the team. Shirking can be controlled when it is encountered.

Monitoring within firms will usually take the form of using measurements of team members' inputs as proxies for the harder-to-measure individual outputs. The firm is able to monitor individuals more cheaply this way and gains an advantage over market-based negotiations. As Williamson (1967) and Calvo and Wellisz (1979) show, monitoring effectiveness diminishes with the size of a firm.

Alchian and Demsetz (1972, 785) summarise their argument by stating that two conditions are necessary for the emergence of a firm. First, productivity may be increased through team production which makes shirking difficult to detect. It is then most economical to estimate factor productivity using input measures as proxies. The presence of these conditions:

'... leads to the contractual organisation of inputs, known as the classical capitalist firm with (a) joint input production, (b) several input owners, (c) one party who is common to all the contracts of the joint inputs, (d) who has rights to renegotiate any inputs contract independently of contracts with other input owners, (e) who holds the residual claim, and (f) who has the right to sell his ... residual status.' (Alchian and Demsetz, 1972, 786, emphasis added).

McManus (1975, 348) argues that neither non-separability nor any other property of a production function would create a monitoring problem if all the relevant characteristics of outputs could be costlessly measured. According to him, the firm seems to have no particular advantages over markets in Alchian and Demsetz's paper.

'Monitoring costs are incurred because the technical characteristics of inputs and outputs are such that it is costly to measure them precisely. Team production ... does not create a monitoring problem, although "teamness" is a source of gains from trade, which provide the motive for seeking a solution to the problem.' (McManus, 1975, 349).

This argument appears to be quite wrong. Alchian and Demsetz argue that non-zero cross partial derivatives in the production function will encourage shirking by making it more costly to measure individual productivity. This is both plausible and testable as an assertion. The monitor's residual-claimant status and the use of proxy measures of productivity may then give the firm particular advantages over market transactions.

Alchian and Demsetz are able to make predictions and observations about a number of types of firm. Among other predictions, they state that we should find the income of the monitor linked to profits. An observation about profit-sharing firms is that the incentive to shirk will

be smaller the smaller is team size, since in a large team the profit reduction from an individual's shirking will be spread thinly among many members. A further prediction is that socialist firms in countries like Yugoslavia will develop substitute management techniques to replace the residual-claim incentive for monitors. These examples suffice to show that Alchian and Demsetz succeed in developing their particular transaction-cost account of the nature of the firm to the point where they can explain both the creation of firms and details of their organisation.

Alchian and Demsetz point out that the authority relationship apparent in firms is an illusion, in contrast to Coase's view. The firm has no power of fiat in relation to its employees, in their analysis. The employer (monitor) can hire, fire and sue. This amounts to being able to reward or punish by offering or withholding future business or by seeking redress for contractual breaches through the courts. Alchian and Demsetz argue that this is exactly analogous to the situation of a consumer in a goods market. They also argue that long-term contracts are not an essential aspect of the firm: employers and employees must continually offer each other gains from association or else their contracts will simply break down. This view of the firm in which there is continual negotiation between employer and employee runs counter to subsequent developments in the theory. Alchian has now changed his views to accept the importance of long-term contracting (Alchian, 1984).

Their contractual view of the nature of the firm is testable. For example, licensing and franchising links between firms could be seen as arising when 'teamness' is absent so that the productivity of different inputs is separable. If such arrangements showed non-separabilities then Alchian and Demsetz's view of the origin of

transaction costs which favour the creation of firms would be challenged.

Considerations of team-production issues adds details to Coase's insight into the nature of the firm. There are connections here with the economics of property rights because of the concern with the residual-claimant status of owners of firms (De Alessi, 1980). Recent writers on the economics of organisation such as Williamson (1985) tend to add further examples of transaction costs which affect the choice of organisational form. The common concern is to avoid the charge that transaction costs can be so broadly defined as to make any analysis which is based on them devoid of any power of systematic explanation.

Asset specificity and transaction costs

Klein, Crawford and Alchian (1978) (henceforth KCA) develop observations made by Williamson (1975) and Teece (1976) on the vulnerability of idiosyncratic exchanges to opportunistic behaviour. Idiosyncratic exchanges are ones not regularly repeated in markets; assets are usually created which cannot be easily moved to other uses and are specific to the transaction. Opportunism refers essentially to a tendency towards cheating which may be observed in human behaviour. The vulnerability of idiosyncratic exchange embodying specific assets may make the use of markets a costly means of supporting a transaction.

KCA construct a theory of vertical integration in which as assets become more specific, appropriable quasi rents are created, the gains from opportunistic behaviour increases, and the costs of contracting in markets rise above those attached to internalising some activity. Vertical integration arises when firms internalise some

aspect of the supply of new materials, production or marketing attached to their product. According to KCA, we are more likely to observe vertical integration where assets are transaction specific.

The example which they give assumes that a specialised printing press is owned by Mr. A. and is hired by Mr. B. at a contracted rate of \$5,500 per day. The depreciation of the fixed cost of the machine is \$4,000 per day and it has a value in its next best use of \$1,000 per day. Operating costs are \$1,500 per day and are paid by A who prints finished pages for B. KCA define the quasi ('as if') rent arising on the transaction as the difference between revenue and opportunity costs for A. This is the difference between revenue minus operating costs and the salvageable value in the next best use ($\$5,500 - \$1,500 - \$1,000 = \$3,000$). Fixed costs are excluded from this comparison and only serve in the example to show that A makes normal returns in the long run.

A may be vulnerable to the following kind of hold up. B may realise that A's only alternative for the machine is the \$1,000 a day that it could earn elsewhere (in some less suitable use). He could, after striking a contract with A on the above terms, reduce his payment from \$5,500 to almost \$2,500 ($\$5,500$ minus $\$3,000$ quasi rent). A would be locked into the transaction as he would be worse off if he were to abandon it; he faces his fixed costs anyway and obtains some contribution towards them by staying in business. In this example, the quasi rent is 100% appropriable by B who demonstrates post-contract opportunism by his behaviour. The appropriability of the rent arises because of the conjunction of opportunism and asset specificity.

In general, not all of the quasi rent may be appropriable. There may for example be another publisher, Mr. C, who would pay up to \$3,500 per day to hire the machine. In this case, the appropriable quasi rent amounts to the difference between the market price paid by B and the one that would be paid by C (\$5,500 minus \$3,500 = \$2,000). All this says is that B could lower his payments to what C would pay, extracting that amount of quasi rent from A, and get away with it.

The entire example is relevant only when we are dealing with specialised assets and consequently with bargaining among small numbers of economic agents. There may then be a useful distinction between the next highest valuing user and the value of the machine in its next best use. Specificity applies in relation to both the use and user of an asset.

KCA argue that the existence of appropriable quasi rents will give an incentive for vertical integration. In principle, individuals could protect themselves against post-contract opportunism in three ways. First, they could write explicit contracts and enforce these in the courts. It is likely that the costs of specifying, monitoring and enforcing explicit contracts increase with the size of appropriable quasi rent as it becomes more worthwhile to renege on contracts. KCA quote the work of Macaulay (1963) which shows that it is unusual for contractors to rely on litigation to settle their differences; generally, it is common for them to augment agreements with implicit contracts which are unwritten understandings backed by market sanctions. KCA quote the then contemporary case of Westinghouse, which refused to fulfill contracts for the supply of uranium on the grounds of commercial impossibility, to illustrate the difficulty that explicit contracting anyway faces when it is subject to review in the courts on the grounds of reasonableness.

In practice, it may be hard to distinguish such things as genuine cost changes affecting a supplier from opportunistic attempts to extract quasi rent (KCA, 1972, 303).

Some interesting examples of implicit contracting are given by KCA. The discounts given to firms who have entered into agreements with firms who make specific products for them can be explained in this way. These represent an attempt to render the long-run market relationship so valuable that opportunism is suppressed. Sometimes companies post 'bonds' against good behaviour; they might buy an asset whose value is then made dependent on the actions of the group threatened by opportunistic behaviour. The bond is lost if opportunism is detected. KCA expect that the cost of these devices would need to rise as appropriable quasi rents increase (KCA, 1972, 307).

However, there is no reason to suppose that the costs of internalising a threatened transaction within a firm should vary in any systematic way with the size of appropriable quasi rents. We obtain a very precise and testable prediction: any advantages possessed by long-term contracting through implicit or explicit contracting will be eroded in value as asset specificity increases the size of appropriable quasi rents. Therefore, we should expect vertical integration to be more likely the higher are appropriable quasi rents. Sometimes, of course, no organising method is of sufficiently low cost to support a transaction.

KCA give a series of case analyses in support of their theory. In particular, they analyse the merger between Fisher Body and General Motors in 1926. When car production in the USA began to use closed metal bodies in 1919, General Motors entered into a long-term agreement

with Fisher. This contained an exclusive dealing clause in order to encourage Fisher to undertake the required specific investment without risking post-contract threats from General Motors to buy elsewhere. General Motors was protected from threats from Fisher to interrupt supplies unless price changes were agreed by cost-plus pricing provisions of the agreement. However, the contract did not work out in practice as the pricing system was complex. Also, Fisher's importance to General Motors increased with the growth of the use of closed metal bodies. General Motors wanted body plants to be located next to its own factories, which Fisher was not prepared to do because of the highly specific investment implied. In the end a merger was agreed (KCA, 1972, 310).

This case along with others covering the ownership of oil pipelines, investments in human capital, industrial leasing contracts, and the renting of brand-name privileges by franchisees succeed in illustrating KCA's theory. Furthermore, it is important to note that their approach is firmly based in the study of contractual relations.

'Once we attempt to add empirical detail to Coase's fundamental insight ... we find that his ... distinction between transactions made within the firm and transactions made in the marketplace may often be too simplistic. ... It may be more useful to ... examine the rationale for different ... contractual relationships.'

(KCA, 1978, 326).

This is precisely the point made by Richardson (1972) in his rejection of the usefulness of a view of the firm which sees

'islands of conscious power in [an] ... ocean of unconscious cooperation.' (Robertson, 1930, 85).

Richardson also accepts that introducing a spectrum of contractual relationships between firms and markets is not

necessarily inconsistent with Coase's insight (Richardson, 1972, footnote 1, 896).

Monteverde and Teece (1982) point out that KCA do not provide a theory of vertical integration. Rather they have a theory of quasi-vertical integration. It is not necessary for firms fully to integrate to overcome problems of asset specificity. Instead the vulnerable party can insist that the other make the specific investment. This supposes that the costs of quasi-vertical integration may often be found to be lower than those of full integration. An example of quasi-vertical integration is when a car producer provides the specialised equipment that a component supplier must use: such arrangements are widely observed. Monteverde and Teece find some statistical support for the relevance of quasi-vertical integration. At the very least they introduce a new contractual device into the spectrum analysed by KCA. Quasi-vertical integration allows specialisation advantages to be retained where these would be lost if more than just the specific assets were integrated.

KCA's work on the conjunction of asset specificity and opportunism, although it deals only with one type of transaction-cost problem, has had a far-reaching influence. This is very noticeable in comparing Williamson's early and later work (Williamson 1975 and 1985), as we do below.

Principal and agent

A number of authors have developed models of agency which illuminate certain aspects of contractual relationships. Rees (1985a and 1985b) and Arrow (1985) give excellent surveys of this work. An agency relationship arises when one party, called the agent, acts

on behalf of another, called the principal. The use of the terms principal and agent is looser and more general than in law where an agent acts as an exact proxy for his principal. In the economist's sense, agency relationships are pervasive. Doctors, subcontractors, employees, managers, politicians and bureaucrats are all examples of agents.

Arrow (1985, 37) gives an up-to-date version of what Ross (1974, 134), calls the canonical agency problem. There are now two main variants of the basic principal-agent relationship. Primary focus has always been on a model in which the agent's actions are not directly observable by the principal. In this variant, the payoffs to both parties are affected by the agent's actions and by chance factors. The principal must choose a fee structure to reward the agent and maximise the principal's returns. This turns out to involve a compromise between optimal risk-sharing and leaving some incentive for the agent to be diligent (Rees, 1985a, 6). Arrow (1985, 38) calls this variant the 'hidden-action' principal-agent problem. It is often referred to as involving moral hazard. In the case of a risk-neutral principal and a risk-averse agent, it can be shown that is it efficient for the principal to bear all risks, guaranteeing the agent a fixed income. However, no contract of this kind could ever be struck since the agent would then have no incentive to act diligently. Therefore, the agent will be left to bear some risk (Ricketts, 1986, 233). Moral hazard refers to the problem which arises when insurance alters the behaviour of the insured.

The second variant of principal and agent is called the 'hidden-information' problem by Arrow (1985). This arises when the agent has information which is not available to the principal. The model is also referred to as involving 'adverse selection'. The principal will find

it hard to prevent the agent from misrepresenting information to him. There are links here with Akerlof (1970) who shows that information asymmetries in markets can drive out above-average-quality products. An optimal contract in this case requires the application of a 'revelation principle' (Myerson, 1979). This states that if some contract leads to false revelation then some other can be found to induce truth telling without reducing anyone's welfare. An example is given by Rees (1985b, 86). Suppose that under some fee schedule, the agent reports state x when y has occurred in order to get payment $v(y)$. Then the principal could give him $v(y)$ when x occurs and be no worse off. Developing the principle that the agent should be no worse off when telling the truth than when lying leads to an efficient fee schedule in which the agent's income and efforts vary correctly with the states which arise. The contract is then said to be incentive compatible.

Two further points are worth making about the content of principal-agent models before moving on to discuss their relevance to the economics of organisation. First, principals are normally regarded as competing for agents so that fee scheduling has to recognise reservation incomes for agents. Also, it is possible to introduce monitoring activities and show rewards as varying with the monitored item.

In his evaluation of agency theory, Arrow (1985, 48) concludes that it gives good reasons for the existence of many types of contract but that it tends to predict rather more complex contracts than the ones commonly observed. An example is the sharecropping contract. In this a tenant farmer pays his landlord an amount $w(m)$ which depends upon the harvest and retains $z = p - w(m)$ for himself, where p represents the total harvest. $w(m)$ can be set to share risks between the two. For example, if

$w(m)$ is constant then the tenant bears all risk, and if z is constant then risk is borne by the landlord. However, Arrow points out that such contracts are remarkably standard across farms and regions and are regulated by custom. Yet agency predicts variation with climate and with relationships between effort and output.

Arrow (1985, 50) identifies three limitations to agency theory. First, the cost of operating contracts which are derived is not explicitly considered; yet, clearly, transaction costs arise or there would be no problems of uncertainty on which the models depend. Secondly, monitoring options available are difficult to define. Finally, non-pecuniary reward structures are important in the world but cannot be included in agency models.

What then is the relationship between agency theory and other theories of economic organisation? First, it may be said that agency is an extreme contract-theoretic view of economic organisation. In this, it is not felt useful to define boundaries between firms and markets: recall the earlier quotation from Jensen and Meckling (1976) referring to the firm as a 'nexus for contracting relationships' (they also refer to 'a set of contracts'). This is quite consistent with the growing recognition among writers like Cheung, Alchian, Demsetz and Williamson (see the next section) that the 'firm' is a somewhat fuzzy concept best replaced in serious work by the study of contractual relationships.

Secondly, there is similarity between the treatment of some topics in agency analysis, and elsewhere. Thus, Williamson's (1975) use of 'information impactedness' parallels agency notions of information differences. 'Shirking' is Alchian and Demsetz's (1972) version of adverse selection (average payments to team members may

select below average performance). Knight's (1921) work on uncertainty-bearing and control within firms has similarities with the idea of a risk-sharing contract. The idea of incentive compatibility has spread out into the economics of organisation, just as the idea of posting a forfeitable bond against good behaviour has spread beyond game theory.

Agency theory does on occasions provide detailed analysis of transaction costs. The work of Jensen and Meckling (1976) demonstrates that when an owner-manager of a firm sells shares in it he will bear the full agency costs of his action as long as equity markets anticipate agency effects. Agency costs are such things as monitoring costs faced by new principals over their agent and bonds that the manager may have to post to convince shareholders of his good behaviour. Buyers of shares will adjust their offer prices downwards to reflect these costs. The manager has an incentive to minimise these agency costs. However, in other respects transaction costs are neglected, as when considering the costs of specifying contracts. It would appear that agency costs are a category of transaction costs.

The major difference between agency theory and other economic theories of organisation may also be seen in Jensen and Meckling's example of agency costs. According to Williamson (1985, 28), the agency approach shows 'ex-ante incentive alignments in superlative degree'. Contracts are completely settled in advance, even where information or actions remain hidden. The tendency in Williamson's own work is to pay considerable attention to execution stages of contracts. This contrast, whilst broadly true, has exceptions as in Fama's (1980) treatment of ex-post settling up in managerial labour markets.

Markets, hierarchies and hybrid transactions

The transaction-cost analysis of contractual relationships is most developed in the work of Williamson. This began with early papers on questions of hierarchical control and on vertical integration (Williamson, 1976 and 1971), became firmly established with his analysis of markets and hierarchies (Williamson, 1975), and has more recently developed into a full analysis of the continuum of contractual relationships which lie 'between market and hierarchy extremes' (Williamson, 1985, 83). It is now inaccurate to describe his work as an analysis of markets versus hierarchies.²

Williamson (1985, 2-12) acknowledges his debts to earlier economists like Coase, Commons and Hayek, and recognises that connections can be made between his work and that of economists like Alchian and Demsetz (1972), Arrow (1969), Demsetz (1967), and Jensen and Meckling (1976). He has also been influenced by organisational theorists like Simon (1961) and Barnard (1938), by business historians like Chandler (1962), and, especially in recent years, by legal analysts such as Llewellyn (1931), Macauley (1963) and Macneil (1978). Although his work draws on the experience of writers from these disciplines, this is really in terms of raising questions and noting failures of explanations. Williamson remains firmly based within economics, offering economic explanations of organisational matters and attempting to expand the scope of economics. This is particularly true if we accept Buchanan's (1975, 229) claim that economics is more a 'science of contract' than of choice. The task Williamson (1981, 1546) sets for himself is to

² Many of the papers leading to the 1975 and 1985 books are contained in Williamson (1986). An important recent paper not summarised in one of the books is Riordan and Williamson (1985).

'operationalise' the insight of Coase into transaction costs and the nature of the firm.

In his earlier work, Williamson (1975, 39) emphasises that markets and hierarchies are alternative structures for organising economic activity. The choice of which is used depends upon the transaction costs attached to each, following Coase. Transaction costs are held to depend on human and environmental factors. Williamson's human factors are bounded rationality and opportunism. The environmental determinants of transaction costs are uncertainty and small-numbers effects. It is the joining of uncertainty and small-numbers with bounded rationality and opportunism that may give disadvantages to markets as organising devices in some instances. These four terms are now explained in more detail.

Bounded' rationality refers to the limits which exist on the human capacity to handle information. Williamson (1975, 21) quotes Simon to state that bounded rationality refers to human behaviour that is:

'... intendedly rational but only limitedly so'.
(Simon, 1961, xxiv).

There may be limits both on the capacity of individuals to receive, store, process and transmit information. These do not amount to a problem of costly information: one could have a limitless budget for information gathering and processing and yet be unable to improve information flow because of, for example, the low IQs of some individuals. It should be noted that bounded rationality need not rule out attempts at maximisation solutions to economic problems. It does, however, rule out comprehensive contracting as not all contingencies can be specified in even relatively straightforward transactions.

Opportunism refers to 'self-interest seeking with guile' (Williamson, 1975, 26). The concept extends standard self-interest assumptions in received theory to allow for strategic behaviour by economic agents who may be led to mislead, to lie to, or to cheat one another. Opportunism rules out self-enforcing promises as a basis for contracting between individuals. This does not suggest that the naive view is always wrong, but only that opportunism is a sufficiently widespread phenomenon so as to warrant recognition by economic theory.

Small-numbers effects play a role in Williamson's transaction-cost economics which is similar to that of monopoly in received neoclassical theory. If an idiosyncratic exchange arises then the parties to it are not driven together by competitive forces. Instead they must think it out for themselves and may encounter problems of bounded rationality and opportunism. Williamson (1975, 28) distinguishes between ex-ante and ex-post small-numbers problems. Ex-ante situations are those that are found at the start of contracting. However, often ex-ante large numbers can be transformed into small numbers. A 'fundamental transformation' can arise in which once trading parties form their links they become locked into their relationship; specific investments may be made, a firm may acquire first-mover advantages, or change may just be very disruptive. An example arises of the fundamental transformation when a franchisee signs with a franchisor: he will sink specific investments into his business and thereafter the returns offered by other franchisors become less relevant to him. The fundamental transformation is an idea which is important in Williamson's later work.

Williamson (1975, 24) regards uncertainty as a pervasive feature of economic life. He refers to uncertainty about states of the world and makes no

distinction between risk and uncertainty in contrast to Knight (1921). In his later work (1985, 57) Williamson introduces the idea of behavioural uncertainty, which arises because of individuals' inability to predict the incidence and form of opportunism.

It is the joining of bounded rationality, opportunism, uncertainty and small numbers which must cause problems for markets and favour the internal organisation of transactions. If one of them is absent then there are solutions which can, in principal and at a cost, support market transactions. However, for internalisation to actually arise there must be an efficient organisational form that can overcome whatever problems are hindering market trade. This last point is what makes Williamson's work a comparative institutional analysis.³ One of his important achievements is to have recognised that difficulties which may beset market trade must be reduced within the firm by the creation of appropriate decision-making structures and incentive mechanisms.

It is interesting to compare Williamson, et al (1975) with Alchian and Demsetz (1972) on the question of team production. Alchian and Demsetz derive a monitoring problem from the team-production requirements of technological non-separabilities. Williamson, et al, comment on this as follows:

' ... it is not non-separability ... that occasions the problem ... it is this in conjunction with ... opportunism and ... information impactedness that poses the difficulties.' (Williamson, et al, 1975, 253).

Information impactedness is a condition that derives from bounded rationality, uncertainty and opportunism. It

³ In the sense of Demsetz (1968).

exists when one party to a transaction has more information than another. Secondly, Williamson, et al (1975) argue that most tasks are separable anyway; except for the smallest teams, buffer inventories can be used to sever the connection between stages of production. I would argue that separability is not a key issue in generating a rationale for the firm if a behavioural basis such as opportunism and bounded rationality is used to generate transaction-costs. This is not to deny that separability has an effect on the magnitude of such costs.

There is some oversimplification involved in describing even Williamson's 1975 position as concerned with markets versus hierarchies. He is quite aware that authors like Richardson (1972) and Macauley (1963) make a strong case for examining a continuum of contractual relationships.

'I nevertheless urge that focussing on the significant differences between normal sales and hierarchical relations is useful ... terms of reference will emerge which will permit the cooperative properties of intermediate forms of contracting to be more accurately assessed.'
(Williamson, 1975, 109)

Markets and Hierarchies represents a beginning and points the way 'toward a transactional paradigm' (Williamson, 1975, 249).

A number of refinements are evident in Williamson's later work. He starts to write in emphatically transaction-cost-economics terms rather than in organisational-failures terms and to talk of the 'governance of contractual relationships' rather than of markets versus hierarchies (1979, 233). By 1981, he writes of the importance of specifying the dimensions along which transactions can differ (1981, 1546). By 1985

this leads to a clearer and better developed version of Williamson's theory of economic organisation.

Williamson argues that in choosing one form of economic organisation rather than another, individuals economise on the sum of production and transaction costs. This formulation holds revenues constant across different organisations. It is conventional to hold production costs constant as well so that the idea of net-benefit maximisation translates into transaction-cost minimisation (Williamson, 1985, 22). Neither assumption is entirely justified and both may be relaxed along the lines of Riordan and Williamson (1985) who use a model in which production costs are affected by the choice of organisational form. Generally speaking, Williamson does not believe that technology is decisive in determining organisation; although sometimes economies of scale or scope may be traded off against organisational savings. According to Williamson:

'Transaction-cost economics poses the problem of economic organisation as a problem of contracting. A particular task is to be accomplished ... in any of several alternative ways. Explicit or implicit contract and support apparatus are associated with each. What are the costs?' (Williamson, 1985, 20)

Transactions costs may be usefully divided into ex-ante and ex-post types. Ex-ante costs cover the drafting, negotiating and safeguarding of an agreement (safeguarding may involve the posting of bonds by parties anxious to demonstrate their good intentions). Ex-post costs cover such things as using courts, arbitration or other negotiation devices to settle disputes. Williamson insists that ex-post transaction costs must not be ignored as in principal-agent analysis.

Transaction costs have three principal dimensions, according to Williamson (1985, 52). First, there is the condition of asset specificity which we have already discussed in relation to Klein, Crawford and Alchian (1978). Williamson believes that the importance of asset specificity cannot be overstated. Asset specificity replaces small-numbers bargaining in Williamson's (1985) work, although he gives no guidance on how it might be measured. The second dimension of transaction costs is uncertainty: transactions may differ by the uncertainty surrounding their physical and human environment. Finally, transaction costs may differ according to the frequency of transactions; for example, non-standard recurrent transactions are likely to require specially constructed arrangements for their governance.

However, Williamson's approach is methodologically individualistic. There is no omniscient planner comparing transactions along these dimensions. Rather individuals and firms contract with each other in specific circumstances. In practice therefore transaction-cost comparisons are made by individuals assessing the costs and benefits of particular contracts. Generally speaking this involves the comparison of discrete totals; the method is not essentially marginalist (Williamson, 1981, 1544). The contracting occurs between individuals and firms and within firms. In different circumstances it may cover discrete transactions or continuing ones, use explicit agreements or implicit contracts, and rely on third-party intervention (using courts or arbitration) market sanction or administrative procedures for resolving disputes.⁴

⁴ These distinctions are covered by Macneil's (1978) concepts of classical, neoclassical and relational contracting. See also Goldberg (1980).

Williamson analyses the world of contract maintaining his behavioural assumptions of bounded rationality and opportunism. Contracting takes place within an uncertain world in which specific assets may be created. The fundamental transformation from Markets and Hierarchies continues to apply but to the creation of transaction-specific assets.

The contracting forms of planning, promise, competition and governance are defined by Williamson (1985, 31). Uncertainty is assumed to underlie all four situations, which vary according to the absence of one of bounded rationality, opportunism or asset specificity.⁵ Where parties are unboundedly rational but opprtunistic, complex incentive-alignment issues arise if locked-in features of contract exist. However, all relevant aspects of contract can be set at the ex-ante bargaining stage. Contract execution problems, either do not arise or are solved by interpretation of the original agreement in the courts or through compulsory arbitration. This is the sort of world described in agency theory where a type of market mediation is achieved. This is what Williamson calls planning, where this refers to the individual's ability to plan for future contingencies.

If opportunism is lacking, but the other attributes of contracting are present then we have a case where a person's word is his bond. All that is needed to safeguard against contract gaps is a self-enforcing general clause under which parties agree to behave efficiently and to seek only fair returns. Williamson uses the term 'promise' to cover this type of contract.

⁵ These cases dominate cases with two attributes of contract missing.

If it is asset specificity that is missing then parties to exchange have no interest in the continuing identity of one another. They do not become locked in to trading with each other. Williamson argues that this describes the world of efficient discrete market exchange. This is in fact the world described elegantly by Baumol, Panzer and Willig (1982) as one of contestable markets, where a leading assumption is that sunk costs do not exist for producers. Sunk costs are the same thing as specific assets.⁶

When bounded rationality, opportunism and asset specificity co-exist in an uncertain world, contracting parties must tailor organisational forms to meet particular problems. Williamson calls this form of contracting 'governance'. Governance structures are the private orderings of economic relationships that individuals create to cope with contractual hazards. These typically embody implicit and explicit features of contract. Williamson argues that the organisational imperative is:

'Organise transactions so as to economise on bounded rationality while simultaneously safeguarding them against the hazards of opportunism'. (Williamson, 1985, 32)

This is clearly consistent with profit maximisation but carries an organisational emphasis.

Williamson distinguishes between the governance and measurement branches of transaction-cost economics. The Economic Institutions of Capitalism concentrates on governance, which is concerned with organising transactions so as to encourage efficient adaptations as

⁶ Baumol recognises the different ends of this matter from which contestability and transaction-cost analysis start (Baumol, 1986).

events unfold. The measurement branch is concerned with the devices that people use to link actions with rewards and is well developed in the work of writers like Cheung (1983) and Barzel (1982). The two aspects of transaction-cost economics often interact. Ex-post costs of contracting will most often be costs of governance.

Williamson has increasingly come to emphasise the importance of asset specificity in his work.

'To be sure, asset specificity only takes on importance in conjunction with bounded rationality ... and opportunism and in the presence of uncertainty. It is nonetheless true that asset specificity is the big locomotive to which transactions economics owes much of its predictive content.' (Williamson, 1985, 56)

Williamson seems to be saying that asset specificity is of tremendous empirical significance. Certainly it crops up repeatedly in Williamson's development and application of his theory. It underlies the hostage model of exchange (Williamson, 1983) which is an important recent development. Williamson holds the neglect of asset specificity in the traditional analysis of competition in industries to lead to an 'inhospitality tradition' towards business practices (Williamson, 1985, 19). It may be enough to identify asset specificity qualitatively for it to be a useful concept in economics.

The essence of any hostage model is that individuals may post bonds to reassure one another that they will not behave opportunistically. One case arises if a producer may achieve production cost advantages by adopting a technology embodying specific assets. He may not be prepared to adopt the technology if he feels vulnerable to post-contract opportunism from his buyer (or buyers). A very simple hostage is posted if the buyer agrees to pay damages to the producer in the event of failure to

purchase (or threat thereof). It turns out that this type of hostage is vulnerable to expropriation hazards as the producer could be tempted to contrive cancellation. Hostages are more likely to be in kind, making use of devices like the creation of reciprocal asset specificity between buyer and seller (Williamson, 1985, 179).

Williamson applies his theory to a number of areas. These include the governance of relationships within the modern corporation. It is well known that this includes the study of the organisational innovation known as the multidivisional firm (M-form) which emerged in the USA in the early twentieth century in companies like Du Pont and General Motors. An M-form arises when separate operating divisions are created subject to strategic control from an head office. A part of the study of the M-form analyses the role of internal labour markets in setting incentives and that of hierarchical structures in handling information and decisions in an efficiently decomposable manner.

Less well-known areas of application include: vertical integration; natural monopoly; and antitrust enforcement (competition policy enforcement). Most recently, Williamson has called for more applications of transaction-cost methods to the study of economic relationships like franchising which lie between the extremes of markets and hierarchies (Williamson, 1985, 84). The empirical work of this study addresses just this issue.

Williamson's work is to be contrasted with technologically-based and power-based accounts of organisational change. According to Williamson, absent its costs of organisation technology is not decisive in determining contracts. Power is a badly defined concept. Furthermore, accounts of organisational change based on

power fail to explain how vested interests are often not served by organisational changes (Williamson, 1985, 124). Where it is market power that is referred to by explanations of contractual practices, the charge still stands that the notion is badly defined.⁷ With respect to all of these alternative explanations, Williamson argues that they are of secondary importance compared with transaction-cost aspects of organisation (Williamson, 1985, 129).

Williamson's work may be regarded as a body of theory aimed at systematically explaining key features of economic organisation. It succeeds in moving transaction-cost analysis forward to a position where it may be applied to a number of interesting questions concerning the spectrum of contractual relationships existing in any economy.

Some general conclusions

A body of theory has developed which contains early contributions on the nature of the firm, agency theory, and the work of modern economists like Alchian, Klein, Jensen and Meckling and Williamson. This treats the firm as a nexus of contractual relationships and is able to study contracting within and between 'firms', and between individuals. Some influence has come from writers on legal affairs.

Klein makes this point:

'The question of what is the essential characteristic of a firm now appears to be unimportant. Thinking of all organisations as groups of explicit and implicit contracts among

⁷ Except in game theory where the power of a coalition is measured by the costs it can impose.

owners of factors of production represents a fundamental advance.' (Klein, 1983, 373)

I agree and add that if this contractual perspective enables us to explain more of the features of economies using developments of economic methods of analysis, then economics is showing scientific progress in the sense of Lakatos (1977).

The similarities between many of the writers on the economics of organisation outweigh their differences on careful inspection. Recent contributions try to operationalise the idea of transaction-cost savings as an incentive for the selection of some organisational forms rather than others. Measurable dimensions for such costs are specified, especially by Williamson. This leads the way to the type of applied study undertaken in this thesis.

CHAPTER 2

The Economic Theory of Franchise Contracts

A franchise is created when one party, the franchisor, allows another, the franchisee, to use his trade name in return for a fee which may be levied in a number of ways. Associated franchise services, such as store design may also be provided. Franchising is usually associated with retailing where it accounts for around 4% of sales turnover in the UK. A product may be supplied by the franchisor although this is not essential to the relationship.

Franchising is an example of the kind of economic relationship which lies between the extremes of markets and hierarchies, according to Williamson (1985,84). It has not been widely studied in economics and yet it is an empirically significant phenomenon which is likely to provide a fruitful area for the application of ideas drawn from the economics of organisation. Before moving on to explain the purpose and methods behind the empirical study of franchising which is undertaken in this thesis, it is important to outline the existing small body of specialist economic theory which is concerned with franchising. This theory has been developed from agency theory and from ideas about asset specificity and opportunism. In the outline which follows, occasional comparisons are also made with less rigorous work which has arisen from sociological (or marketing) perspectives on franchise systems.

Franchising as a problem of monitoring and control

Rubin constructs an interesting theory of franchising which explains some of the features of franchise

relationships. First, this relationship is defined in terms of a written franchise agreement covering the sale by a franchisor of the right to market his product or service in a particular location. These contracts are held to have a number of more or less standard clauses covering the managerial assistance offered to the franchisee, control by the franchisor of methods of operation, payments to the franchisor, and termination of the agreement. Rubin constructs his theory on the assumption of profit maximisation by both franchisor and franchisee.

First, Rubin (1978, 225) discusses the capital-market argument which is commonly put forward as an explanation for franchising. It is often alleged that a firm may franchise its outlets rather than set up wholly owned branches in order to raise capital for expansion. This idea is associated with some economists (Caves and Murphy, 1976) as well as with writers outside the subject (Vaughn, 1972; Mendelsohn, 1978). Rubin argues that this explanation of franchising makes no sense unless we assume that the franchisor is more risk averse than the franchisee, which is implausible as the franchisee normally sinks most of his personal wealth into starting his business.

Rubin demonstrates his point by making the strongest case possible for the capital-raising explanation. We simply assume that franchisors cannot use normal capital markets and wish to rely on branch managers for capital.

'Even in this case, the franchisor would do better to create a portfolio of shares of all outlets and sell these to his managers.'
(Rubin, 1978, 226).

The reason for this is that the portfolio would diversify risk for the managers but impose no costs on the

franchisor. Franchisees must expect higher returns from undiversified investments if they are risk averse and this implies smaller returns for franchisors. Any capital-market advantages must come from shifting risk to the franchisee, and this only makes sense if the franchisor is the more risk averse. Rubin concludes that capital-market arguments do not explain franchising.

Rubin's alternative explanation is a direct descendant of the work of Jensen and Meckling (1976) and Alchian and Demsetz (1972) on monitoring and control within the firm. His starting point is that a theory must explain the choice of the franchising relationship by both parties and must predict key features of observed franchise systems.

Rubin observes (1978, 226) that franchising is normally used in retail networks where the franchisee is remote from the franchisor's head office. In this situation, monitoring of the franchisee is difficult and it pays to develop an incentive system which leads him to be efficient and to avoid shirking. This is achieved by a profit-sharing agreement which gives the franchisee sufficient profits to make any shirking too costly for it to be in his interests. It can be reasonably assumed following Alchian and Demsetz (1972) that the system will show more total profit if shirking is controlled so that there are gains to both parties from the profit-sharing agreement.

Franchisors will not wish to pay any more profit to franchisees than is necessary to remove any incentive to shirk. Rubin assumes that there is a competitive supply of would be franchisees who are attracted by a difference between what they could earn as branch managers in similar occupations, and the profit stream accruing to them as franchisees. The franchisor can effectively auction his

franchise territories. Franchisees would be willing to pay up to the difference between their reservation earnings (e.g. a market wage rate) and the present value of the profits available from a franchise in order to become franchisees. An example may be given in which a franchised outlet pays \$15,000 a year for 10 years. At a discount rate of 6% this profit stream has a present value of \$110,000. If a franchisee could earn \$12,000 a year as a manager (present value over 10 years = \$88,000) then he would pay up to \$22,000 for the franchise. This predicts that the franchisee will pay a lump sum to obtain title to all profits from a franchised outlet. The value of the lump sum may be stated formally as follows:

$$F = \sum_{k=1}^n \frac{R_k - W}{(1 + r)^k}$$

Where F = the franchise lump sum fee,
 $R_1 \dots R_n$ = the franchise profit stream,
 W = the franchisee's reservation earnings,
 n = the term of the franchise agreement,
and r = the rate of interest facing franchisees

Rubin implicitly assumes that the interest rate, profit stream and reservation wage facing all franchisees is the same.¹

Rubin acknowledges (1978, 227) that we do not observe franchise contracts of this kind. Instead we typically find that franchisees pay a lump-sum initial fee and a continuing royalty payment which is related to sales. In

¹ Presumably, unsuccessful bidders have higher reservation wages.

return, they do indeed claim residual profits. The most plausible explanation, argues Rubin, is that the franchisee wishes to protect himself against poor post-contract performance by the franchisor. To make a once-and-for-all payment to the franchisor would be to risk his opportunistic failure to perform efficiently the parts of the relationship which rely on him. These duties cover such things as providing managerial support and the monitoring of standards of operation throughout the franchise system. Rubin emphasises the importance which franchisees are likely to attach to system monitoring, given that they typically buy rights to a national trademark:

'What is involved is a classic externality problem ... if one franchisee allows ... quality ... to deteriorate, he benefits by the full amount of the savings from reduced quality maintenance; he loses only part of the costs, for part is borne by other franchisees.'
(Rubin, 1978, 228)

This externality is described by Mathewson and Winter (1985, 506) as horizontal free riding. It occurs where franchisees free ride on each other's efforts.

Rubin's theory is one in which franchisees buy the title to a stream of profit earnings in order to solve a monitoring problem that the franchisor would face with direct employees. A different monitoring problem, concerned with horizontal externality, is then solved by making the franchisor's income partly dependent on continuing royalties linked to sales. This income is then perfectly sensitive to cross-system effects.

Monitoring costs are undefined in Rubin's paper. It is also noteworthy that he makes no attempt to find the optimal division of the franchise payment between the fixed and variable components. Both of these deficiencies

are best described as omissions of detail rather than of substance.

The theory generates some predictions about franchise systems. The most interesting of these is that we may observe franchisors buying back their outlets as their chains become more mature and the density of outlets increases. Under these circumstances, simple at-a-distance monitoring costs become lower per outlet and it may pay to integrate fully. Simple monitoring costs would be decisive here as the externality problem may be controlled as well by full forward integration. The cited buy-back phenomenon is observed (Rubin, 1978, 229).

However, Rubin does not have a complete explanation of the franchise relationship. A major problem with his approach is that its application fails to distinguish profit-sharing employees from franchisees. Profit sharing here would also solve at-a-distance monitoring problems, at least where these concern the employee's effort level, by giving a reward which is linked to effort. The full integration that it would involve would solve any problem of horizontal externality. It is notable that Rubin does refer to franchisees as being rather like employees in his paper (1978, 230).

Another serious shortcoming of this approach is that the franchise relationship is related to the written franchise agreement. Implicit-contract aspects of the relationship are not considered. Yet we know from Macauley (1963) that business contracting tends to be governed by its long-term value to the parties: it is 'relational' and adjusts over time as circumstances change (Macneil 1978). Concentrating agency style on the ex-ante aspects of the franchise relationship may rule out important features like reputation effects which can affect behaviour over time.

Finally, it is not clear in general that horizontal externality must be solved by a residual-earning monitor like the franchisor. Whether market-based externality control is feasible depends upon the type of horizontal externality and the costs of organising particular solutions to the problem. It has been noted already that monitoring-cost details are omitted from Rubin's theory. However, the point about the nature of the horizontal externality is more fundamental. Specifically, we need to know whether we have a case of separable or non-separable externality.

Separability of the profit functions of different franchisees does not rule out an horizontal externality relationship. For example, we could have the following functions:

$$R_i(q_i, q_j) = f(q_i) + g(q_j)$$

$$R_j(q_j, q_i) = h(q_j) + k(q_i)$$

where R_i = the profits of the i th franchisee,

R_j = the profits of the j th franchisee,

q_i = the inputs of the i th franchisee,

and q_j = the inputs of the j th franchisee.

A separable externality occurs reciprocally between the two franchisees because each profit function can be written as the sum of two separate functions each containing just one of the arguments. An important corollary of this is that the externality $g(q_j)$ or $k(q_i)$ does not affect the franchisees' marginal cost functions. It is well known in externality theory (Davis and Whinston, 1962 and 1966) that it is feasible in the separable case to control externality by establishing

charges which may be levied unambiguously on an offender. Whether it is desirable to do this then depends on the costs of operating such schemes relative to other methods of control such as, in this case, the establishment of a franchisor with residual claims (Coase, 1960).

In the case of non-separable externality, cross partial derivatives like $\partial^2 R_i / \partial q_i \partial q_j$ do not equal zero and it is not possible to unambiguously link a penalty with individual franchisee effort. No expression may be found for a franchisee's external effect on another which only contains variables under his control. Under these circumstances a complex set of simultaneous adjustments is required to control the externality. It is very difficult to imagine these being achieved through individual negotiation. Some form of integration is likely to be chosen due to monitoring-cost and negotiation-cost advantages.

Rubin does not distinguish between the two types of externality. Neither form guarantees the selection of franchising as a means of control although non-separability strongly suggests some form of integrated management of the externality problem.

Nevertheless, Rubin must be credited with questioning the existence of capital-market incentives for franchising. Also, he analyses aspects of the franchising relationship which are likely to be important in a more complete theory. The franchisee's reservation earnings, externalities, monitoring costs and the need to protect contract performance from opportunistic behaviour are important aspects of franchising.



Franchising as a case of principal and agent

Mathewson and Winter (1985) attempt a rigorous analysis of franchise contracts using an agency approach which concentrates on ex-ante contracting. Their model has hidden-information and hidden-action aspects (Arrow, 1985). They see the principal ingredient in any franchise contract as the franchisee's right to use a national brand name in exchange for paying a share of profits to the franchisor. They also note that franchisors typically provide national advertising and training programs, and retain the right to terminate the franchise agreement. Quality standards are usually imposed on franchisees. Normally, the franchisee pays a lump-sum initial franchise fee as well as a profit share (as a sales royalty) to the franchisor. Mathewson and Winter try to explain the franchise relationship and some of their observations on variations in franchise contracts.

The franchise contract is likely to be incomplete according to Mathewson and Winter (1985, 504). If all decisions could be completely specified in an explicit agreement then the franchisor and franchisee could guarantee their joint profit maximisation at the start of the relationship. Since there are enforcement and monitoring costs attached to a complete contract some decisions will not be covered by it but will be undertaken later on the basis of the unconstrained self-interest of the parties. These incomplete areas will nevertheless be governed by incentive structures set in the agreement. To the extent that incomplete contracts deviate from joint-profit maximisation, they move away from a 'first-best' optimum towards the second best (agency costs are created). An important example of agency costs in franchise systems arises when the franchisee can shirk over the quality of his input. The franchisee will not capture the whole benefit of his efforts, some of which

accrues to other franchisees and some to the benefit of the franchisor. The horizontal externality between franchisees is treated much as in Rubin (1978). Mathewson and Winter also introduce a vertical externality between franchisor and franchisee.

Mathewson and Winter construct a basic model of franchising by deriving necessary and sufficient conditions for the emergence of franchise contracts. They assume that a franchisor sells the right to use his brand name in an exclusive territory. The quality of both the franchisor's and franchisee's production and marketing efforts will affect the level of output in the franchise system. There are both horizontal and vertical externalities possible in the model which emerges. In contrast to Rubin (1978), Mathewson and Winter argue that horizontal externalities (free riding on fellow franchisees' efforts) are not necessary to explain franchise contracts. This is because monitoring difficulties may arise for the franchisor even when there is only one territory: vertical externality (chiselling on the franchisor's standards) is an ever present problem. They therefore concentrate on vertical externality and regard the horizontal type as a complication rather than as a necessary condition for franchise contracts containing monitoring rights for the franchisor.

Their model contains the kind of information asymmetry which is usual in agency models. Local demand at a franchised outlet is subject to uncertainty and may take on an high or a low state. However, the franchisor cannot costlessly identify any ruling state of demand. The franchisee has better local information and may attempt to reduce the quality of his effort in high demand states and try to pass off the resulting low output as due to a low demand state. He will only do this if it is profitable for him.

The franchisee is assumed to pay a franchise fee to the franchisor which may contain a fixed and a variable component. The contract specifies this payment scheme plus the quality of the franchisee's input in good and bad demand states. Mathewson and Winter (1985, 510) derive the first-best contract and then show why profit sharing arises to replace this.

Mathewson and Winter agree with Rubin (1978) that the dominant (first-best) contract between franchisor and franchisee would be one which leased the brand name in return for a lump-sum payment, if this were feasible. The franchisor would then establish mechanisms to monitor chiselling and free riding and would have the incentive of maximising his fees to encourage efficient ex-ante contracting of monitoring arrangements. Each franchisee would pay a fee conditional on the value of the brand name and therefore dependent on the optimal amount of monitoring which reduces free riding and chiselling to levels that maximise joint profits given monitoring costs. The lump-sum sale of a franchise is not ruled out automatically by problems of monitoring the franchisor if contracts may be complete.

If it is not feasible to cover all aspects of the franchise relationship in an explicit contract, profit-sharing will emerge. In their basic model, Mathewson and Winter do not necessarily attribute this to opportunism on the part of franchisors. Rather, they introduce a constraint on the wealth of franchisees which prevents them from being able to sink a large investment into their franchise outlets. This constraint, which may be empirically relevant and which is subsequently relaxed, means that franchisors must rely on rewards rather than the penalty of termination to maintain franchisees' standards of operation. It is notable that as with Rubin

(1978), it is profit sharing that is regarded as the essence of a franchise contract.

Mathewson and Winter (1985, 512) set out their basic model using the assumption that the franchisor's levels of brand-name investment and monitoring can be completely specified in the contract. The franchisor seeks to maximise his profits:

$$(1) \quad R_f = F + (1-f) \sum T_i X_i(q_i, Q) - C(p) - Q - G, \quad i=1,2$$

- F = the lump-sum franchise fee
 f = the percentage of sales revenue retained by the franchisee
 T_i = the probability of state i , $T_1 + T_2 = 1$
 X_i = demand in state i , $X_1 = X_1(q_1, Q)$, X_2 is the better demand
 q_i = franchisee effort in state i as a cost to him
 Q = the franchisor's investment in brand creation
 G = the franchisor's sunk costs
 C = the franchisor's monitoring costs, $C = C(p)$
 p = the frequency of monitoring

Equation 1 shows the franchisor's profits on the assumptions that the franchise relationship lasts for one time period and is with one franchisee. The franchisor's profits depend upon the lump-sum fee, plus the royalty $(1-f)$ applied to the expected value of sales, and minus his costs. However, equation 1 is constrained by the need to provide incentives for the franchisee:

$$(2) \quad fX_2(q_2, Q) - q_2 - F \geq (1-p)\{fX_2[q_1(1-y), Q] - q_1(1-y) - F\}$$

where y = the proportion by which the franchisee understates demand

Equation 2 is a constraint which prevents the franchisee from chiselling. The profit accruing to the franchisee

from correctly declaring the better demand state and applying the correct effort level q_2 is shown on the left side of equation 2. This must exceed the profit from wrongly declaring the poor state and adjusting effort downward to $\hat{q}_2$ where $\hat{q}_2 = q_1 (1-y)$. The fraction y is set by the franchisee to ensure that the franchisor observes a consistent output and declared demand state whenever the franchisee is misleading him over the demand state and effort level.

Equation 3 spells out the definition of y and may be substituted into equation 2. This is another constraint on equation 1.

$$(3) \quad X_2[q_1(1-y), Q] = X_1(q_1, Q)$$

Equation 4 implies that the franchisee has a limited (i.e. zero) level of wealth as the contract must give sufficient profit to him to pay a royalty fee.² Equation 4 shows the franchisee's wealth constraint as a further constraint on the maximisation of equation 1.

$$(4) \quad fX_i(q_i, Q) - q_i - F \geq 0, \quad i = 1, 2$$

The model refers to one finite period of time and to the relationship with one franchisee. Horizontal externality is therefore excluded.

Kuhn-Tucker methods may be used to solve this model for the franchise fees, franchise effort in each state, the level of brand-name investment by the franchisor (his advertising) and the frequency of monitoring (Mathewson and Winter 1985, 513). The results are simply summarised here. In the better demand state, the franchisee's effort

² This appears also to imply that any lump sum is not paid until the end of the time period.

level maximises the joint profitability of the franchisor and franchisee by setting marginal effort cost equal to marginal revenue. In the poor demand state an effort level ($\hat{q}_1$) is set below what would be the joint profit maximising one (q_1) under complete contracting. This is because the franchisor recognises that the franchisee's incentive to chisel ($y\hat{q}_1$) in the better demand state increases as $\hat{q}_1$ approaches q_1 . The franchisor's investment in the brand sets his expected marginal revenue from this advertising equal to his marginal advertising costs. Optimal monitoring occurs where its marginal cost equals its marginal benefits to the franchisor; where marginal benefits are measured by a movement of $\hat{q}_1$ towards q_1 which increases in monitoring make possible.

At the optimal effort level and brand investment in the poor demand state, the franchisee should be just able to pay the fixed fee after paying the royalty and his other costs. The royalty should leave him better off if he reveals true demand in the better state.

Mathewson and Winter's basic model demonstrates that a precise description can be given to a profit-sharing franchise contract derived from a zero-wealth constraint for franchisees. They also consider the effect of relaxing this constraint.

The removal of the zero wealth constraint from Mathewson and Winter's model opens up the possibility that franchisees could post bonds sufficiently large to guarantee their good performance. Mathewson and Winter refer to Williamson's (1983, 1985) work in this respect. They argue, however, that this will not lead to the hypothetical first-best contract in which a franchise is sold outright for a lump sum. This is because the franchisor might behave opportunistically by, for example, misdeclaring a contractual breach so that he may escape

having to deliver any franchise services in exchange for the lump sum.

Mathewson and Winter note that franchisees typically do pay a lump-sum fee (1985, 516). They therefore ask what determines the size of this. An answer is that the expected value of the lump-sum must be outweighed by the profits accruing to the franchisor if he delivers his services. The basic condition for this according to Mathewson and Winter is:

$$(5) (1-f)[\sum T_i X_i(q_i) + H_k(Q)] - Q - C(p) \geq rF$$

where r = probability of successful cheating,
 $r(0 \leq r \leq 1)$,

$H_k(Q)$ = expected value of the brand name,
 where $k=n$ for a new and e for an
 established franchise.

This condition states that the franchisor's quasi rents (returns on contract specific investment) must exceed his expected profit from cheating.³ In the case of a new franchise, we would expect competition among franchisors for franchisees to reduce the franchisor's expected profits to zero. In that case:

$$(6) F + (1-f)[\sum T_i X_i(q_i) + H_n(Q)] - Q - C(p) - G = 0$$

The condition that the franchisor should not wish to cheat may be written (from 5 and 6) as:

$$(7) F \leq G/(1+r)$$

Mathewson and Winter's model is for one period which may be repeated. They assume that at the end of one

³ This condition appears to wrongly omit F , the lump sum, from the left hand side. Its inclusion does not materially affect the argument.

period new franchises have either made losses or profits, which is entirely consistent with the ex-ante requirement for zero expected profits. Only the profitable survive to become established franchises. These may operate for another period after recontracting between franchisor and franchisee. Ex-post profits for the first period will be expected in the second:

$$(8) \quad F + (1-f) \left[\sum T_i X_i(q_i) + H_e(Q) \right] - Q - C(p) - G = R_f$$

where R_f is profit.

It follows that the condition for no cheating becomes:

$$(9) \quad F \leq (G + R_f) / (1+r)$$

It follows from (8) and (9) that the lump-sum franchise fee can be higher in the case of an established franchise.⁴ A corollary of this is that the sales royalty will be lower in the case of an established franchise. Mathewson and Winter (1985, 517) note that unproven franchise systems typically do carry low lump-sum fees.

We may note that the lump-sum fee which is constrained by equation 5 has no bond-posting characteristics. It acts purely as a profit-sharing device. A bond is a valuable item put under the control of the franchisor by the franchisee. The franchisee loses this if he misleads the franchisor. Williamson (1983) refers to this type of practice as using hostages to support a transaction, as discussed below. In fact, Mathewson and Winter observe that no pure bond would be placed by the franchisee in their model. This is because no reputation effects are allowed for in their one-period formulation. It could be said that they should allow for this since they use the idea of repeating the period to

⁴ If F is included on the left side of (5), (7) becomes $F \leq G/r$ and (9) becomes $F \leq (G + R_f)/r$ and the conclusion still follows.

distinguish established from new franchises. Their point is that the franchisor suffers no penalty in their model if he absconds with a bond at the end of a period, so no franchisee will deposit one.

Mathewson and Winter explain observed lump-sum franchise fees in terms of what may safely be passed to the franchisor without giving him an incentive to cheat. A weakness of this approach is that it fails to account for indirect bonding effects on the franchisee which are likely to arise when lump sums are deposited. The franchisee's behaviour is not altered by his payment of any lump sum in the model. The likelihood is that his commitment to his business and his efforts would be affected by sinking some of his wealth, whatever the direct explanation for the payment. This possibility is simply ignored in Mathewson and Winter's model.

Mathewson and Winter (1985, 520) also consider a franchise contract which is incomplete as far as the franchisor's performance is concerned. They adjust their model to allow for incomplete contracting over monitoring and brand-name investment (national advertising by the franchisor). Incompleteness refers to the explicit contract; it is now the case that franchisors may establish a national advertising fund and take general monitoring powers but that no specific amount of advertising or monitoring is stated. In this setting, the optimal contract is still one that maximises the franchisor's profits whilst giving the franchisee at least his next-best alternative earnings. The contract must be self enforcing in the unspecified aspects of the franchisor's behaviour: for example, the franchisor must be left to choose the optimal advertising level.

A first-best optimum would be achieved in this model by the franchisee posting a bond to ensure his true

declaration of local-demand states and by giving the franchisor residual claim to all profits in excess of the franchisee's alternative earnings. The franchisee is penalised if he misrepresents and the franchisor receives the full marginal benefit of the levels of monitoring and advertising that he undertakes. Mathewson and Winter rule out this solution on the grounds that performance bonds are infeasible in their model: as it is for one period, franchisors can make off with any bond without incurring recruitment problems with future franchisees.⁵

As with the earlier versions of their model, Mathewson and Winter introduce additional constraints which move the optimal contract away from the first best in agency-theory terms. The method of establishing the model is to make behavioural assumptions about how the franchisor and franchisee react to aspects of the contract and introduce these into an ex-ante optimisation problem. As a result the contract is either explicit or self enforcing at the start. It is incomplete in its explicit points with gaps being filled by implicit (self-enforcing) aspects. There is no question of the contract being relational: that is to say being one which develops through time (Macneil, 1978).

The franchisor will set his advertising expenditure so that:

$$(10) (1 - f) H'_k(Q) = 1.$$

Where H'_k is the partial derivative of $H_k(Q)$ with respect to Q . This states that the retained marginal revenue from investment in the brand must equal its marginal cost to the franchisor. This immediately shows that established

⁵ In a repeat of the model there must always be a last period which would support this idea.

franchisors will invest more in the brand name for a given royalty fee (f) if $H_e > H_n$ as Mathewson and Winter assume. Equation 10 implies $Q = Q(f)$ and the optimal contract will be designed with this in mind and using equation 10 as a self-enforcement mechanism.

The basic conditions governing monitoring by the franchisor are less easily derived by Mathewson and Winter (1985, 523). The following ex-post maximisation problems are defined by assuming fixed unit monitoring costs $C(p) = p.M$ and by remembering that the franchisee may choose to chisel in the better demand state. The franchisor has the problem of choosing his monitoring intensity (p) to maximise his expected profit given that there is a possibility (d) that the franchisee may misdeclare demand:

$$(11) \text{ Maximise } T_1 R_{F1} + (1-p)dT_2 R_{Fc} + [1-d(1-p)]T_2 R_{F2} - Q - pM$$

p ,

where,

R_{F1} = franchisor's profits net of advertising and monitoring costs in the poor demand state,

R_{Fc} = franchisor's profits net of advertising and monitoring costs when the franchisee cheats,

R_{F2} = franchisor's profits net of advertising and monitoring costs when the franchisee does not cheat in the better demand state

The franchisee has a related problem:

$$(12) \text{ Maximise } T_1 R_{A1} + T_2 [d(1-p) R_{Ac} + (1-d) R_{A2}]$$

d

where,

R_{A1} = franchisee's profit in the poorer demand state,

R_{Ac} = franchisee's profit when he cheats,

R_{A2} = franchisee's profit when not cheating under the better demand state

Equation 11 may be solved to yield:

$$(13) d = H/[T_2(R_{F2} - R_{Fc})]$$

and equation 12 gives:

$$(14) \quad p = 1 - \frac{R_{A2}}{R_{Ac}}$$

Given that these results will motivate the franchisor and franchisee once a contract has been established (ex-post) equations 13 and 14 must be added as constraints to an ex-ante constrained optimisation problem similar to the basic model of equation 1.

It is, however, the constraints of this final model of Mathewson and Winter which are of interest. In particular, we may note that equation 12 contains no component reflecting rewards to the franchisee if he cheats and is detected. This implies that under such circumstances the franchise contract will be terminated with an expected payment to the franchisor of:

$$dp \quad T_2 \quad R_{r2} - Q - pM.$$

In summary, Mathewson and Winter succeed in constructing a rigorous model of the franchise contract which explains the level of royalty and lump-sum franchise fees in terms of removing the franchisor's incentive to abscond with a lump sum and of encouraging true demand-state revelation by the franchisee. This model treats monitoring as a rational profit-seeking activity on the part of the franchisor. When incomplete explicit contracting over monitoring and brand building is considered, the possibility of contract termination by the franchisor emerges.

However, this model is rather narrowly constructed. The franchise relationship does not develop over time under arrangements which can cope with unforeseen changes. The model is very much one of agency with concentration on

ex-ante contracting. It shows the sort of approach which ".... features ex-ante incentive alignments in superlative degree". (Williamson, 1983, 28).

It is also notable that the model excludes the use of explicit bonds as a means of guaranteeing desirable modes of behaviour. Also, possible commitment-enhancing aspects of lump-sum franchise fees, paid to transfer profits, are ignored. The testable implication is that commitment ordering is not an important aspect of franchise relationships. This turns out to be incorrect in the empirical work reported in later chapters.

Finally, as with Rubin (1978), there is nothing in Mathewson and Winter (1985) to explain the selection of the franchise relationship rather than an employment relationship. Some observed features of franchise contracts are explained but this does not predict when such relationships will arise.

Hostages and unfair contractual arrangements

Klein (1980) has developed some of his earlier work on asset specificity and opportunism (KCA 1978) in an explicitly contractual setting. Along with Klein and Leffler (1981) this work analyses the conditions under which implicit contracts can operate to fill gaps in complete explicit contracts between transactors. Klein's (1980) paper is of particular interest as a franchising example is used to illustrate how contractual provisions which are often regarded as unfair in the law in fact have important transaction-cost implications for efficiency. This area of study brings bond posting, or - as Williamson (1983) calls it - hostage taking, into the theoretical analysis of franchise contracts.

Klein's starting point is that completely explicit contracts, which would be fully contingent and costlessly enforceable, are not usually feasible. This is because uncertainty implies a large number of possible contingencies and because some aspects of contractual performance are difficult to measure. Individuals will often have an incentive to renege on their agreements, holding up their trading partner as explained by KCA (1978). This involves taking advantage of unspecified or unenforceable aspects of contract. Klein (1980, 357) argues that the solution emphasised in KCA (1978) will not always be observed. First, human capital may be involved so that integration is outlawed by the usual prohibition of slavery. Second, transaction costs may permit alternatives like implicit contracts which are self enforcing and rely on market sanctions for their governance. A final possibility is that cheating may be anticipated' in the design of an incomplete explicit contract.

An example of anticipated cheating shows that, where the damage which results is not too severe, the victim can simply adjust prices to protect himself. Klein (1980, 357) argues that when an employee is known to cheat a certain amount each period, the resulting damage can be deducted from his wage. Thus, if some existing employees are 10% underefficient the employer can cut their wages by 10% and hire 10% more people and not suffer any losses. Also, the employees' behaviour suggests that they prefer the leisure to extra income. The solution is one of voluntary contracting to maximise gains from trade.

Generally, such ex-ante contract adjustment cannot be used. Sometimes the cheater imposes costs in excess of his benefits upon the cheated. In that case the damage cannot be anticipated in the payment which is offered. In general, deficiencies in the quality of some service may

not always be compensated by alterations in payment quantity. Individuals will be willing to spend resources to safeguard high quality performance. The question then is how this will be achieved. In particular, what implicit contractual arrangements can be used to plug the gaps in explicit contracts and encourage high quality supply?

One method which can be used to safeguard high quality performance is to have a potential cheater post a bond which is forfeited if he cheats, argues Klein (1980, 358). This bond might arise in a less obvious form if the cheater is required to make one of his investments in an highly firm-specific form which will have a very low salvageable value. In both cases, the same purpose is served. I will use the more recent term "hostage" to refer to these bonds (Williamson 1983, 1985). I also think that it is useful to say that in the case of the obvious bond an explicit hostage is taken. When the bond is less obvious, an implicit hostage is taken. Klein (1980) uses aspects of franchise contracts to illustrate some basic principles of hostage taking.

Franchise contracts typically require franchisees to pay lump-sum fees to franchisors and to make highly specific investments in equipment. The franchisor usually takes the right to terminate the contract at will should he believe that the franchisee is not maintaining quality standards. It is a common observation that these contracts appear to favour the franchisor's interests.

'...franchisees see contracts as being weighted in the franchisor's favour. Amongst the current ... sample ... less than one in 10 franchisees saw it as weighted in their favour'. (Stanworth, 1985, 93).

An obvious question is why franchisees freely agree to such arrangements. Another question concerns the

possibility that the franchisor might try to opportunistically appropriate any hostages that he takes, as pointed out by Rubin (1978) and Mathewson and Winter (1985).

Klein (1980, 359) notes that for any hostage to be effective in deterring cheating it must set the franchisee's expected gain from, for example, free riding on the brand name equal to zero. This implies that hostages will be worth much more than the actual gain when monitoring costs are positive. In such cases, there is a probability of detection which affects the franchisee's expected gain. If there is only a 10% chance of my poor standards being detected then 10 times the gain to me must be taken from me whenever I am caught to render my expected gain equal to zero. The contract will necessarily be constructed 'unfairly' in this sense. Klein's analysis would appear to accept the possibility that hostages could impose net costs on cheating franchisees.

Reverse cheating by the franchisor is controlled by the possible increase in operating costs which he faces if he tries to appropriate hostages opportunistically. If this is communicated throughout his franchise network he will lose franchisees and find it hard to recruit new ones. Klein assumes that there is some operating or monitoring cost advantage to franchising. The present value of this cost advantage over the life of the contract acts as a hostage put up by the franchisor. As long as the franchisee's bond is greater than the franchisee's expected gain from cheating and is less than the cost penalty which is imposed on the franchisor on moving to some other organisational form, an hostage can support their relationship. This explains why franchisors cannot just insist on arbitrarily high hostages, which would certainly deter quality chiselling: the problem is that

these would prove too tempting for franchisors and franchisees would never agree to them.

The advantage of using an hostage to support the franchise relationship is that it may replace costly monitoring and enforcement with a device likely to have low monitoring costs attached to it.

An interesting aspect of Klein's analysis (1980, 361) is that he is able to explain how one side of a franchise contract will be more incomplete. It is often observed, as in the empirical work in the later chapters of this thesis, that rather more of the requirements of franchise contracts apply to the franchisee. Often the franchisee's obligations are fairly well specified whereas the franchisor's are stated in broad terms. This can be explained by noting that the franchisor may have a large investment in his brand-name capital which can act as an implicit hostage. This is at risk if he develops a reputation for poor provision of franchise services. Thus, well established franchisors will find that franchisees will trust them over the exact details of the services which they provide.

Klein is also able to explain the use of professional limitation clauses in franchise agreements. These usually prevent a franchisee from competing in his area for some period after he has left the franchise system. When a franchisee joins, his future level of skill is not known. If he becomes highly adept at his business, he may be tempted to set up on his own. The professional limitation clause prevents him from simply removing the franchisor's investment in his training into another business: this human capital acts as an hostage as it is only of value to him as a franchisee. An alternative view of the nature of this type of hostage emerges in later chapters of this study.

Klein (1980, 361) also recognises that implicit contracting enables sequential adaptation to occur within a contractual relationship. When change is required, market sanctions encourage movement towards greater efficiency.

Williamson (1983, 1985) develops the analysis of hostage taking which Klein begins. He argues that:

'not only are the economic equivalents of hostages widely used to effect credible commitments, but failure to recognise [this] has been responsible for repeated policy error'. (Williamson, 1985, 168).

The policy error occurs when hostage arrangements are seen as unfair aspects of contract or as restrictions on competition. In the UK, for example, a major reason why the Office of Fair Trading requires franchise agreements to be registered is concern over professional limitation clauses.⁶

Williamson's analysis of hostages sees them as having ex-ante screening properties as well as ex-post bonding aspects; he emphasises the latter as does Klein. He states that his hostage model refers to a finite period of time, whereas others like Klein and Leffler (1981) assume that self-enforcing contracts can exist indefinitely. Williamson's model (1985, 170) shows hostages being used to support the use of the most efficient technology when this involves specific assets which might be vulnerable to opportunistic behaviour. Klein (1980) uses hostages to support the most efficient organisation, which is not necessarily based on highly specific assets. Both Klein and Williamson posit some link between the creation of a

⁶ This is stated in a letter to the author.

hostage and the profitability of the relevant transactions.

Williamson makes some suggestions concerning likely hostage selection. In kind (implicit) hostages are likely to be less vulnerable to opportunistic appropriation by trading partners compared with pecuniary hostages (1985, 178). He does not explain this well, but seems to refer to situations where hostage taking is not attractive to its holder. This can only mean that the hostage is valued by the person who gives it but it is of no value to the person who holds it. In Williamson's terms (1985, 177) it is an 'ugly princess': the father of two equally cherished daughters is safer posting the ugly one as a hostage. Thus, a firm may be reassured into using a franchise system if franchisees make specific investments with no salvageable value if they fail but where no terminal value accrues to the franchisor either. Clearly, the franchisor has no incentive to contrive a contractual breach in order to appropriate the hostage. This type of hostage may well be virtually costless to create. It may also cause trading partners to use methods of production which would not be their first choice: for example, a franchisee might prefer a less brand-specific site design.

Klein and Williamson construct broadly similar theories of hostage taking. Both have opened up a potentially fruitful avenue of enquiry with their novel insights into apparently unfair contractual provisions. Suitable hostages may be used to support movement towards first-best franchise contracts which are ruled infeasible by Rubin (1978) and Mathewson and Winter (1985). Since franchisors do not typically sell franchises for fixed fees, however, that first-best seems largely irrelevant. Nevertheless, hostages may be found to support aspects of the franchise relationship which enhance its value to both the franchisor and franchisee. Introducing hostage theory

into the more traditional analysis of franchising offers the prospect of detecting crucial aspects of the franchise relationship which could not be duplicated by alternative organisational forms.

Motor-car distribution and vertical pricing

Up to this point in this chapter the explanations of some of the details of franchise systems which we have considered have been based, at least implicitly, on transaction-cost considerations. Those schooled in a more traditional analysis of industrial structure might argue that this risks ignoring explanations of franchise systems based on market imperfections or pricing considerations. A problem with this type of analysis, however, is that it is not likely to be decisive in explaining observable franchise systems. To illustrate this we can consider a traditional analysis of motor-car franchised dealerships put forward by Pashigian (1961). A similar analysis is used by Koo (1959).

Many franchise systems have the property that a manufacturer sells a product which franchisees must purchase from him. These are dealership-type franchises. The question which Pashigian analyses in traditional terms is why car manufacturers should choose to sell their product through a restricted dealer network rather than just sell cars to any retailer.

Pashigian (1961, 26) notes that vertical pricing considerations give no reason why a manufacturer should either forward integrate or sell restrictively if he is faced by a competitive retail market for his product. All he need do is subtract a competitive distribution-cost margin from the retail demand function for his cars and maximise his profits with respect to this. He could do no better if he restricted retail sales. In particular, if

he has a monopoly in production, this will effectively carry through to the competitive retail market without any additional measures. In equilibrium:

$$(1) \quad P = AC_D + W = MC_D + W$$

where

P = retail price

AC_D = average cost of distribution

MC_D = marginal cost of distribution

W = wholesale price

This is only satisfied when each retailer is at his point of minimum long-run average costs. If the manufacturer is a monopolist he just sets his marginal revenue (derived from W) equal to marginal production costs and sells to anyone wishing to buy at the implied wholesale price. Pashigian concludes that there must be some market imperfection in retailing which explains the manufacturer's restriction on sales which a dealer network implies.

One case which Pashigian analyses (1961, 53) arises when a manufacturer faces a known number of retailers for his vehicles each of whom has a local monopoly.⁷ The simplifying assumption of local monopoly may be relaxed in favour of other forms of imperfect competition. Pashigian claims that the theory of vertical pricing shows that as long as the dealer has some monopoly power he will always wish to sell fewer cars than the manufacturer would wish. This may lead to a phenomenon known in American markets as 'forcing' or known in the UK as 'pushing steel', where the manufacturer gives his dealers all-or-nothing choices over stock levels. It may be noted that all-or-nothing choices

⁷ Pashigian also has an account of the determination of the number of dealers (1961, 28).

are also part of the contracting in agency analysis as discussed above.

The model contained in Figure 2.1 illustrates the application of this theory. A gross profit function can be derived for each dealer by subtracting total operating costs from total revenue. Gross profit is calculated before the wholesale price of the car is deducted by the dealer. The gross profit function is shown as section (a) of the figure. Its shape would reflect long-run economies and diseconomies of scale for the dealer, and price reductions that he must make to expand his sales. Section (b) shows marginal gross profit (MGP) and average gross profit (AGP) derived from the gross profit function (vertical scales are different for the two figures). If a wholesale price of W_1 is paid for vehicles then area W_1STN shows the net profit accruing to the retailer. (AGP is the demand curve net of average retailing costs). Given that the retailer is aware of the effect of his local sales on his prices he will always set the wholesale price equal to marginal gross profits to determine his output. The marginal gross profit function is therefore the demand curve from the retailer as far as the manufacturer is concerned.

The total demand curve faced by the manufacturer is shown in section (c) of the figure. By horizontally summing the individual demand functions, the total demand function DD_1 may be constructed. Profit maximisation by the manufacturer then sets DD_2 , which is marginal revenue drawn to DD_1 , equal to marginal costs (MC). If the manufacturer just responds to the retailers' given demand function for vehicles he will produce Q_1 vehicles. The price W_1 does, in fact, then emerge.

However, Pashigian argues that manufacturers may be able to increase their profits by exerting bargaining

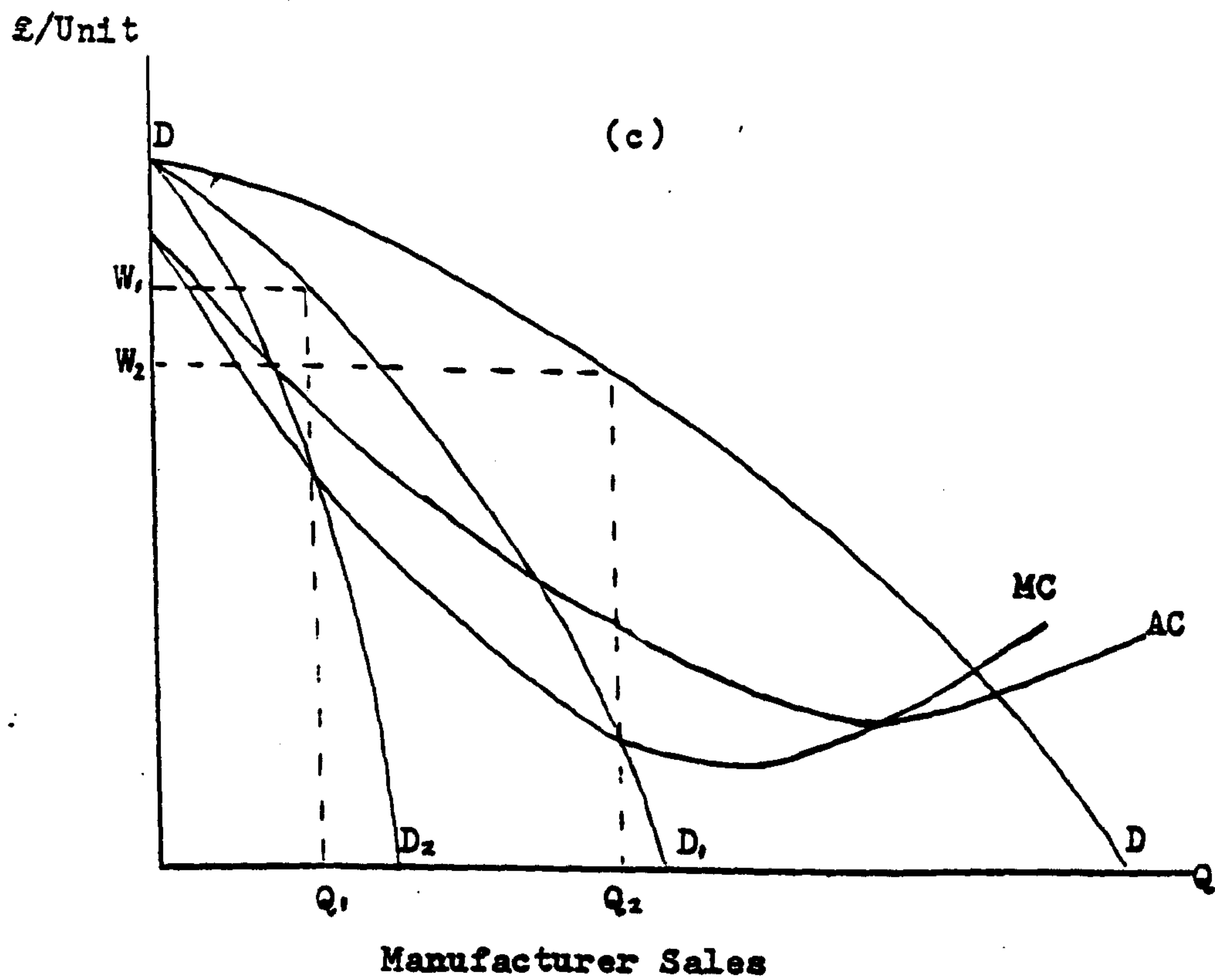
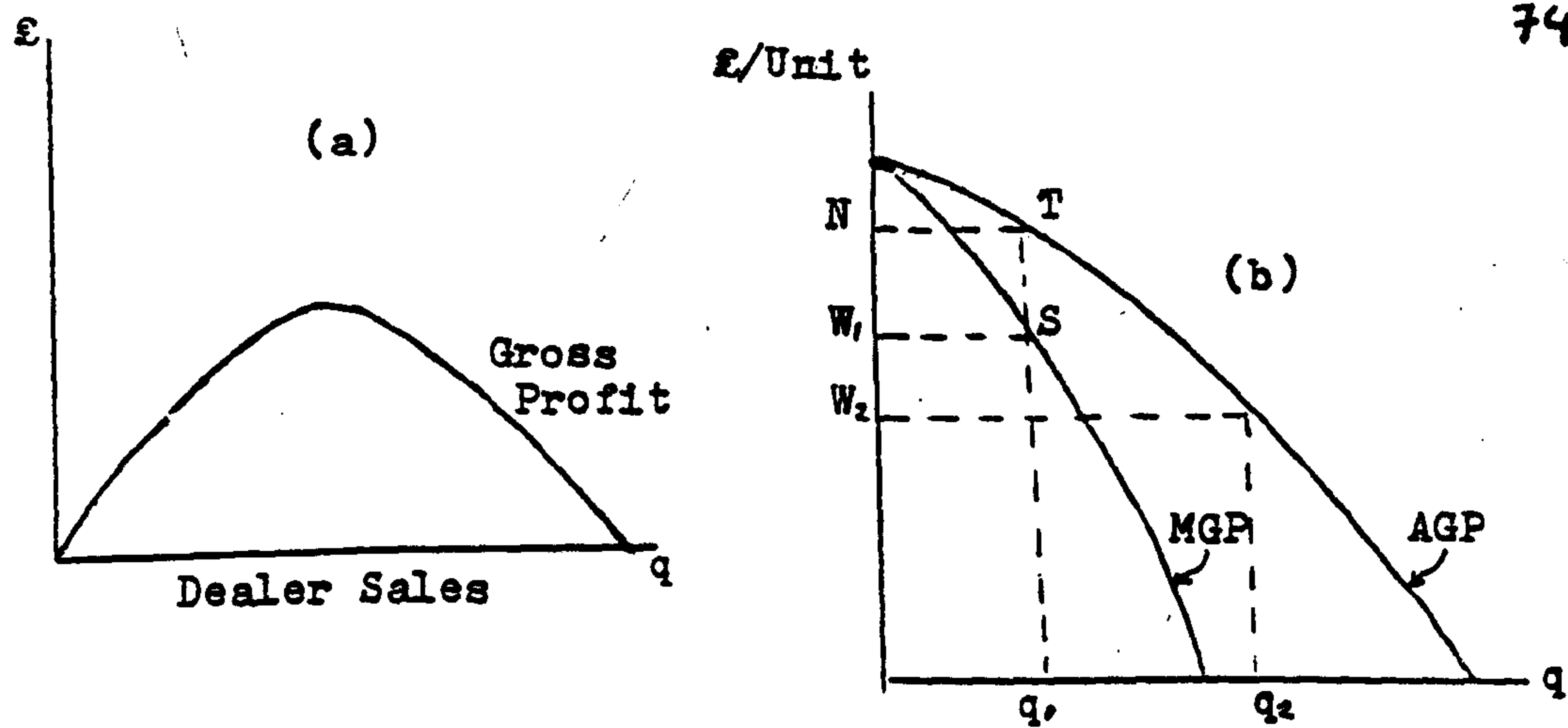


Figure 2.I A Model of Forcing

power over retailers. This happens if they can offer an all-or-nothing bargain where they set the wholesale price and a quota of vehicles which the retailer must sell. If the manufacturer can make retailers work from their average gross profit functions (AGP) in section (G) of the figure then the total demand in section (c) shifts to DD. The manufacturer's marginal revenue then becomes DD_1 , the old total demand function. The manufacturer again equates marginal revenue and marginal cost. This leads to a wholesale price of W_2 and a higher output of Q_2 . The retailer is obliged to buy a quota of q_2 vehicles at W_2 in section (b). This gives him just normal profits. It is sometimes asserted that rent extraction is an incentive for the establishment of a franchise system if this helps the manufacturer to make all-or-nothing offers (Stekler, 1962, 333).

This type of model does not give any decisive reason for the establishment of a franchised dealer network. The manufacturer could exercise his all-or-nothing option with independent retailers. In a franchise network, support will be given to franchisees, the brand will be developed and specific investments will be made. This sort of consideration does not enter the vertical-pricing type of model. Sophisticated contractual details are just ignored.

This is not to argue that vertical pricing considerations are never relevant to franchising systems. However, an interesting observation can be made. The possibility of economic rent being extracted by a manufacturer may make retailers reluctant to deal in products like vehicles, particularly as they may make highly specific investments in equipment. It may be that the difference in interests which exists between manufacturer and retailer gives an incentive to use franchising as a means of protecting the retailer. For

example, the costs of switching between retailers may be higher for the maker when there is a restricted franchised network; this may constrain the manufacturer's behaviour. This type of possibility is explored later in this study.

Towards a theory of the franchise relationship

A small body of economic theory exists which explains many individual aspects of franchising. This could be more unified if elements of hostage theory and agency analysis were combined. The agency approach is strong on ex-ante contract design but weak in relation to incomplete contracts that may need to be adaptive. There is scope for this type of unification which can draw on the more general work from the economics of organisation which is discussed in Chapter 1.

A particular problem with recent analyses of franchising is that there is little regard to the exact form of franchising contract under discussion. The empirical work reported in later chapters shows that varieties of franchising exist and that these must be considered in detail if the franchise relationship is to be properly understood.

CHAPTER 3

Methodological Aspects of the Study

Mathewson and Winter (1985, 505) note that the economic literature on agency has both focussed on observed contractual practices and has tried to characterise the precise nature of some aspects of these. It is possible to say this about the economics of organisation in general. Theory is in fact predominantly aimed at understanding observed practices. A certain amount of prediction then occurs as a result of building up explanations. It is a general impression of this author that the observation of contractual practices is a relatively casual affair, typically using very broad categories. Subsequent theorising is relatively sophisticated. There is a case for improving the care with which observations are made.

The empirical work of this thesis focusses on observable practices in UK franchising systems. Case studies are constructed which allow current theoretical insights to be checked for explanatory and predictive power and which allow new ones to be developed. The cases are constructed carefully and make use of tailored versions of methods which have been developed throughout the social sciences for the analysis of qualitative data (Glaser and Strass, 1967; Lofland, 1971; Miles and Huberman, 1984; and Reid 1986, 1987a and 1987b, Chapter 3). This is not just a matter of gathering anecdotal evidence. Rather, great care has to be taken to ensure consistent treatment across cases and to make relevant data gathering as complete as possible.

Franchising is chosen as an area for investigation for several reasons. First, it is an interesting area for the study of contractual relationships as it lies between the market and the firm. Williamson (1985, 84 and 130)

argues that attention to this part of the spectrum is likely to help our understanding of complex economic organisation; the area is currently not well understood. Secondly, specific theories of franchising cover the ground rather patchily and without sensitivity to contractual detail; Williamson (1985, 105) calls for more detailed work on contractual relations. Thirdly, franchising is young in the UK, having developed since the 1970s, so it is an interesting business development in its own right. Finally, the population of UK franchisors is relatively small so that case-study work can be fairly extensive in its treatment of relevant categories.

The principle aim of the study is analytical rather than purely descriptive. The intention is to develop theoretical insights into franchising that can add to the existing body of knowledge in the economics of organisation. Methods of case-study construction are used which are intended to lead to these theoretical insights by giving every opportunity for connections to be made between details.

The decision to gather qualitative data

Qualitative data are those which are insufficiently defined to enable quantified comparisons to be made (Reid, 1986, 5). In general, this will mean that dimensions either do not exist or that they cannot be measured. Height is a dimension that can be measured to generate quantified data. Honesty is a character trait that cannot be so easily measures; it is more likely to be indicated by certain attitudes. The economics of organisation concentrates on contractual relationships which necessarily generate qualitative data. There is no choice in this matter: if economics is to extend its scope and

explain these relationships it must contend with qualitative data¹.

Ideas like asset specificity (measured as a cost attached to relocating assets) are essentially quantitative. However, opportunism, bounded rationality, trust, flexibility and contractual incompleteness are all examples of qualitative aspects of contractual situations.

Many dimensions of transactions costs, like uncertainty and even frequency of transactions may be difficult to measure in practice and may be best treated as qualities. Qualitative aspects of the economics of organisation are likely to dominate even in theoretical work. When it comes to empirical work, given the usual difficulties attached to statistical data collection of poor records and reluctant informants, the chances are that much all collectable material will concern qualities.

It is necessary and desirable to work with this type of data. This does not imply that data collection and analysis should be anything less than rigorous. Rigour is still possible through care with detail and through the development of logical explanations of observations.

Establishing the field of study and the samples

The fieldwork which is displayed and analysed in the chapters below covers franchise relationships in 19 franchise systems. These cover a broad range of franchises in areas like fast-food retailing, vehicle hire, mobile garage services and car distribution. In most cases, well-known brand names are involved. Three franchises are relatively small-scale at the start-up

¹ Which may be represented mathematically in theoretical work by making simplifying assumptions.

phase of their life-cycle (Barstock, Garage Door and Quikframe). The franchise systems form the field or area of study.

Within the field of UK franchise systems, the sample of 19 particular franchises represents those who were willing to cooperate with the study. A first step in locating these was to contact individuals in the field who seemed likely to act as 'gatekeepers' (Burgess, 1984, 48). These are contacts who are thought able to provide access to the targeted group. Two gatekeepers were contacted by a letter which explained the project in broad terms: these were Mr. Tony Dutfield who was the Director of the British Franchise Association (BFA) and Mr. William Bryson who is the Secretary of the Scottish Motor Traders Association (SMTA).

In both cases, a letter was sent which was deliberately kept simple although not vague. It was written on University headed paper, was kept to a readable length of one side, and outlined my intention to conduct an 'independent survey of franchising practices' in the UK. The letter to Mr. Bryson was tailored to refer to motor manufacturers and traders. In both cases, I identified myself as a university lecturer in business economics and asked for their organisations' support. The letters were personally addressed and signed. I was not able to offer much of an incentive to my contacts for their participation but had to more or less appeal for their help to get started. The letters were designed to convey information in a friendly and business-like manner.

Mr. Dutfield provided some details on the BFA and a list of member franchisors and of those seeking membership. The BFA is an association of established business-format franchisors which may be joined by a franchisor meeting minimum standards over the written

contract used, number of franchisees, and length of time in business. Using the BFA list I was able to choose 25 franchisors who appeared to cover the range of businesses offered to franchisees. A further three franchises were added to the list in order to include some non-BFA franchisors in the sample. Total BFA membership stood at 72 in 1987. In addition there may be as many as 25 franchisors who are not members, most of whom run very small businesses. The sample represents a high proportion of the franchisors in the field. Business-format franchising attempts to provide every aspect of support for a franchisee's business, and typically provides manuals covering sales, operating and administrative procedures.

In choosing the 25 franchisors, the main guide was to cover the spectrum of business types but to avoid likely duplication of categories. The type of fieldwork undertaken in this study corresponds to what Glaser and Strauss (1967) call theoretical sampling. In this, the field is studied from as many angles as seem likely to generate useful information either on theoretical requirements or on the relevance of some piece of theory. This is what they call the 'slices of data' approach. Merely adding data does not automatically improve information flows in this process. In the case of the BFA list, it seemed reasonable not to study franchises which are virtually identical except for the brand name. Otherwise there would be a strong probability that redundant data would arise. Thus, Bally was approached whereas Clarke's Shoes was not. Similarly, Mobiletuning was chosen rather than Home Tune. Often there was a secondary reason for choosing one company rather than the other: in both of the above cases I thought that the newer company, with more need for brand-name exposure, would be more likely to cooperate. Judgement was used to fill the categories of well established and less well established,

large versus small businesses, and UK-wide as opposed to locally based.

Each franchisor was contacted by personal letter using the Mailmerge option on the Wordstar word-processing package designed for the IBM personal computer. The letter was so designed that it could also be used to contact franchisees at a later stage in the study. It would in each case contain a personalised address and the name of the director or senior manager responsible for franchise development (or, later, the franchisee). This contact was to provide information on the franchisor's side of the business. The franchisor acts as a secondary gatekeeper in providing access to franchisees.

A copy of this letter is attached to this chapter as Appendix 3.1. In the letter, a brief outline of the project is given, published details of the franchise and a franchise contract are requested, and the usefulness of the study is indicated. I also requested a short interview emphasising that I was well organised and would not let this run beyond 50 minutes. The response was extremely good. 17 franchisors granted interviews. Of these, only four refused to pass over copies of their franchise agreement. The sample of 17 business-format franchises was quickly established.

However, I was not so lucky with the approach to the motor trade through the SMTA. I met with Mr. Bryson who provided a list of SMTA members. This comprised most established motor traders in Scotland. My intention was to begin by contacting new car dealers rather as I contacted franchisors. I would then contact manufacturers later. With business-format franchising I worked from the top down. With the motor trade I was working upwards. The letter to motor traders was simply adapted to motor-trade terminology (e.g. 'dealers' not franchisees) and

sent to 15 motor dealers covering all makes in the Scottish central belt. I could quote Mr Bryson's support. It was my intention to use local dealers as far as possible just to economise on the transport costs of the research. Also, I had a few motor-trade contacts and felt that I would be able to pilot my research methods at an early stage.

The response from motor dealers was poor. I made a total of five cooperative contacts, including two that I had started with. The reason for this difference in response is not hard to diagnose. In the business-format case, franchisors were happy to devote time to explaining their systems to an academic researcher because of the brand-name exposure that could follow from any publications. The motor dealers had no such interest.

I therefore tried a different approach. I contacted Ford, Austin Rover, International Motors and Renault either at the relevant Scottish or UK main offices. I then treated the manufacturer just like a business-format franchisor in my study. Only Ford and Austin Rover responded and both proved willing to discuss and document their franchised dealer networks and to provide contacts with dealers. The top-down approach worked better. The poorer response from manufacturers is explained by the fact that they have little to gain from inclusion in a study compared with business-format franchisors. Manufacturers run large businesses with well established brand-names; inclusion in a study would probably have small returns to them. I was relying on their goodwill.

It was a disappointment not to obtain the participation of an importer (International Motors imports most Japanese cars). However, no other importers were approached as by this stage it had become clear that my investigations in the motor trade were concerned with a

particular kind of franchising. Distance between the manufacturer and his dealer network is not a key distinguishing feature of this category. Anyway, in many respects the international division of production within Ford covers any issues attached to car importing.

In both the business-format and motor-trade cases, the contact with franchisees which was provided by the franchisor proved completely fruitful. Each franchisor usually provided between two to four names which fitted a description of likely candidates which I provided at interview. Sometimes I was told to just choose any. I asked for franchisees whom I could reasonably easily visit, which meant preferably a Scottish or London location. I also asked that they should between them cover the range of franchisees operating in a network; small-scale operations, larger businesses, recent recruits and established franchisees all should be represented. Sometimes uniformity in certain respects suggested a smaller number of franchisee contacts. In one case (Pizza Express) the franchisor would only establish contact with one franchisee. Where franchisors also ran a network of company outlets, I asked to visit one of these. The franchisor often asked to be allowed to circulate particulars of my imminent contact. In all cases except one, franchisees agreed to participate in the study after I made contact with them by telephone. The franchisee version of the draft approach letter proved unnecessary when using this top-down approach.

Telephone contact with franchisees followed guidelines suggested by Frey (1983) and used by Reid (1986, 17) in setting up a study of small businesses. These include: identifying the caller; mentioning the field contact; explaining why the call is being made; emphasising confidentiality; mentioning any benefits of

participation; and giving a reasonably free choice on the date and time of interview.

I believe that the virtually 100% response rate is explained by franchisees' wishes to appear cooperative in any venture which has the approval of the franchisor. In this sense there is a benefit to the franchisee from participation. This is not fear of the franchisor. The complete participation extends across franchises where the franchisees were suggested to me and where I freely chose them. In some cases, I felt that franchisees thought they might check how their franchise compared with others through me, although this was not suggested or intended on my part.

It is possible that some franchisors attempted to select franchisees who they thought would provide a favourable picture of the franchise. I do not believe that any serious bias arises on this account. Plenty of selected franchisees offered criticisms of the systems that they were in. Also, I genuinely feel that franchisors looked for franchisees who fitted my description: often they explained their reasoning as they went along; nobody was evasive; eye contact was not avoided.

A complete list of franchisors along with associated data are given in Table 3.1. This shows a total of 77 distinct categories of significant relationship within the sample. There are 19 franchisors, 47 franchisees, and 11 significant company-owned outlets. Mobiletuning's one company van is run for information purposes and is not really a separate aspect of the case, this is also true of Dampcure's operations in the Watford area. One of Ford's company-owned outlets became franchised during the course of the fieldwork. Hence, the 14 company outlets give 11 separate points of contact. Barstock has no franchisees as the company is in the process of establishing a

franchise network. Two of the 47 franchisees left their networks during the fieldwork (thus filling an additional category).²

Table 3.1 Franchising sample

Franchisor	Franchisees Contacted	Company Owned Outlets	Contract Provided	Product or Service
Apollo	3	No	Yes	Window Blinds
Avis	3	Yes	Yes	Vehicle Rental
Austin Rover	3	No	Yes	Motor Vehicles
Bally	3	Yes	Yes	Shoe Retailing
Barstock	N/A*	Yes	No	Bar Stocking
Budget	3	No	Yes	Vehicle Rental
Computerland	2	No	No	Computers
Cure 30	2	Yes	No	Damp Proofing
Ford	4	Yes	Yes	Motor Vehicles
Garage Door	3	Yes	Yes	Garage Doors
Midas	3	Yes	Yes	Garage Services
Mobiletuning	4	Yes	Yes	Car Tuning
Nationwide	3	Yes	Yes	Detectives
Olivers	2	Yes	Yes	Coffee Shops
Pizza Express	1	Yes	Yes	Restaurants
Quikframe	1	Yes	Yes	Picture Framing
Youngs	3	Yes	Yes	Dress Hire/Sales
Yves Rocher	2	No	No	Cosmetics
Wimpy	2	Yes	Yes	Fast Food

*Not applicable

² Two interviews with Renault and International Motors franchisees were not used as the franchisor failed to cooperate. One interview with a Ford main dealer undertaken at a pilot stage was not subsequently used.

Instruments of investigation

Miles and Huberman (1984, 42) point out that opinion differs over the value of having any preconceptions about a particular piece of fieldwork. On the one hand, Glaser and Strauss (1967) emphasise that theoretical sampling should be able to generate leads if the researcher simply gets out into the field and begins methodical observation: an ultimate inductive method is advocated in which the researcher does not know what he is looking for at the start. The advantage of this is that certain biases may be discouraged; it is less likely that facts will be selectively ordered to fit some already-preferred theory. Miles and Huberman point out that the literal adoption of such a method could be very costly. The research problem is unbounded and the researcher risks becoming overloaded with unstructured data.

Miles and Huberman advocate imposing some preconceptions on a fieldwork exercise and then guarding against selective data gathering and the avoidance of surprise. In this study of franchising, my initial interests were theoretical anyway, which made it impossible to avoid some idea of what was of interest. The advantage of having some idea is that research costs can be cut: for example, by not devoting time to issues that have been logically demonstrated to be of little or no consequence. Capital-raising arguments for franchising are not likely to be a fruitful line of enquiry, according to this view. Some prior direction can be given to fieldwork without ruling out alternative lines of enquiry.

Therefore, this study of franchising makes use of transaction-cost concepts arising in the economics of organisation. Does a contractual arrangement involve hostage taking to protect against opportunism, for example? However, plenty of space is left for

alternative, non-transaction-cost factors to emerge as important in the study. This is achieved by designing in opportunities for surprise and for known alternatives to show through. It seems only sensible to start from a basis of already-established coherent theory, with an awareness of gaps that it has and accepting that alternatives may be relevant. The approach is supported by Kirk and Miller (1986, 51) who argue that non objectivity is not a problem unless what may turn out to be a partial view of some issue is assumed to be the global view.

Miles and Huberman argue as follows:

'... knowing what you want to find out leads inexorably to the question of how you will get that information.'

(Miles and Huberman, 1984, 42)

This is indeed the next question on the list. However, the connection is not as important as Miles and Huberman suggest. They proceed wrongly to argue that the prior direction of a field study will determine the instruments of investigation that will be used. Instruments are such things as tape-recorded interviews and questionnaires. Thus, if a study investigates local crime the field worker would do things like interview suspects at the local police station. A criticism here is that there are many other things that could also be done. Also, having no idea of what is looked for, the field worker could still select instruments for his widespread observations.

This study of franchising does make use of prior instrumentation in the narrow sense. The technique of analysis on which the case studies are based is the semi-structured interview. In this, an interview agenda is followed which focusses data collection but which leaves room for unanticipated points to develop. This technique

is economical, encourages consistency between separate cases, and corrects for researcher bias. It also allows the quality of data to be assessed. Was an interviewee particularly earnest in pressing a point? If so perhaps it carries more weight than a contradictory but disinterested response, or it may suggest a vested interest.

Alternative instruments were felt to suffer from drawbacks from the point of view of this study. Participant observation was thought to be too costly to undertake. Questionnaires can suffer from low participation rates and from misunderstood questions. If questionnaires are administered, they can make data gathering rather too determined in relation to theoretical sampling. Given the time devoted to interviewing (around 90 hours plus much more travelling and debriefing time) it was not thought to be feasible to attempt a multi-method approach. In addition, the author had built up interviewing experience in previous studies and was happy with this method.

An interview agenda was constructed with the aim of getting out into the field as soon as possible. Decisions had to be made early about how data would be gathered. One possibility was to tape record interviews and to consider methods for reducing the data so collected. Reid (1986) considers the disadvantages of this. Respondents may become inhibited if they believe that their answers are being recorded. A further problem may be that redundant data are unselectively recorded. Also, the fieldworker may tend to become careless in listening to respondents, believing that everything is going on to tape. The quality of data may suffer as a result of taping according to Reid (1986, 32).

Set against the disadvantages of taping is the advantage that the accuracy of recording cannot be disputed, at least in a narrow sense. Also, the interviewer is not distracted by notetaking. These advantages were not felt to be sufficiently valuable to outweigh the disadvantages. Furthermore, I felt that methods based on note taking were more likely to encourage attention to contextual detail on my part. In taking notes during an interview, opportunities continually present themselves for recording details of the environment and of any events which are occurring. These may be important: is a respondent as well treated by customers as he claims; is he distracted by events?

The interview agenda is given at the end of this chapter as Appendix 3.2. It has six main headings and an additional three minor headings. The main enquiries concern the respondent's background, the nature of the franchise, the contract between franchisor and franchisee, the perceived advantages of franchising (to the franchisor and to the franchisee) and the fees and returns accruing in a particular case (which are partly quantitative data). In each interview, enquiries were also made about the popularity of the franchise and about further contacts in the field. Time was also left to pursue any unanticipated points. An early version of the agenda was used with one of the motor dealers. This piloting exercise very quickly refined my ideas about detailed areas of enquiry and enabled the agenda in Appendix 3.2 to be produced.

The agenda has a number of key features. First it may be used with both franchisees and franchisors to encourage consistency in questioning. Secondly, it uses deletion options where questions are likely to produce yes or no answers; section 6.3, for example, asks whether sales targets are set (for a franchisee). The interviewee can still elaborate on these points, but the deletion device

may make note taking easier. Importantly, questions are not rigidly set: each main section contains probes. A probe is a heading covering an area of enquiry. Precise questions are formulated, more naturally, in the interview itself. It is this probe-based method that leads to the description 'semi-structured interview'.

The probes exhibit two important aspects: they are often nested and always coded. Nesting refers to the hierarchy which has been imposed on the probes in Appendix 3.2. Probes nest within each other as they move from the general to the more specific. Coding is used to indicate the hierarchy and to facilitate subsequent data handling. A good example of the coded nesting occurs in the second section of the agenda. This example is shown in detail in Table 3.2.

Table 3.2 Coded and nested probes

<u>Code</u>	<u>Probe</u>	<u>Hierarchical position</u>
2	Nature of franchise	General category
2.4	Who pays what?	Focussed on financial relationship.
2.4.3	Specialised assets	Specific financial issue
2.4.3.1	Stockholding costs	Focussed on an example

The example demonstrates the move from general questioning on the nature of the franchise through to questioning about a specific example of one of the financial relationships. This level of organisation enabled interviews to proceed in a businesslike manner and meant that subsequent data handling was not just based on unstructured case notes. Where subdivision did not seem useful (for example 1.1 on company history) it is not

used. The pilot exercise and some early interviews suggested the exact form of the hierarchy.³

The probes attempt to provide an interview structure which respondents could be expected to follow. Some are deliberately included as a means of relaxing the interviewee and gaining his involvement. Section 1 gathers a lot of background material which is likely to be interesting but which may not generate important findings about contractual relationships. The point is that everyone can relate to such questions, trust and confidence is built, and successful completion of the exercise is encouraged.

Otherwise, the probes reflect key issues that have arisen in the economics of organisation and in the analysis of franchising. There are probes on contracting practices which seek information on implicit and explicit contracting, monitoring and enforcement. The advantages of franchising to each party are sought by probes covering such things as risk sharing, specialisation economies, maintenance of product or service standards, monitoring economies, and brand-name sharing. Franchise fee schedules are investigated in Section 6.

An important part of the structure of the agenda is the extensive use of open probes. No hierarchical category is closed down. This shows in the use of an open probe simply labelled 'other'. In Section 6 on the fee schedule, four two-digit probes (6.1 - 6.4) cover the royalty payments, lump-sum fees, tied-in sales and the possible fixing of sales targets which might be used to

³ Early interviews lead to some minor changes in the probes. Mostly this was a matter of realising what were sensible ways to ask questions; for example, stockholding costs became nested within specialised assets.

transfer economic rent to the franchisor. These items might not exhaust all possibilities so a fifth category, labelled 'other' (6.5) is added. Lower down the hierarchy, enquiries are similarly not sealed off. Thus, within 6.1 which covers the royalty the possibilities exist for this to be linked to sales (which is usual) or to profits. Codes 6.1.1 and 6.1.2 cover these possibilities. Code 6.1.3 is labelled 'other': this allows for a royalty type not envisaged at the start of the fieldwork. In fact, an important unforeseen royalty-payment method did emerge as one franchisor was found to use a fixed weekly royalty (Nationwide Investigations). By being careful not to close off avenues of enquiry, prior instrumentation can avoid biases which are feared by fieldworkers who prefer more freestyle methods.

I am later able to show that important elements of the franchise relationship, particularly concerning the nature of lump-sum payments, came to light during the fieldwork. This is attributable to the deliberately chosen open-ended probe structure. The probes allow lateral and vertical movement towards unforeseen possibilities. They do not guarantee such findings; if these are present in a case skills of detection and luck are still required. A particularly helpful respondent may make a great difference to the findings. Also, the fieldworker's skills of being able to listen well and to spot clues are important.

Carrying out the interviews

Franchisors and franchisees were interviewed at a time and place of their choice. The interview agenda was printed so that space remained for making notes. These were generally extremely rough and made extensive use of abbreviations. Notes were made of impressions of business sites and of any aspects of the respondent's business

practices or behaviour that might have a bearing on the data being provided. The respondent's raw data were recorded in normal prose notes. Any observations of mine were printed in capitals within square brackets to avoid confusion (e.g. [SEEMS BUSY]).

Interviews were enjoyable experiences. Most people were very happy to discuss their franchises and we built up a good rapport. I was self-disciplined over keeping to the interview time of 50 minutes. Very occasionally, things over-ran but this was usually due to an interviewee's zeal in discussing and showing me his business.

The interview agenda worked well. It lent just enough structure to proceedings to keep things moving. It also lent authority to my interviewing without intimidating respondents.' The agenda document was used just as it is given in Appendix 3.2; the use of University headed paper as a front cover helped to reassure respondents that they were participating in a proper academic study.

Most of those interviewed spontaneously offered to give more time to the project if this would help. Interviews took place over a 12-month period.

Data handling

As soon as possible after conducting an interview, it was written up as fully as possible as a short prose note. The codes proved invaluable in doing this debriefing and saved a lot of time by giving headings for future quick reference. They also enabled cross references to be made. A typical section of one of these notes is shown on Table 3.3. It arose from an interview with Mr. D. Peacock, who is the Managing Director of the Quikframe picture-framing franchise.

Table 3.3. Extract from interview note.

3. Contract

- 3.1.1 The 20-page franchise agreement was described as 'very general'. This is being overhauled as there have been problems (franchisor wants tied-in sales and an head-lease arrangement). [THE WISH FOR TIE-INS SUGGESTS A WISH TO CONTROL MATERIALS QUALITY OR POOR FRANCHISOR RETURNS. THE CONTROL OF FRANCHISEES' LEASES COULD BE TO STOP THEM LEAVING]. The agreement has a seven-year term.
- 3.1.2 The franchisor thought that 'informal understandings are dangerous'. However, he mentioned franchisees' expectations of growth in public awareness of the brand name.

The interview notes proved invaluable. As pointed out by Reid (1986, 32) recall of field events develops with practice. However, it is important to minimise reliance on memory when different parts of a case study may be constructed at different times. In my study, economy dictated that I interviewed participants in order of their geographical locations. A typical day's fieldwork might be to interview two franchisors and an unrelated franchisee. When I later interviewed at other sites forming part of these distinct cases, it was important that data could be matched up.⁴ The code system and diligence in debriefing proved essential parts of building cases.

⁴ In this study a case refers to one franchise system and a site refers to one franchisor or franchisee. Site is often used as a synonym for case: see Miles and Huberman (1984, 78).

Data reduction and presentation

Miles and Huberman (1984, 21) point out that some data reduction is an inevitable part of analysing written-up case notes. This refers to the next logical step in handling the data. Once whole sets of case notes are assembled, a decision has to be made about how to present and use the data which has been collected. I decided to write each case using a traditional narrative prose style. In doing this, I tried to consciously minimise data loss so as not to compromise the theoretical-sampling method. However, we should note that some data loss is automatically implied by the decision not to tape record interviews.

Data reduction means that some data are lost as the fieldworker selects and analyses the case material he intends to use. I decided not to write up cases in an highly analytical manner. Instead, I concentrated on analysing within-case issues and on providing a consistent descriptive treatment of each case at this stage. The 19 individual case studies comprising the second volume of this study are the result. The primary aim of these is to give as full a picture as is possible of the franchise systems, as a basis for subsequent theorising in an organisational framework.

Some analysis does take place in the case studies of Volume 2 wherever issues arise which either seem peculiar to a franchise or which establish possibilities for further analysis. In this narrow sense, the case studies are single-case analyses.⁵ Thus, the impact of head-lease arrangements on franchisees, where they are required to

⁵ 'Within-site' analyses in Miles and Huberman (1984, 79).

lease premises through the franchisor, is analysed in the cases. In addition, the possible purposes behind these arrangements are also discussed. However, the main purpose of the written-up case studies is to provide high quality data which are analysed in the remainder of this first volume of the study.

The data are not greatly reduced from the case-note stage in Volume 2. In two respects they are augmented as I have added an analysis of the franchise agreement (written franchise contract) to the case in instances where franchisors provided copies of these, and as subsequent recall of details was often valuable. The point is that data losses are kept to a minimum right up until the last minute, when analysis begins. This maximises the chance of drawing theoretical inferences from the cases. It also provides case material which is interesting in its own right in a reasonably pure form.

Written-up cases follow a uniform narrative style. Main probes (one digit) are used as section headings. In the interests of ease of reading codes are not used within the text. Inverted commas (") mean only one thing in Volume 2 and that is recorded interviewee comments or quotations from franchise agreements.

Franchisees are given codes (e.g. H₂) in the cases of Volume 2. This is to protect the confidentiality of the data that they furnished. Place names and other details are blanked out where this could aid identification. All franchisees are similarly treated.

Within-case analysis and case comparisons

Within-case analysis and case comparisons⁶ represent the analytical heart of this study. Most of the remaining chapters of this volume comprise them. The aim here is to feed the empirical findings back on to existing theory in the economics of organisation in general, and the economic theory of franchising in particular, and to provide a basis for modifying and extending existing theory. Within-case analysis refers to the understanding of contractual relationships within a particular franchise. Case comparisons assess the manner in which different franchises contractually support similar or different types of transaction.

A major question in all of this is whether insights into the economics of organisation based on transaction-cost reasoning can lead to explanations of the details contained in the cases. Alternatively, is some other theory suggested, such as one based on monopoly power rather than efficiency considerations? Whatever theoretical view emerges, can predictions as well as explanations be offered?

The analytical chapters which follow are microanalytic in the extreme. This means that they focus on individual contracting practices rather than on averages. Some of the ideas which may be drawn from guides to qualitative data analysis suit such an approach admirably. In particular, in looking at case material for connections between details it has been urged that 'outliers' may be important in verifying conclusions (Miles and Huberman, 1984, 237). An outlier is an aspect of a case which does not fit in with an overall trend that has been noted. An example arises in this study when one franchisor uses a unique royalty system (Nationwide Investigations). The

⁶ Case comparisons are 'cross-site' analyses in Miles and Huberman (1984, 151).

reasons put forward for this are entirely consistent with the observations on most franchises.

Nevertheless, often it is widespread practices which prove to be of interest. Klein makes this point:

"When all firms in a particular industry use similar contractual provisions, it is unlikely to be the result of duress or fraud and should not be considered as evidence of collusion. Such uniformity suggests the existence of independent attempts to solve an important common problem and signals the presence of a prime research prospect." (Klein, 1980, 362)

Klein is clear here that he is thinking of many independent but similar responses rather than of some kind of average trend that might mask variety.

Within-case analysis and case comparisons allow the useful outliers and the common issues to be detected.

Miles and Huberman (1984, 215) make a number of useful suggestions to aid the drawing of conclusions from qualitative data. In many respects these amount to an exhortation to the fieldworker to be honest with himself and diligent in pursuing his detective work. Apart from the use of outliers, they suggest the following practices: looking for rival explanations; looking for negative evidence; attempting replication; triangulation; and seeking intervening variables. Discussions with other researchers can be useful in this process. Looking for rivals implies more than the requirement for an open mind which is attached to data gathering; alternative theoretical explanations must be actively attempted. Then, when the field of possibilities appears to have narrowed to indicate a theory, it can be useful to attempt its rejection by looking for counter examples. Replication refers to the practice of seeing whether other

researchers can come to similar conclusions based on similar data. By triangulation, Miles and Huberman refer to checking to see whether different instruments of enquiry produce similar conclusions. Intervening variables are ones which can feature in causal links between apparent connections (or disconnections) although they may be initially hidden.

Summary and conclusions

The empirical work on franchising comprising the remainder of this volume is based on detailed case-study methods which assess contracting practices in terms of their transaction-cost and other properties. This work is firmly based within a tradition of attempting to explain observed contracts which has been noted within the economics of organisation (Mathewson and Winter, 1985).

The case studies are reported in Volume 2 with a long-term aim of providing a record of contemporary franchising practices which will be used beyond this study. This is the data base on which the work of the rest of this volume depends. Some of the recent guidelines which have been established for qualitative data analysis, in particular by Miles and Huberman (1984) are incorporated into the study. The within-case analysis and case comparisons which follow are informed by these.



UNIVERSITY OF EDINBURGH

Department of Business Studies

WILLIAM ROBERTSON BUILDING, 50 GEORGE SQUARE, EDINBURGH EH8 9JY.

031-667 1011 ext.

Telex 727442 UNIVED G

101

[Address and date]

Dear [contact's name]

I am undertaking an independent academic study of business-format franchising in the UK. In particular, I am trying to gather data on the advantages of franchising to both franchisors and franchisees.

I am writing to ask for your help in completing this study. Could you possibly send me any written details of your franchise such as an information pack or a draft contract?* I wonder if you would also be prepared to grant me a short interview of approximately 50 minutes in order to discuss your franchise? My questions are well structured and should not waste time. Any information will be treated confidentially where this is required.

The completed study should provide a useful impartial review of franchising practices for participants. It will cover matters which are not raised in any existing study.

I have discussed my study with Mr Tony Dutfield of the British Franchise Association, who has expressed support for it.**

Thank you in advance for any help which you may give to me. I look forward to hearing from you.

Yours sincerely,

A Dnes
Lecturer in Business Economics

* The request for data is omitted with franchisees.

** Mr Bryson (Scottish Motor Traders' Association) is cited in a version sent to motor dealers.



UNIVERSITY OF EDINBURGH

102

Department of Business Studies

WILLIAM ROBERTSON BUILDING, 50 GEORGE SQUARE, EDINBURGH EH8 9JY.

031-667 1011 ext.

Telex 727442 UNIVED G

Franchisees are usually party to arrangements which tie them to one franchisor. These may fall short of business format franchise schemes. Occasionally, more than one franchise is held. The project analyses the contractual relationships involved in franchising.

Franchise Contractual Arrangements.

Interview with _____

Franchisor/Franchisee

Date _____

Interviewer _____

Appendix 3.2 Interview Agenda

Appendix 3.2 (cont.)Outline of Project.

Franchisees are usually party to arrangements which tie them to one franchisor. These may fall short of business format franchise schemes. Occasionally, more than one franchise is held. The project analyses the contractual relationships involved in franchising.

Appendix 3.2 (cont.)1. Background Questions

Probe on (PO)

1.1 Company history1.2 Current markets

Appendix 3.2 (cont.)1. Background Questions (Cont)1.3 Involvement with other organisations

1.3.1 Franchisor/Franchisee

1.3.2 Suppliers

1.3.2.1 Components

1.3.2.2 Materials

1.3.2.3 Finance

1.3.2.4 Other

1.3.3 Employees

1.3.4 Subcontractors

1.3.5 Multiple/Fractional franchises Y/N

1.3.5.1 Organisational economies of scope

1.3.5.2 Production economies of scope

1.3.6 Other

Appendix 3.2 (cont.)2. Nature of Franchise

PO

2.1 Franchisee obligations2.2 Franchisor obligations2.3 Franchisor involvement in franchise format

2.3.1 Full business-format franchising Y/N

2.3.2 Sales

2.3.3 Service

2.3.4 Parts

2.3.5 Training

2.3.6 Finance

2.3.7 Administration

2.3.8 Other

Appendix 3.2 (cont.)2. Nature of Franchise (Cont)2.4 Who pays what?

2.4.1 Advertising

2.4.2 Training

2.4.3 Specialised assets

2.4.3.1 Stockholding costs

2.4.3.2 Other

2.4.4 Other

2.5 Transfer of franchise

2.5.1 Allowed Y/N

2.5.2 Subject to franchisor approval Y/N

Appendix 3.2 (cont.)3. Contract

PO

3.1 Method of contracting (spectrum)

3.1.1 Explicit

3.1.2 Implicit

3.1.3 Mixed Y/N

3.2 Enforcement by

3.2.1 Franchisor

3.2.2 Franchisee

Appendix 3.2 (cont.)3. Contract (Cont)3.3 Monitoring by franchisor/ franchisee

3.3.1 Results observed Y/N

3.3.2 Inputs observed Y/N

3.3.3 Daily conduct observed Y/N

3.3.4 Cost

3.3.5 Other

3.4 Known cases of dispute

3.4.1 Cases

3.4.2 Common causes

3.4.3 Dispute resolution

Appendix 3.2 (cont.)4. Franchisee Advantages/Disadvantages

PO

4.1 Why not go it alone?

4.1.1 Brand name important Y/N

4.1.2 Risk bearing by franchisor

4.1.2.1 Franchisor behaviour
when market is bad

4.1.2.2 Franchisor debt factoring Y/N

4.1.2.3 Other

4.1.3 Other advantages/disadvantages

Appendix 3.2 (cont.)4. Franchisee Advantages/Disadvantages (Cont)4.2 Factors checked when seeking franchise

4.2.1 Financial commitment

4.2.2 Financial help from franchisor

4.2.3 Observation of other franchises

4.2.3.1 Franchisor's model franchise Y/N

4.2.4 Other

4.3 How was franchise obtained?4.4 Territorial limits Y/N

Appendix 3.2 (cont.)5. Franchisor Advantages/Disadvantages

PO

5.1 Why not fully integrate or freely sell?

5.1.1 Product debasement (avoidance)

5.1.2 Specialisation economies

5.1.3 Franchisee entrepreneurship Y/N

5.1.4 Risk bearing by franchisee

5.1.4.1 Franchisee behaviour
when market is bad

5.1.4.2 Other

5.1.5 Other reasons

5.2 How did franchisor decide on franchising?

Appendix 3.2 (cont.)5. Franchisor Advantages/Disadvantages (Cont)5.3 Franchisor operates own franchises Y/N

5.3.1 Operates in geographically concentrated areas?

5.3.2 Characterised by other links

5.4 Monitoring costs

5.4.1 Compared with forward integration
or open sales

5.4.2 As proportion of total costs for franchisor

5.5 Perceived capital-raising advantages Y/N

Appendix 3.2 (cont.)6. Franchise Fees and Returns

PO

6.1 Royalty

6.1.1 On sales

6.1.2 On profits

6.1.3 Other

6.2 Lump-sum payments6.3 Sales targets set Y/N6.3.1 Forcing Y/N
(franchisee's profits higher at lower sales)6.4 Tied-in sales Y/N

6.4.1 How is transfer price set?

6.5 Other6.6 Returns higher than normal?

Y/N/franchisor/franchisee/both

Appendix 3.2 (cont.)7. Franchise Popularity

PO

7.1 Is there a queue of potential franchisees?

8. Contacts

PO

8.1 Suggested names

9. Unanticipated Points

CHAPTER 4

The Nature of Franchising

Existing theoretical treatments of franchising assume that systems are broadly similar to each other. The empirical work of this study shows this to be unjustified. In this chapter we analyse some of the key features of the sample of 19 UK franchisors which was described in Chapter 2. The features are reported in detail in the case studies of Volume 2.

We are interested in what franchisors achieve with their systems and what franchisees receive in return for their fees. It emerges that there are three principal types of franchise which differ in terms of the products or services linking the franchisor and franchisee and in terms of the payments flowing between them. In addition, the sample contains a unique system (Case I, Nationwide) which is an interesting outlier. Broadly, the sample divides into product, brand and specialised-input franchises.

In a product franchise, a manufacturer appoints dealers to sell and possibly to service a product which he makes. In a brand franchise, the franchise relationship does not revolve around a product but simply around the franchisee's use of a brand name which the franchisor has developed. In a specialised-input franchise some specialised service other than the provision of a brand name is at the heart of the system. The commonest type in the sample is the brand franchise.

In all franchise systems, instead of supplying a product or service directly or distributing through any willing outlets, a franchisor contracts with his

franchisees over payment schedules, maximum prices, system controls, the maintenance of corporate image, and quality standards. The franchisor and franchisee have obligations towards each other which would not be duplicated by transactions on open markets nor within a system of full forward integration. Williamson (1981, 1548) calls this an example of obligational contracting.

The three main franchise systems identified here take us beyond a distinction which is commonly made between business-format franchising and dealerships. In the business-format case, franchisors offer a total business model which they claim allows an inexperienced franchisee to set up from scratch. Dealerships correspond to product franchises in terms of the analysis developed below. Such claims are not made for dealerships; the manufacturer of a product literally appoints sales agents who usually are seen as being already experienced in some line of work. In fact, there is evidence in what follows that this conventional distinction is of limited use.

Many of the differences between the systems revolve around the payments or fee schedule which is used. The fee schedule is a statement of the payments flowing between the franchisee and franchisor as a result of the franchisor's supply of services to the franchisee. Generally, payments flows are from the franchisee to the franchisor, although sometimes they are reversed if a sales-bonus system operates. Derivation of a fee schedule is a central aspect of principal-agent analysis, of which the paper of Mathewson and Winter (1985) is an example. In this chapter, the fee schedules used by the sample are illustrated and partly analysed. Further analysis of fee schedules is undertaken in the following chapter.

Product franchises

When a manufacturer appoints a specialist retailer to sell his products, usually in an exclusive sales territory, a product franchise or dealership is created. The maker normally restricts the other sales lines which may be carried by his dealer. Traditionally, the franchisor provides support over his product and not much else. The franchisee has sales targets and provides service facilities for the product if these are relevant. The earliest franchise systems fall into this category, which in modern times includes petroleum companies, vehicle manufacturers and computer manufacturers. The Singer Sewing Machine Company practised this type of franchising as early as the mid-nineteenth century in the USA (Jack, 1957).

The product franchisees in this study are listed in Table 4.1 along with some of their features. They are the two car makers, Austin Rover (AR) and Ford (Cases R and S) and Apollo, Bally and Yves Rocher (Cases A, H and Q) from among business-format franchisors. We will consider the car makers first.

Table 4.1 Characteristics of product franchises

	Lump Sum (£'000)	Tied Sales	Sales Targets	Sales Bonuses	Business Format	Service Aspects
Austin Rover	No	Yes	Yes	Yes	No	Yes
Ford	No	Yes	Yes	Yes	No	Yes
Apollo	4	Yes	No	Yes	Yes	Yes
Bally*	No	Yes	Yes	Yes	Yes	Yes
Yves Rocher	7.5	Yes	Yes	Yes	Yes	No

* Bally also charges a royalty on sales of 3%

Both AR and Ford have major shares of the UK new-car market (around 18% and 25% respectively in 1987). Their dealers can run businesses based entirely on sales of the single make and are required to do so. In both cases, minimum vehicle stocks, parts stocks, and service levels are specified by the manufacturer. Second-hand car sales are not only permitted by the makers but act to support new-car sales, allowing the customer to trade in his old vehicle. In fact, one of the principal advantages to the makers of using dealerships rather than forward integrating into their own retailing ventures is that dealers have specialist car-trading skills, according to views expressed in Case R and Case S.¹

Territorial exclusivity is practised by both Ford and AR. In the Ford case, this is very precisely defined: dealers may not have more than five dealerships; these must not be contiguous; and no dealer can advertise outside his territory. Ford will not sell products to retail customers except through its dealers. The AR practice is similar but discretion is exercised over the number of dealerships which may be held. Dealers may sell vehicles to customers who are located anywhere in the UK in both cases.

Territories may help to keep distribution costs to a minimum. If they are chosen to provide a market which will support local scale economies in retailing then dealers can afford to pay higher wholesale prices for their products compared with a system of more numerous and smaller-scale retail outlets. However, territorial limits are common to most forms of franchising, and to systems of forward integration into retailing, and are not a distinguishing feature of product franchises.

¹ This is further discussed below. Alfred P. Sloan, former President of General Motors, also held this view (Sloan, 1986, 279)

The distinguishing feature of a product franchise is simple: a product is made by the manufacturer and distributed by a network of franchised dealers. All franchises in Table 4.1 show these product transfers which appear as compulsory (or 'tied-in') product sources from the franchisee's viewpoint. The obligations of the dealer are seen in terms of selling and supporting the product. Those of the manufacturer are seen in terms of developing the product and giving dealers sales and technical back up over it. The Managing Director of S₂ (a Ford main dealer) puts this clearly in his interview: the dealer's obligation is to provide a quality local sales and service operation within the context of the maker's sales targets. He expects Ford to ensure that the product is competitive with others.

Dealers are rewarded by profits made on selling products and associated services after meeting the wholesale transfer prices set by the maker. They have every incentive to be cost conscious since all cost savings accrue to them, given a transfer price. The maker can be confident that the retail network is encouraged to run at minimum cost levels. This is important to him as it means that he has the maximum scope for extracting pure profit from the retail end of the business, as we saw in Chapter 2 where Pashigian's (1961) model of vertical pricing was considered. The maker will, not extract more than retail economic rent (return over and above that necessary to keep factors in their present use) or he risks losing his dealers. Recruiting new ones would be costly to the manufacturer. Bonuses may be used to restore dealer profitability if transfer pricing mistakes are made and this becomes too depressed. Mr. Tilley, Ford's Regional Director for Scotland and Northern Ireland points out in Case S that dealer returns are monitored for just this possibility. Bonuses are used to shift economic rent back to dealers in a discriminating way to ensure

that profit maximisation coincides for both parties. There will then be no allegations of forcing as is true of the current period compared with the early 1980s in the case of both AR and Ford. It is notable that makers always set sales targets for their dealers. As explained in Chapter 2, forcing occurs if a dealer is made to sell beyond his profit-maximising level as a condition of participating in the franchise system.

Any economic rent shifted from the retail end of the business may become a normal return on the franchisor's assets including his brand-name (i.e. reputational) capital. This depends upon the alternative returns available to him from using his assets elsewhere. Car dealers comment in their interviews that current pricing structures do not leave them high returns. One Ford dealer, S, even tolerates losses on his new-car sales because of servicing work which is generated. Car makers may be profitable, as in Ford's case. AR has made losses for some years. Car making is not widely regarded as an highly profitable activity. It is likely that normal returns (at best) are achieved by makers partly through shifting economic rent from the retail end of the business.

AR and Ford use transfer pricing to ^{obtain} their business returns from new-car sales. No lump-sum payments from dealers or royalties on sales are used. Dealers have every incentive to keep their costs down since savings imply increased profits for them. The makers' profits are linked to product sales giving an incentive for advertising, product development and dealer support. In these simple terms, the arrangement appears to be broadly incentive compatible. Sales bonus schemes are used to overcome the need for forcing by increasing compatibility.

At this point we may introduce a slight complication. Three business format franchises in this study have product franchise (dealership) characteristics as their dominant features. In the cases of Apollo, Bally and Yves Rocher, a product is made by the franchisor and this is made the centre of the franchise relationship. In addition both Apollo and Bally have servicing aspects to their operations (blind and shoe repairs) although these are simple in comparison with vehicle maintenance.

Details of the three business-format product franchises are recorded in Table 4.1 alongside data covering the two car makers. Apollo, Bally and Yves Rocher rely mainly or entirely on transfer prices for products to generate the franchisor's income. Furthermore, franchisees from Bally and Apollo report that sales bonuses are used; these link product discounts to sales targets. Sales targets are set less formally than with car dealers but are used in two of the cases.

There are some differences compared with car dealers: Bally levies a percentage royalty on franchisees' sales turnover; Apollo and Yves Rocher charge lump-sum initial fees to franchisees. The three systems have these features which are usually found with the other franchise types discussed below. The lump-sum and royalty payments are not the main means of generating the franchisor's income, however, and tend to be smaller sums than the ones used in other types of franchise systems. Franchisees believe that Apollo and Yves Rocher spend the lump sums in starting their businesses.

Neither is the business-format nature of the Apollo, Bally and Yves Rocher systems a major distinction compared with the car dealers. It is clear throughout Cases R and S that AR and Ford provide high levels of guidance, covering sales, administrative and operating procedures, for their

franchisees. These are at least equivalent to the provisions of franchisors using the business-format label. The same is true for training; there is more support from the car makers, if there is a difference. It is also clear from the interview with S₃ that Ford has a scheme to set up new franchisees from scratch through its Dealer Development Scheme. Ford even gives financial support, according to S₃.

The crucial aspects of product franchises, or dealerships, is that tied-in products are transferred and that the fee schedule revolves around their prices. This is not to argue that other considerations are irrelevant but only that they are less distinguishing. All franchisees interviewed in the five cases of Table 4.1 state that their major benefit from belonging to their networks comes from the brand name in some form or another. However, this is not a particularly distinguishing characteristic as the same claim is made by virtually every franchisee in the study. In product franchises, the brand is associated with tangible products.

Brand franchising

Brand franchising arises when the principal feature of a franchise system is its brand name. The franchisor builds a national reputation for the provision of a standardised service. It is then the right to sell the branded service that is sold to franchisees. The sale of branded products is not a key feature of brand franchises although products may often be supplied to franchisees either for their convenience or to maintain service standards. The business model followed by franchisees contains much which is aimed at standardising service levels throughout the franchise network.

A good example of brand franchising is the system run by Wimpy (Case F) which is an established chain of walk-in restaurants. In the early 1980s, Wimpy launched a new counter-service restaurant network using franchisees. It already had a well known brand name due to its existing chain of table-service restaurants, which were all franchised. The existing network was beginning to become somewhat run down at a time when eating habits were changing. Wimpy decided to move over to an American-style fast-food format, aiming the service at young people in population centres where they are concentrated.

Mr. McGlashan, Wimpy's Scottish Franchise Manager states that the Wimpy system of operation and the promotion of the network are geared to reassuring a customer that he will be buying the same thing regardless of location. When a franchisee operates a counter-service restaurant, he knows that he benefits from a brand name which conveys this information to passing customers. The information may be particularly valuable to travellers who are in a strange town and lack knowledge of the quality of local restaurants (e.g. in choosing food for children). Information transmission of this kind is the principal result of investments in creating brand image (Demsetz, 1982).

As in the case of Wimpy, a brand franchisor normally must run some outlets and become heavily involved in national advertising before franchisees are attracted to the network. Wimpy opened flagship counter-service operations which it owned and ran prior to selling franchises in the new network. Regular advertising of the chain is undertaken using a variety of media. Wimpy is a successful brand franchisor showing growth in its restaurant business; not all brands become strong as some examples in this study show.

Once franchisees are taken on in a brand franchise, monitoring systems are needed to ensure that standards are maintained. Franchisees must not horizontally free ride on the reputation of other franchisees, nor vertically free ride on the brand-name investment of the franchisor. Thus, in the case of Wimpy an essential aspect of the brand image is that restaurants are kept very clean. An operations procedure is followed where cleaners work continually in a restaurant to achieve this. Wimpy makes regular visits to franchised restaurants and, among other things, checks this standard of operation.

Table 4.2 identifies the franchises in this study whose characteristics are predominantly of the brand type. Details are also given of aspects of the fee schedule which are discussed shortly. Note that not all of the franchises possess major brand names: Olivers (Case L) is well known but is not a major force in chain restaurants; Garage Door (Case B) is a small franchise with just four franchisees; Cure 30 (Case N) is well known only in the south east of England; and Quikframe (Case M) is an unsuccessful franchise which has lost its franchisees possibly due to a failure to build up brand awareness. Nor should it be thought that there are no other aspects to these franchises. The important point is that they illustrate a type where franchisees enjoy benefits which flow directly from trading services under the brand name. They are rather like dealers but the franchise relationship does not centre around a manufacturer's products. The table identifies nine such franchises in the sample. Tied-in sales are recorded in the case of Olivers even though they result from an informal understanding (Case L).

The brand franchises use a fee schedule based on a fixed percentage royalty which the franchisor levies on the sales turnover of franchisees. Franchisees must feel

confident that franchisors will at least maintain and preferably develop the brand image. A percentage sales royalty links the franchisor's rewards to his success in promoting the brand. We can also regard the franchisor's efforts in training the franchisee as a form of brand investment. Sales royalties are always uniform percentage charges on sales revenue and are not generally associated with sales bonus-scheme payments or with sales targets. Separate advertising levies are generally not charged by franchisors even when they are specified in the franchise agreement. Olivers charges a 1% advertising levy which is recorded as part of the royalty in Table 4.2

Table 4.2 Characteristics of brand franchises

	% Sales Royalty	Tied-in Sales	Sales Targets	Lump Sum (£'000)
Wimpy	8.5	Yes	No	10
Garage Door	5	Yes	No	2.5
Avis	10	No	Yes	15
Young's	10	Yes	No	10-15
Pizza Express	4	Yes	Yes	12.5
Budget	10	No	Yes	c20
Cure 30	15	Yes	No	9
Quikframe	12.5	Yes	No	6
Olivers	7	Yes	Yes	10

Linking the franchisor's rewards to sales performance through a fixed percentage sales royalty leads to a problem. The franchisee may alter his level of sales from the one which the franchisor prefers. This happens if the royalty influences the profit maximising level of output for the franchisee, which is normally the case. As Ozanne and Hunt (1971, 46) point out, it is generally more efficient if a franchisor shares directly in the

franchisee's profits as a fixed percentage profit levy will not alter the maximum-profit level of output. However, profit royalties are never observed as Table 4.2 suggests. The reason is that profit is altered by cost elements which a franchisee could inflate to keep profit for himself. Profits are hard to monitor relative to sales levels.

All cases in Table 4.2 use a sales royalty which is a constant percentage regardless of the level of sales. It is also the same for all franchisees within a given system. This uniformity between franchisees is not a prediction of current franchising theory as it is developed in the work of Mathewson and Winter (1985). We return to this point in the next chapter, where the expectations of franchisors and franchisees are shown to be important in explaining uniformity.

Table 4.3 Lump sums in brand franchises

	<u>Franchisor's comments</u>	<u>Franchisee's comments</u>
Wimpy	Half of costs	Franchisor spent fee
Garage Door	Aim to cover costs	Franchisor spent more
Avis	Aim to cover costs	Franchisor spent more
Young's	Aim to cover costs	Franchisor spent more
Pizza Express	Profit target = 20%	Franchisor made profit
Budget	Cover costs on average	Franchisor spent more
Cure 30	More than spent	Franchisor spent more
Quikframe	No view	No view
Olivers	Aim to cover costs	Franchisor spent more

The table also shows that all of the brand franchisors charge lump-sum initial franchise fees. These range from a modest £2,500 (Garage Door Company) up to around £20,000 (Budget) although very different opportunities are being offered in the various franchises. Table 4.3 analyses the reasons advanced by franchisors for these fees together with franchisee's comments; costs refer to launch costs incurred by the franchisor.

Lump-sum payments do not amount to a large proportion of the business returns accruing in the franchises. Sales turnover for a Wimpy counter-service restaurant is around £500,000 a year, to take an example. Wimpy's sales royalty is £50,000 of this and we know from F₁ that the franchisee is left with a profit which is just better than a financial sector return on his investment. The franchisee's profits are probably at least £60,000 a year. Against all of this, Wimpy charges a lump sum of £10,000. Even if this were not spent by Wimpy at the start, it is not significant compared with returns. This picture is general, as far as can be judged given the reluctance of many respondents to discuss their profits and revenues.

It is possible to find a consensus of franchisee views in each case. As with product franchises, franchisees generally believe that franchisors spend at least the lump sum in starting them up in business. Franchisors do not always claim this; for example, Mr. Dell of Pizza Express candidly states (Case G) that he has a target profit rate on the lump sum, and this is supported by G₁ who is the franchisee interviewed in the case. Pizza Express is unusual. More often franchisors aim to cover their launch costs with a franchisee, at least on average. Sometimes they incur initial net launch costs: Wimpy can spend around £25,000 on setting up a franchisee and only charges a lump sum of £10,000. Neither the franchisor nor franchisee in the case of

Quikframe (Case M) expressed views on this matter. The Quikframe franchisor sold back his interest as franchisor in the franchisee's shop; this implies that the lump sum did not extract all expected economic rent.

Not all franchisees expressed views on the lump sum in the sample of brand franchises. Those that did usually had thought about the matter. Sometimes they had even made careful calculations: C₃ even had his accountant check to ensure that his lump sum was fully spent. Hardly any franchisees, apart from G₁ believe that their franchisors make profits from the lump sum. When they do, they do not suggest that substantial economic rents are transferred. Given that lump sums are usually fixed amounts, this is still consistent with the franchisor more or less covering his costs of setting up a new franchisee.

Lump sums are closely related to set-up costs for the franchisor. Existing theory does not predict this. Mathewson and Winter (1985, 519) use the lump sum as a means to extract economic rent from franchisees which is partly related to set up costs. Klein (1980) makes no connections with these costs in his hostage-based analysis of some features of franchising. Klein (1980, 359) instead links lump sums to the expected value to the franchisee of free riding within his system: the lump sum acts as a hostage against this. Table 4.3 shows a clear regularity which is also supported by interviews with the franchisees in Apollo and Yves Rocher. We return to this issue in the next chapter, where a development of Mathewson and Winter's (1985) model is used to explain this regularity.

All of the franchisors in Table 4.2 except for Avis and Budget make use of tied-in sales. The question therefore arises whether transfer prices on tie-ins feature significantly in providing the franchisor's

income. Table 4.4 gives details of the tie-in systems used by the seven brand franchisors concerned. The table shows that tie-ins are used to safeguard standards of service associated with a brand name. This is true for Wimpy, Young's, Pizza Express, Cure 30 and Olivers. In each of these cases there is an advantage to controlling inputs.

Table 4.4 Tied-in sales in brand franchise systems

	Add-on Charges	Franchisor's Reasons	Franchisee Comments
Wimpy	'normal'	Standards/ Buying Economies	'1/2 price'
Garage Door	5%	None	Increased Cost/ Controlled Lines
Young's	'small'	Standards	Product Good
Pizza Express	'small'	Standards	Product Dearer
Cure 30	Yes	Guarantees	Necessary
Quikframe	Yes	None	No Advantages
Olivers	5%	Standards/ Buying Economies	Cheaper

Cure 30 guarantees the workmanship of franchisees, and chemical companies guarantee materials. By specifying a company and insisting that invoices go through head office, Cure 30 can be sure that liability is clearly defined. Franchisees do see the need for this. With Wimpy, the intention is to provide uniform standards nationally. In addition, buns, meat and other items are very attractively priced for franchises: F₁ comments that many lines are half price compared with alternative sources. Wimpy passes on economies of scale and the tie-ins would be self selected by franchisees. Young's is in the fashion business and a part of the service associated

with the brand is the exercise of buyer's judgement each season. This basis to the brand image requires control over franchisees' purchases. E₃ gives the clearest statement of the benefits from this in Young's. Pizza Express is rather like Cure 30 in protecting standards by insisting on purchase control. Although G₁ complains that he could purchase equal quality at lower prices, the existing system may be the cheapest method of ensuring system-wide quality adherence.

The Garage Door Company appears to increase franchisees' costs with tied sales for no apparent reason other than providing some profits for the franchisor; in fact, tie-ins also act as a check on true declaration of sales values. The relevant cost increase is in the form of administrative work arising in sending orders through head office according to B₃. The franchisor may be worried about losing franchisees in the longer run if they established their own links with suppliers once they have learned the business. B₁ and B₂ comment on the franchisor's close guard of his supply lines. The Quikframe franchise had no apparent reason for tying in inputs other than distributing profit to the franchisor. Franchisees very quickly saw that there were no advantages to belonging to the franchise system. Mr. Peacock may have been distorting his franchisees' costs.

In the Olivers system, franchisees are allowed to buy away from company supply lines if a case can be made. Otherwise the position is as with Wimpy: franchisees would self select the supplies due to their cheapness, according to L₁ and L₂.

As for the charges levied by franchisors for tie-ins, Table 4.4 suggests that these are modest. Where franchisors are prepared to describe their charges they refer to 'normal returns', 'small profits' or quote a

specific handling charge of 5% on invoice values. Franchisees either benefit from the practices or complain of modest nuisance costs.

The prices of tie-ins are not likely to be used to extract economic rent from a franchisee. This is because his input use will be distorted as input prices change if his production techniques are in any way variable.² Cheaper inputs will be substituted for artificially dearer ones and real production costs may rise. In a brand franchise, some inputs, like labour, are not under the franchisor's control. Substitution is possible: a Wimpy franchisee, for example, could use labour to reduce materials waste below economically justified levels. Distorted pricing also increases incentives to break with quality standards by buying substandard materials from the elsewhere. System costs are likely to rise if the franchisor takes his income in this way. Where tie-ins give a cost advantage it is best passed on. Where they are a means of surveillance or standards enforcement only, costs should not be raised more than is essential to run the system.

Product franchises transfer a branded good for which the franchisee cannot substitute other inputs. Therefore, input distortions do not arise from profit-shifting transfer pricing in the cases shown in Table 4.1.

Avis and Budget have no particular reason for tying inputs. Franchisees must buy quality vehicles but can purchase identical ones from many sources. Poor quality would be instantly noticeable (old vehicles or cheap imports). The hire companies do exactly what they need to do: they make supply lines as cheap as possible and leave self selection to ensure that proper quality vehicles are

² See Plott (1966)

bought. This is outlier evidence that tie-ins are for standards control in brand franchises where input quality may vary.

Sales royalties are used in brand franchise systems as the principal means for generating the franchisor's income. The royalty links his performance of brand services with his rewards in an economical manner which reassures the franchisee that services will be delivered. There are not the ambiguities that could affect profit based calculations, and nor are there input distortions which would follow input price variation. Compared with sale of the franchise outright for a lump sum, the royalty economises on bounded rationality by avoiding the need for uncertain returns to be estimated at the start of business; it also leaves the franchisee reasonably protected against opportunism on the part of the franchisor, who might be tempted to abscond with a large lump sum.

From the franchisor's viewpoint a fixed percentage sales royalty implies that the franchisee is rewarded by a profit residual after paying costs including the royalty. The franchisee therefore is motivated to operate at his lowest possible cost function (in an envelope sense) as cost savings become profits for him. The brand franchise works much like the product franchise in this respect.

Table 4.2 gives some weak support for the possibility that franchisors may use sales targets as a means of moving franchisees to the franchisor's preferred level of sales: Avis, Budget, Pizza Express and Olivers use these. The case-study material generally reveals that franchisees are exhorted by the franchisor to achieve these targets. Positive incentives exist in that franchisees may hope for additional franchises (i.e. to become multiple franchisees) if the franchisor approves of their

performance. Negative incentives exist in that franchisees may fear non-renewal of their franchise agreements; this can have implications for their wealth as we analyse in later chapters. If the royalty is not the primary means of controlling the franchisee's sales then it may be applied arithmetically to divide profits after optimal sales have been established. The royalty must then become the subject of bargaining or be imposed by the franchisor. It will still provide the franchisor with an incentive to develop the brand.

Brand franchises may be distinguished by their reliance on fixed percentage sales royalties as the principal means of rewarding the franchisor. The main service provided by the franchisor is the maintenance and development of a brand image which increases the franchisee's sales. Franchisees can be put into a situation where they earn normal business returns by appropriate choice of a royalty figure and a lump-sum initial fee. The royalty payment can influence the franchisee's production level in relation to the one which the franchisor desires. Brand franchisees sell services rather than products. Any tied-in sales are usually a means of supporting quality standards but may have other advantages such as low cost for the franchisee.

Specialised-input franchises

A specialised-input franchise is one where the franchisor supplies an identifiable specialised input other than the brand name or a branded product to the franchisee. The franchise relationship is based on such an input in four cases in this study. These are Mobiletuning, Midas, Barstock and Computerland (Cases D, K, O and P). The classification is, as in the case of product or brand franchises, a question of emphasis. Each of the four systems has brand-name aspects. However, case

detail suggests that other specialised inputs are of greater importance.

Table 4.5 records the principal features of the specialised-input franchises in this study. Each system charges a lump-sum fee (or plans to charge one). Fixed percentage sales royalties are also a feature of this group. In the three systems which are already operating, the franchisor sells products which the franchisee may choose whether or not to purchase: there are no tied-in sales. Neither sales targets nor sales bonuses are a feature of the franchises. The sales royalty recorded in the table for Midas includes a 6% levy towards the franchisor's advertising costs. Computerland's royalty includes a 1% advertising levy.

Table 4.5 Characteristics of specialised-input franchises

	% Sales Royalty	Voluntary Purchases	Sales Targets	Lump Sum (£'000)
Mobiletuning	10	Yes	Not Enforced	3
Midas	12	Yes	No	10
Barstock	Planned	Not Planned	Not Planned	Planned
Computerland	4.5	Yes	No	13 - 20

Mobiletuning is an excellent example of this type of system. Franchisees all report that the major benefit of belonging to the network is the training provided by the franchisor. This covers initial instruction in the electronic tuning of vehicles and continuing up-dating. Training has become very important in recent years as vehicles have adopted sophisticated electronic ignition and carburation systems. Interviews with motor dealers reinforced this view of training in the motor trade. Mr.

Rowntree, who established Mobiletuning, is aware that training is the most important aspect of his business: 'On their own, they would just wither on the vine'. From the franchisee side, D₂ tells us that, particularly at the start of his business, he encountered vehicles that he could not have turned without the continuing technical support which he receives.

Mobiletuning has brand-name aspects although neither the franchisor nor any franchisee emphasises these. Some extra business is likely to be generated by the recognition customers give to the brand. Also, Mr. Rowntree runs a parts service for franchisees which is usually found to be the most economical source at least after allowing for the convenience of dealing with one supplier. Nevertheless, both of these advantages could disappear and yet we would still find this system operating. 'It is doubtful whether the current level of brand awareness or the parts service could support the franchise relationship on their own.

Barstock is included in the list of specialised-input franchises because Mr. Watson intends to start franchising based on the advantage attached to his computerised bar-stocking system. This franchise is scheduled to begin recruitment in late 1988. Mr. Watson believes that his systems have cost advantages in providing a fast service. He has copyright in the computer programmes that he uses. It is possible that others could create a similar system and it may be that Mr. Watson has little to offer franchisees.

Computerland is an interesting example of specialised-input franchising. Until 1986, Computerland charged an 8% sales royalty and provided its supply lines more or less at cost to franchisees. It is probable that Computerland expected to build up a brand presence

comparable to its American one. The franchise was initially launched with an emphasis on its brand name. However, the system has not developed as intended. Franchisees perceive their main advantages to lie in the bulk-buying undertaken by Computerland. The franchise has become increasingly orientated towards this bulk buying and the associated product discounts which are passed on. Both P_1 and P_2 tell us how the franchisees have required Computerland to link its profits to its skills in this area. In 1986, the royalty was reduced to 4.5% after representations from the franchisees and Computerland is now expected to draw more of its returns from its transfer pricing.

Midas is a borderline case. It has brand-name aspects which are relatively important. However, franchisees state (see K_1 and K_2) that their main benefits come from the parts discounts negotiated by Midas. In many ways, Midas is one side of a border and on the other is Olivers. Both companies provide discounted supply lines which are highly valued by franchisees. In both cases, the brand name is not very strong in the marketplace, at least at present. However, Olivers exercises more control over supplies as a means of influencing standards probably because food inputs are subject to wide quality variations. Midas only exercises a right to withhold approval on parts' sources; quality is easier to assess and preserve in the case of vehicle parts and supply lines need not exist primarily for standard maintenance. As supplies are used more to support the brand image in the case of Olivers, it lies in the brand franchise camp. Midas supplies are relatively more concerned with economy and it is therefore better classified as a specialised-input franchise.

Table 4.6 analyses the supply of products in all four cases. As we have said, Midas, Mobiletuning and

Computerland sells products which franchisees are free to choose whether to purchase. With Mobiletuning discounts on parts are not the major benefit to franchisees. The franchisor candidly admits that he makes a profit on these sales; however, prices are competitive with alternative suppliers. With Midas, discounts are the main benefit but franchisees have recently insisted that the franchisor should drop a handling charge of 5% and allow them to deal directly with suppliers, according to K₁. In the case of Computerland, we have already noted that there is a profit element in transfer prices. In all of the cases, prices must be competitive with alternative sources, as products are not tied in. High prices would cause franchisees to switch their suppliers.

Table 4.6 Product sales in specialised-input franchises

	<u>Franchisor's Comments</u>	<u>Franchisee's Comments</u>
Mobiletuning	Small profit element	Not major benefit
Midas	No profit element	Forced end to handling charge
Barstock	None Planned	Not applicable
Computerland	Franchisees prefer a profit element to higher royalty	Major benefit

All of the franchise systems in this section either charge a lump-sum initial franchisee fee, or, as in Barstock's case, intend to charge one. The lump sums are analysed with the aid of Table 4.7 which shows franchisees' and franchisors' views on initial fees. As with product transfers, it is broadly possible to find a consensus among franchisees within each system over whether or not the franchisor makes profits on the lump sum.

Table 4.7 Lump sums in specialised-input franchises

	<u>Franchisor's Comments</u>	<u>Franchisees' Comments</u>
Mobiletuning	Aim to cover costs	Franchisor spent fee
Midas	Aim to cover costs	Franchisor spent fee (K_1) Franchisor made profit (K_2)
Barstock	Aim to cover costs	Not applicable
Computerland	Modest profit	Franchisor made profit

The comments of franchisors and their associated franchisees are consistent providing Midas's intention refers to covering costs on average. Since lump sums are constant, if set-up costs differ between sites then profits can be made on some sites and losses on others with no total profit accruing. We should note that the variation in Computerland's initial fee is due to the size of the market being served. Computerland admits to a modest profit on set-up costs and it is possible that more economic rent is transferred at the start in larger and more profitable markets.

As with brand franchises, lump sums are fairly close to set-up costs for the franchisor. This is true even where profits are taken this way as in the case of Computerland. Table 4.7 supports the similar findings with product and brand franchisees. The possible links between lump sums and the franchisor's set-up costs are investigated in the next chapter.

Tables 4.5 and 4.6 show that the continuing part of a specialised-input franchisor's rewards comes from a sales royalty except in the case of Computerland, where a

decision has been made to link rewards to buying efficiency. We may interpret Mr. Rowntree's comment that he makes a small profit on product sales to mean that he seeks a normal business return reflecting his opportunity cost of capital.

A sales royalty has the same advantages to a specialised-input franchise that it possesses in the brand-franchise case. The franchisor's income may be linked to his performance without distorting costs and as events unfold. Profit royalties are not used for the same reason that they are avoided by brand franchisors. Profits can be obscured by strategic costing by franchisees.

It is still true in this section that the use of transfer pricing to extract economic rent is inefficient where the franchisee can substitute inputs for the transferred product. This is true in the case of Mobiletuning and Midas where alternative suppliers exist for products. The franchisee may turn to these freely. If the franchisor artificially inflates his prices then the franchisee may adopt dearer products causing the franchisor's profits to fall. Tying sales in would not solve this problem: franchisees would have an incentive to buy elsewhere, so that monitoring costs would be incurred, and they might be able to substitute labour or other inputs in an inefficient way for the materials in question. A Mobiletuning franchisee might waste time cleaning sparkplugs when this is not really justified. In Midas's case, the evidence from K₁ suggests that administering handling charges was not justified by the returns. Generally, it is better to equip franchisees to be as competitive as possible and then to extract economic rent in a manner not likely to disrupt production incentives relative to the franchisor's plans.

Mobiletuning provides training as its principal specialised input anyway. A levy on sales provides the suitable incentive for the franchisor to perform this function just as in the case where it is principally the brand name which is being provided for the franchisee.

Computerland is a little different. The franchisees prefer to have buying skills rewarded by profits on products which they are free not to take. This is like agreeing to reward the franchisor with every pound he saves franchisees on market prices. It is not, on the surface of things, a recipe for efficient operation. However, Computerland supplies branded goods to franchisees even though they are produced by outside suppliers. There are no possibilities for substitution within the franchisee's production function (outside supplies of the same goods must be used and labour cannot be substituted for computer supplies). Consequently, as long as transfer prices are below alternative market prices no input distortions need occur.

However, the Computerland system still uses a small royalty of 4.5%. This may be explained by noting that there are brand franchising elements to this system. If the franchisor's services to brand development were charged to transfer prices, then there could be input distortions as prices rose above those charged by alternative suppliers. This difficulty arises whenever economic rent above product cost savings is extracted.

The Computerland franchisees were dissatisfied with the franchisor's performance as a buyer at the time they pressurised for the move to basing his profits more on efficiency in this area. They now have a system which will never disadvantage them relative to open market prices. The franchisor can still extract virtually all of any economic rent arising at the retail end of the

business. This is due to his lower prices and may be captured by the transfer price upper limit of the market price for his supplies.

Specialised-input franchises are based on the performance of a specific function by the franchisor (other than promoting the brand, or making and promoting a product). In this study examples cover a specialised training function, specialised computer software, and the transmission of bulk-buying economies. Unless input substitution is irrelevant, transfer pricing will not be used to extract economic rent. Generally, we observe products being transferred at cost prices and sales royalties providing the franchisor's returns. Lump sums are used in these fee schedules and reflect the franchisor's set-up costs for a franchisee as in other franchise types.

Nationwide Investigations

Nationwide Investigations (Case I) does not fit in to the categories used above due to a peculiarity in the fee schedule used. The case is best treated as a useful outlier which helps to illuminate some of the transaction costs which affect the choice of fee schedule details.

Nationwide has brand and specialised-input aspects. Franchisees I₁ and I₂ both state that the brand name is of value to them. I₃ implies in his interview that this was true at the start of his business although he now feels able to operate independently. Independent operation is discouraged for all established franchisees by a professional-limitation clause in the agreement which prevents continual use of his client list if the franchisee leaves the Nationwide system. I₁ and I₂ emphasise the role of this negative incentive in keeping them with Nationwide. Professional limitation is analysed

in a later chapter. The point here is that there is a brand aspect to this franchise.

The specialised input is training. I₁ and the franchisor both emphasise the value of this. I₂ and I₃ began as already-trained operators having been previously employed in the business by Mr. Withers (the franchisor). There is evidence from Mr. Withers's interview that he has had difficulties in retaining franchisees. These have probably arisen where training has been completed and the franchisee feels able to work on his own, perhaps changing area to satisfy the professional-limitation clause in the agreement.

The above suggests that we should find Nationwide making use of a royalty which is based on sales turnover along with a lump-sum franchise fee.

Mr. Withers does charge a lump sum of £7,500 which, in line with other franchises, has links with his costs of starting up a new franchisee. I₁ has considered this matter and believes that his fee was spent on the training and equipment which he received. I₂ and I₃ would have required less training as they were already in the business. Training costs probably account for very little of the fee since training is carried out on the job at Mr. Withers' London office. The franchisee pays these costs separately with his own labour. Each franchisee receives a standard start-up package consisting of manuals, computer equipment, and computer software. There is no reason for these costs to vary between franchisees, so the initial fee appears linked to costs. There may be a small profit element as equipment tends to be secondhand.

This franchise is one which has been running for long enough for franchisees to meet contract-renewal conditions. A renewal fee equal to the initial fee is

charged. From I₂ and I₃ it is clear that this is regarded as a re-entry payment charged to the franchisee for continued access to his client list: if he leaves, the he cannot trade in his old area for one year. No equipment is provided by Mr. Withers at this stage. Nevertheless, this contribution to Mr. Withers' income is still linked to set-up costs, since the renewal fee does not generally exceed the cost of establishing a new recruit. In the case of I₃, a special negotiation occurred over creating an additional right of renewal.

Nationwide differs from other franchises in its royalty, which is charged as a fixed weekly payment of £150 (1987). The reason for this is simple. Originally, the franchisor charged a sales royalty which was a fixed percentage of turnover. However, private detection is a delicate business where the client often does not wish anyone to know that he has hired a detective. There are considerable possibilities for unreceipted work. Mr. Withers found it too costly to create the sales monitoring system required by a levy on sales turnover. He believes that he has a good idea of how a business should perform and simply levies his royalty as a flat-rate weekly charge. The new system has virtually zero monitoring costs, as no real monitoring occurs.

The Nationwide system contains a reasonable incentive for the franchisor to develop his brand name. If he performs well then he may charge a higher weekly figure. However, on a continuing basis, once the figure is fixed it does not reward the franchisor for improving the business beyond his initial projections. If continuing reputation is unimportant, the franchisor may be tempted to just collect his weekly sum and make no efforts beyond those necessary to keep the franchisee in business. The system is not as good as a percentage royalty from the franchisee's viewpoint.

The weekly royalty is also badly adapted to deal with any unexpected change in basic demand conditions. If a franchise performs very well in relation to the franchisor's initial expectation, we would expect to see attempts by him to renegotiate terms in some way. This has happened with Nationwide. Early on, agreement carried a £100 charge which is now £150. It seems that Mr. Withers underestimated the potential of his franchised outlets.

From the franchisor's perspective, the fixed weekly charge provides a cost-minimising incentive for the franchisee similar to the one from a percentage charge. The franchisee becomes a residual claimant; any cost savings are additions to his profits.

Nationwide is a franchise system which is distinguished by the franchisor's decision to avoid monitoring costs by moving over to the weekly fixed-sum system of royalty. It gives clear evidence that monitoring costs are relevant to the decision over design of the franchise fee schedule. These costs are transaction costs of measurement in Williamson's (1985) terms. In product franchises a reason for using the price of tie-ins to reward the franchisor is that transactions in these must occur anyway. As long as input distortions cannot arise, the franchisor may as well save additional measurement costs. Percentage royalties are economical where monitoring costs are not severe but where tie-ins must not become artificially costly.

Some common aspects of franchises

Generally, the franchise systems in this study do not charge renewal fees to franchisees at the end of the term of their agreements. The exceptions are the Garage Door Company and Nationwide Investigations. These companies

require exactly the same lump sum for renewal as the one required at the start of the original term. Apollo and Pizza Express specify an administration charge of 10% of the current lump sum for renewal in their agreements; the other franchises do not rule out processing charges of this modest kind. We should note that it was not possible to check on these conditions for Computerland Yves Rocher, or Cure 30 as copies of agreements were not supplied by franchisors. Barstock has not yet written its agreement. Renewal fees are not regarded as a good practice by the British Franchise Association.

A second common feature of the 18 franchises is that franchisees make investments which include sunk costs. These are expenditures which would not be recoverable to them if their business failed for one reason or another. Some estimates of these costs are presented in a later chapter where we analyse issues of asset specificity as these affect franchise contracts. At this stage it is important to note that in all cases except Cure 30, the entire lump sum may be regarded as sunk, if one is charged. Also, except in that case, sunk costs always exceed the lump sum. These comparisons and the estimates given later are based on franchisors' estimates of start up costs for franchisees and franchisees' comments on the recoverability of their costs.

Finally, franchisees do not typically obtain financial support from franchisors. This issue is discussed further in a later chapter.

Summary

Franchises may be distinguished by the extent to which the franchise relationship revolves around a product, a brand name or a separately identifiable specialised input into the franchisee's business. Fee

schedules vary for these types with product franchisors using transfer pricing to take their income. Brand franchises and specialised-input franchises rely on fixed percentage sales royalties to reward the franchisor. Given monitoring costs, these schedules minimise the measurement and other costs of linking the franchisor's rewards to his continuing efforts, whilst giving the franchisee an incentive to be cost conscious and to seek revenue. The franchisee is protected against being abandoned by the franchisor. Also, the schedules allow performance to be rewarded as the value of a business unfolds.

It is broadly true that the fee schedules observed 'Economise on bounded rationality while safeguarding transactions against opportunism' (Williamson, 1981, 1546).

The systems analysed in this chapter show regularities. One of these is that sales royalties are generally of the fixed percentage kind. Another is that lump sums charged by franchisors at the start of a franchise contract are closely linked to the set-up costs incurred by the franchisor in training the franchisee and launching the new business. These regularities are further analysed in the following chapter.

CHAPTER 5

Fee Schedules: Some Further Analysis

This is a convenient point at which to examine the fee structures identified in Chapter 4 in relation to existing theories of the franchise contract. In this chapter, it is quickly apparent that the observed fee schedules are not consistent with the predictions of existing theory. Dealing with this matter here, helps to direct the enquiries of subsequent chapters. In particular, it is argued that a broad view of the franchise contract must be taken. It is misleading to take a narrow view based on the fee schedule alone.

Fee schedule analysis

Mathewson and Winter (1985) have given a detailed analysis of franchise fee schedules as part of their theory of the franchise contract. Their basic model is considered in Chapter 2. In this section, a development of their model is assessed in detail over its predictions concerning royalty and lump-sum payments. There is a sense in which the economic theory of franchising has concentrated on these two aspects of the franchise contract. Klein (1980) represents a different direction.

The model to be considered allows the franchisor to realise that the franchisee may attempt to chisel by free riding on his brand name. The franchisee does this by attempting to pass off good demand states as poor ones to the franchisor and adjusting his own effort level downwards to ensure that a consistent output level is observed. The franchisee may save costs which exceed revenue losses from this under some fee schedules. The franchisor cannot observe the state of demand but can measure output; he can detect chiselling on effort levels by incurring monitoring costs.

As we have already discussed in Chapter 2, the franchisor could require the franchisee to purchase the franchise outright for a lump sum. Competition among franchisees then ensures that this sum leaves them with only a normal return. Strategic behaviour would then cause them to make losses. Mathewson and Winter (1985) rule out this possibility in their model, most generally, on the grounds that a final period for transactions renders reputation valueless to the franchisor; he will then be tempted to abscond with a lump sum and not deliver his support services. Similarly, lump sums placed with the franchisor as hostages against the franchisee cheating are vulnerable to such opportunism. If such payments were possible they could support joint profit maximisation for the franchisor and franchisee, which is the first-best contract (the franchisor extracts all pure profit, or economic rent, arising at the franchisee's end of the business).

Given that this contract based only on a lump sum is ruled out, the franchisor maximises his profits subject to the franchisee making normal returns and having no incentive to chisel. The contract which is required sets the franchise royalty in relation to specified effort levels for the franchisee. This removes the incentive to chisel and recognises the franchisor's costs of monitoring. The franchisor's investment in brand name is also set. The franchisor must have no incentive to abscond. This contract is optimal, given monitoring costs, but is not first best. It is second best as it contemplates the presence of irremovable constraints like the failure of reputation as a disciplining device. If means can be found to circumvent such problems without increasing monitoring or other transaction costs then the

franchise contract could be moved nearer to the first best solution.¹

In order to analyse the fee schedule generated by Mathewson and Winter (1985), we must derive expressions for the royalty and lump sum in a version of their model. Initially we assume that the franchise is of the brand or specialised-input kind, although subsequently we include product franchises in the analysis. The franchisee is required to pay his lump sum at the start of his franchise term (he pre-commits it). This is a model containing no wealth constraint for the franchisee (Mathewson and Winter, 1985, 518). It implies that the franchisee is able to borrow money to finance his business, which in general appears to be the case. The version of this model used here allows complete contracting over the advertising and monitoring activities of the franchisor. These assumptions are just conveniences here as the analysis concentrates on the royalty and lump sum payments.

Using the notation of Chapter 2, the franchisor wishes to maximise his profits (R_F):

$$(1) R_F = F + (1-f) \sum_{i=1}^2 T_i X_i (p_i, Q) - C(p) - Q - G$$

where

F = the lump sum fee

f = the percentage of sales revenue kept by the franchisee

T_i = the probability of state of nature i , $T_1 + T_2 = 1$

X_i = demand in state i , 2 is the better state

¹ The constraint may be of the type II kind (removable and non-technological) defined by McKee and West (1981, 439).

q_i = franchisee effort as a cost to him in state i

Q = the franchisor's investment in the brand

$C(p)$ = the franchisor's monitoring costs

p = the probability of detecting chiselling

G = the franchisor's sunk costs

as before. Thus the franchisor maximises his lump sum and royalty receipts minus his monitoring and other costs. His royalty receipts depend upon his brand development, the franchisee's effort level, and the state of nature. The model refers to a franchise that lasts for one time period and has one franchisee, for simplicity. The price of output is set equal to one in the model, which is consistent with the franchisee being a price taker in his local market.

The problem is constrained by the basic requirement that the franchisee should make at least normal profits (and at least cover his costs):

$$(2) \sum_{i=1}^2 T_i \{ f X_i (q_i, Q) - q_i \} - F \geq 0$$

Solving at this point would give the first-best (joint-profit (maximising) levels for F , f , q_1 , q_2 , Q and p . However, we add a constraint which ensures that the franchisee does not chisel on the brand name by making it not worth while to do so:

$$(3) f X_2 (q_2, Q) - q_2 \geq (1-p) \{ f X_1 (q_1, Q) - q_1 (1-y) \}$$

where

y = the proportion by which the franchisee reduces his effort level when misdeclaring the good demand state (X_2) as the poor one (X_1).

Inequality 3 in this chapter differs from inequality 2 in Chapter 2 by the exclusion of the lump sum (F) from the comparison; this is because F is paid at the start and cannot influence the franchisee's incentive to chisel once the contract is running.

A final constraint is given by the requirement that it must be in the franchisor's interests not to abscond with any lump sum that he receives at the start of his franchise. The form of constraint used here for this purpose differs from the one used by Mathewson and Winter (1985, 518). We have already pointed out (Chapter 2, footnote 3)' that their formulation appears to wrongly exclude the lump sum from a comparison of the costs and benefits of absconding. The constraint used here may be derived as follows:

$$\begin{array}{ll} \text{Expected value of} & \text{Expected value of} \\ \text{performance by franchisor} & \geq \text{absconding} \end{array}$$

or

$$(4) \quad F + (1-f) \sum_{i=1}^2 T_i X_i (q_i, Q) - C(P) - Q - G \geq rF - G$$

Where r is the probability that absconding will be successful and where G may be excluded from both sides as sunk costs are faced either way. Sunk costs are attributable to the particular franchised outlet to which (4) refers.

Substitution of (4) into (1) gives

$$(5) \quad rF - G \leq R_F$$

Which may be rearranged to show the third constraint as

$$(6) \quad R_F + G - rF \geq 0$$

Inequality (6) tells us that $rF \leq R_F + G$

This states that where a franchisor is certain to succeed in absconding ($r = 1$), he may be given a lump sum which does not exceed his sunk costs plus his expected profits. As long as he covers his avoidable costs and makes some profit from performing his franchise services, it is of benefit to him to perform properly. In the case where $R_F = 0$, no pure profits are made by the franchisor and the lump sum may only reflect the franchisor's sunk costs. This case corresponds to the situation where the franchisor makes a normal (opportunity cost) return on his assets after extracting any economic rent from the retail end of the business (i.e. these amounts would be rent if it were not for the opportunity cost of the franchisor's services). Where $R_F = 0$ and $r = 1$, the lump sum (F) cannot exceed sunk costs (G).

The constraint (6) ensures that the franchisor only receives the part of his revenues, early on, which may be safely passed over without inducing opportunistic behaviour on his part. The constraint reflects the logic that there should always be a greater return to the franchisor from staying with his contract, either in the form of a contribution to sunk costs or in this form plus pure profits.

Letting U stand for (2), V stand for (3) and W stand for (6), we may form the Lagrangean expression:

$$(7) \quad L = R_f + mU + kV + hW$$

Where m , k , and h are undefined multipliers. The lump sum and royalty may be assumed to be non zero in finding solutions to the maximisation of (4). These solutions differ from the first-best levels for variables (where revenue partial derivatives equal cost partial derivatives) to the extent that $m \neq 0$ and $h \neq 0$. Assuming these values do not arise (i.e. that the additional constraints bind) the condition $\partial L / \partial k = 0$ shows:

$$(8) \quad f = \frac{q_2 - (1 - p)(1 - y)q_1}{X_2(q_2, Q) - (1 - p)X_1(q_1, Q)}$$

This tells us that the royalty $(1 - f)$ varies with the demand and cost conditions facing a franchise. A basic prediction is that different franchises and different franchisees within the same franchise will face different royalty percentages. The prediction does not accord with the findings of Chapter 4 for franchisees within a given franchise, where a uniform royalty is observed. We return to this matter shortly.

From $\partial L / \partial h = 0$, we find that:

$$(9) \quad F = \frac{G + R_f}{r}$$

This is essentially a rearrangement of constraint (6) given that the lump sum (F) cannot equal zero. This result accords with the findings of Chapter 4, where lump sums are found to be closely linked to sunk costs for the franchisor. In particular if $R_f = 0$ and $r = 1$, we can predict that lump sums exactly equal the sunk costs faced by the franchisor. It should be noted that this correspondence does not arise using the version of (6) proposed by Mathewson and Winter (1985, 518) where

$$(10) \quad F = \frac{G + R_F}{(1 + r)}$$

is used². Equation (10) states that $F + rF = G + R_F$, which is not open to the interpretation in terms of a safe early contribution to the franchisor's profits which is used above.

It should be noted that the model used here, along with Mathewson and Winter's original version, does not depend on the risk preferences of the franchisor and franchisee, nor upon horizontal externality between franchisees, to generate its results. Thus, it differs from much agency theory (Shavell, 1979) and from Rubin's (1978) analysis of the franchise contract. It should also be noted that effort levels (q_1 , and q_2) once set may be perfectly enforced by the franchisor, who has taken the costs $[c(p)]$ of this enforcement into account. The issue of forcing, discussed in Chapter 2, does not arise as franchisees are led to self select the optimal effort level given the state of demand and the franchisor's monitoring.

An explanation of observed initial fees

In Chapter 4 we observed that lump-sum franchise fees are typically very closely related to the costs incurred by a franchisor in launching a franchisee. Franchisees usually report that this is the case. Their reports have surprise value as we might expect them to think otherwise. In all cases, the franchisees responded promptly to questions about this and seem to have thought about the matter. Some could cite accounting exercises which had been undertaken to check whether the fee met costs.

² This is discussed in Chapter 2 (equation 9).

In a few cases (such as Pizza Express and Computerland) the franchisor admits to making a small profit over his set-up costs. It is significant that these franchisors refer to a modest profit on the launch operation: they view these profits in relation to those costs and not in relation to the long-run profitability of the franchise. In some cases (such as Wimpy) the franchisor's launch costs can exceed the fee that is charged.

We must be careful to avoid too simplistic a view of the connection between lump sums and the franchisor's launch costs for a franchisee. We cannot simply explain the fees in cost terms. Even though these fees typically match the costs, they still represent an early contribution to the franchisor's income. The regularity to be explained is that launch costs appear to limit the size of that early contribution.

Let us suppose that a franchisor expects his franchise to make pure profits and that the probability of successfully absconding (r) is less than one. According to equation 9 in the last section, the franchisor should be able to charge a lump sum $1/r$ times larger than the sum of these profits and his sunk costs. His launch costs for a franchisee are the sunk costs in each franchise contract he issues. A problem with this is that no franchisee will accept such a charge unless convinced of the accuracy of the franchisor's forecasts. Another difficulty is that data must exist on which forecasts can be based.

The franchisee has no more reason to trust the franchisor's forecasts than to trust in a demand for a lump-sum sale of the franchise. In both cases, the reputation of the franchisor would underpin such trust. It is reasonable to believe that franchisees make their own forecasts of how their businesses will perform.

It is only necessary that franchisees believe that new franchises will make normal business returns for them to be willing to offer no more than $1/r$ times the franchisor's sunk costs as an initial fee. The franchisor can then devote resources to persuading them of the likelihood of better returns or he may choose to accept their view as the relevant constraint on the contract. The observations on the lump sum in Chapter 4 are consistent with a value for r which is approximately equal to one.

A little reflection on the franchise contract from the franchisee's point of view supports a value for r of one. Given that we are not allowing the franchisee to rely on the franchisor's business reputation in constraining behaviour, he would only have recourse to the common law to stop a franchisor from absconding with a lump sum. The nature of the absconding would be that the franchisor would cease to make consummate efforts to develop the brand or offer support services. These types of service are not usually exactly defined in franchise agreements, as we shall see in the next chapter. It is a reasonable expectation for the franchisee to believe that legal action would be expensive in pursuit of undefined failures on the part of the franchisor. Thus, it is rational for him to assume that absconding would be successful. Therefore, the franchisee would only pay a lump sum equal to the franchisor's sunk costs.

The experiences of Young's and Quikframe franchisees as aspects of the franchisors' services broke down is discussed in the next chapter and in cases E and M in Volume 2. These support the view that legal action is avoided in such circumstances, as does the work of Macauley (1963).

However, in reality, reputation is important as there is no last period of contracting for franchisees. Even where one franchise venture ends another may begin; this is the current situation with Wimpy as it moves from table-service to counter-service restaurants (Case F). The question which arises is how important reputation is in qualifying the conclusions drawn above. The answer is that allowing a value to the franchisor's business reputation leaves the conclusion intact: the lump sum still equals the franchisor's sunk costs for an established franchise with a reputation.

The easiest way for the franchisor to maintain his reputation for not absconding from his duties may be to ensure that he is seen to be in a position where it could not possibly be in his interests to abscond. If he detects any inability on the part of the franchisee to assess either the future benefits to the franchisor from the contract or the extent to which reputation can be relied upon to safeguard a lump sum, the franchisor can create certainty by charging an amount equal to his sunk costs. The franchisee knows that the franchisor must expect at least a normal return on his investment: therefore, if the franchisor only charges his sunk costs at the start for his services, there must be some benefit remaining from proper performance. The basis of this explanation is bounded rationality (see Chapter 1) on the part of the franchisee, who cannot calculate an upper limit for the lump sum given uncertainty over either the franchisor's estimates or over the value of his reputation.

The franchisor always has the option of devoting resources to persuading franchisees that his forecasts and behaviour are trustworthy. As examples, he could invest in smart offices or in expensive advertising campaigns. Following Klein and Leffler (1981), the purpose of this

would be to convince observers that he could not afford these expenditures unless his projections were reliable. However, the franchisor has much better information than the franchisee over his incentive to default. It may well be less costly to him to tolerate small adjustments in the franchise contract rather than devote resources to conveying this information. At any rate, the lump sums in the sample of franchises in this study are consistent with this view.

In allowing reputation into the analysis we do not let it support the full lump-sum sale of the franchise. This is again due to uncertainty over the franchisor's forecasts and behaviour. If the franchisee's rationality is bounded so that he is disadvantaged over information flows then it may prove best for the franchisor to remove uncertainty over his behaviour by adopting a royalty or similar payment method. Doing this simplifies the franchisee's decision-making and enables him to participate in the transactions.

The above analysis predicts similar behaviour from established and new franchisors over lump sums, providing each faces similar costs of establishing a reputation as a franchisor. The new franchisor has no reputation and so franchisees expect him to make normal returns. Reputation may be built by an established franchisor by ensuring that it is clear that it is certain that he loses by absconding; it may be that reputation has benefits in reducing recruitment costs for him. In both cases lump sums are linked to the franchisor's sunk costs. This is consistent with the case-study material. Established franchisors like Wimpy and Budget continue to charge lump sums which are approximately equal to these costs³.

³ Although it may be argued that all franchise ventures are relatively new in the UK.

It does not seem useful to describe the franchisee's behaviour as embodying risk aversion in the above. Risk aversion refers to a situation where the certain expected value of a gamble is preferred to actually taking the gamble⁴. In such a case, people systematically avoid risk by comparing outcomes which they can measure. Such behaviour is completely rational if diminishing marginal utility of income implies that the utility of expected value for a gamble exceeds the expected utility from the gamble. However, the franchisee's behaviour is not like this. He lacks information or information-handling capabilities. The franchisor, with superior information, simplifies the decision problem to allow the franchisee to assess the value of participating in transactions. We have an example of Simon's (1957) bounded rationality, where this leads to problem simplification rather than satisficing⁵. The simplified problem still fits the maximising framework of this chapter.

The franchisor does not reduce his lump-sum charge to zero which would reassure them of his behaviour and forecasts if there were doubt over his sunk-cost estimates. This is because doing so would leave him open to hold up by the franchisee. Specifically, the franchisee would realise that the franchisor faces costs of enforcing an existing contract or of finding replacement franchisees, should he threaten to leave the system. As the franchisor is unlikely to wish to run a network that is based on the legal coercion of franchisees, it is the replacement cost which is relevant. This cost is likely to be similar for all franchisees and

⁴ See Gravelle and Rees (1981,554) for more discussion on this.

⁵ Satisficing arises when decision makers seek aspirational rather than maximal values for objectives.

comprises the franchisor's sunk costs for each outlet. Therefore the sunk costs give the lowest value that the franchisor will accept for the lump sum. If he were not to require this amount, there would be an appropriable quasi rent created for the franchisee on the franchisor's setting up services in the sense of Klein, Crawford and Alchian (1978) whose work is discussed in Chapter 1. The franchisee would attempt to extract this by renegotiating his contract once it began. Anyway, there is not the uncertainty over sunk costs that surrounds projections of returns: the costs are incurred in the present and may be assessed using auditing methods.

The results derived above and supported by the data of Chapter 4 are different from the predictions of Mathewson and Winter. Also they predict that the lump sum will extract more economic rent in the case of established franchisees on the basis of equation 10 in this chapter and on the assumption that established franchises make economic rent.

If we relax the assumption that all sunk costs are attributable to individual franchisees, no substantial change occurs in the analysis of lump sums. It will then still be true that the sunk costs attached to the individual franchisee govern his maximum willingness to pay a lump sum and govern the minimum protection required by the franchisor. There would be some (central) sunk costs for the franchisor which would play no role in limiting the lump sum; it is difficult to find examples of such costs. The assumption used in this chapter is empirically relevant.

There may be situations, which we discuss in a later chapter, where lump-sum payments act to demonstrate a franchisee's commitment to his business. It may be noted at this stage that there is no reason why such a function

cannot overlap with the ones described above. We may also note that the franchisor's sunk costs give an upper limit to the amount of any commitment-demonstrating payment, or hostage, that can be safely passed over to the franchisor without tempting him to abscond. In short, it does not matter why the sum is paid as the upper limit always remains the same for the franchisee. However, we shall reserve judgement on this possible use of the lump sum.

Uniformity in fee schedules

A further observation in Chapter 4 is that franchisors charge uniform fees to franchisees. Regardless of location and associated differences in demand and cost conditions, the same input prices, sales royalties and lump-sum fees are charged within a given franchise system. Yet, we should reasonably expect these to vary as demand and cost conditions change. It may be helpful to continue to think in terms of brand and specialised-input franchises for the moment. Equations 8 and 9 in this chapter indicate that the royalty and the lump sum should vary between franchised outlets, unless demand and cost conditions are uniform. As an example, we might expect higher royalties in more profitable outlets.

The explanation for the observed practices is again found by analysing an information asymmetry which affects the franchise system. When a franchisor markets his network he must have forecasts of the business potential of different areas. It is likely that he is often better able to judge this than the franchisee, although this may not always be the case. As noted in the previous section, it is reasonable to assume that the franchisee makes his own forecasts of the performance of any outlet in which he may be interested.

A reasonable expectation for the franchisee is that a particular franchised outlet will show average return. This is consistent with the view that new franchisors make normal business returns, which was cited in the previous section. The franchisor is then faced with a choice whether to devote inputs to persuading the franchisee of the merits of a particular location or whether to accept the franchisee's estimates and construct his fee schedule based on data for an average outlet.

The observation of uniform fees is consistent with the franchisor's acceptance of average performance forecasts on the part of franchisees. This implies that franchisees in above-average-performance outlets will retain some economic rent whereas others will make losses. There is some evidence for this in the case studies.

Franchisors and franchisees were asked whether they were 'happy' with their profitability. It is franchisees' responses which are of interest here. The franchisees did not mind answering this question, although many of them pointed out their reluctance to quantify their answers. The question was always phrased in such a way that it was clear that the franchisee should consider whether returns were good relative to the alternative opportunities available to him. Given that a subjectivist view is required here (opportunity costs differ between the subjects) the answer to this question is more useful than quantitative reporting of returns. The vast majority of franchisees reported that their returns were good but not overwhelmingly so. A typical franchisee, like F₁ in the case of Wimpy, just betters his next-best business opportunity. If this were true of all franchisees, there would be support for the view that franchisees' demand and cost conditions are similar, unless franchisors have hidden devices through which to collect economic rent. However there is variation, with some franchisees

reporting losses within the same systems in which others report normal or very good returns (for example: Garage Door, Midas, and Nationwide).

The first interpretation of the above observations is that uniform fees leave economic rent accruing to some franchisees whilst others make losses. It may be too costly for franchisors to operate information systems that would enable differentiated fees to be applied. However, average fees cause too few problems for this to be plausible. If demand and cost conditions do vary between outlets within the same system then the data are consistent with the franchisors using other instruments (such as property charges) as devices through which to extract economic rent. In particular, we might expect franchisors to attempt to introduce devices as a franchised outlet become established and its profitability becomes clear. In the case of loss-making outlets, these devices may need to work in reverse.

The financial implications of contractual devices other than the franchise fees are discussed in the next chapter.

Product franchises and fees

Product franchises may be brought into the analysis of this chapter quite easily. Their distinguishing characteristics are that the franchise relationship revolves around a branded product and that fees are included in transfer prices. We have already explained in Chapter 4 how it can be more economical to add charges for franchise services to product prices when no input distortions arise as the product transactions are occurring anyway.

It is economical to use the prices of tied-in products to transfer any economic rent when inputs, including the tie-ins must be used in fixed proportions. This means that inputs must be used in fixed proportion to the output as well. In turn, this means that a mark up on the cost price of the manufacturer is exactly equivalent to a fixed percentage sales royalty as long as the transfer price is specified as a fixed percentage of the final price. In product franchises this is typically what happens. Manufacturers state their transfer prices as a proportion of their list retail prices. Thus, Austin Rover dealers receive a 17% discount on the recommended price of the Rover Sterling. This indicates that the maker is applying the same kind of charge as the other types of franchisor but finds it most economical to link his royalty to the product transfer.

A distinction with product franchisors like Austin Rover, Ford, Apollo, Bally and Yves Rocher is that there appears to be a pricing issue with franchisees. To avoid forcing, sales bonuses are used to encourage the franchisee to lower his prices and sell more than he otherwise might. These act as a low-cost method of reducing the implied sales royalty to expand sales. This issue could be included in the type of mathematical model used in this chapter by specifying that the franchisee's effort level is a function of the royalty.

Apollo and Yves Rocher both charge franchisees a lump sum which may be analysed in the terms used above. However, all product franchisees pay a less obvious lump sum which is the cost of their initial stocks. The franchisor and franchisee typically agree an annual sales plan and the franchisee then holds and pays for stocks to support this plan. Such payments have classic lump-sum aspects: an estimate is made of the future value of sales and some of this is passed to the franchisor early on.

Changes in stocks then represent the continuing reward to the franchisor. Stocks are highly valued in the case of Bally and the vehicle dealers. They are small in the case of Apollo franchisees, who are close to central stores.

The principle that the franchisee will hand over income only when convinced that the franchisor still has an incentive to deliver his services may be extended to the study of stocks. We are concerned to explain their ownership not their size⁶. The franchisee certainly will be prepared to pay the franchisor's sunk costs in supplying his stocks, because the franchisor is definitely left with an incentive to promote the brand and help in actually selling the product. In addition, the franchisee may pay some or all of the difference between these costs and the wholesale price of the goods if he is confident that this will not impair the franchisor's incentive to provide outstanding services. The least payment acceptable to the franchisor (unless some other device reassures him) is his sunk costs in providing the stocks. This suggests that Bally, Yves Rocher and Apollo franchisees believe that the continuing value of their association with the franchisor exceeds the value to him of chiselling on the services required to sell products which they buy at wholesale prices. However, the car dealers are different.

The car dealers pay stocking fees that are less than the wholesale value of the cars which they may keep in stock. Austin Rover requires an interest payment (less sales bonuses) on the stock cost of one and a half month's sales, based on the previous 12 months. Some dealers have no limits on the size of stock they may keep, subject to model availability. Similarly with Ford, a deposit of 13%

⁶ Optimal stocks minimise the sum of holding and ordering costs whoever owns them. See Baumol (1977,9).

of the previous 12-months sales is placed by the dealer who may then hold stocks up to six times this value. In practice, the dealers keep stocks deemed suitable by the makers.

The car dealers are different because if they were to own the stocks which they keep they would bear the full risk of being unable to sell outdated models at profitable prices. This could alter the maker's behaviour away from maintaining secrecy over model abandonment until the last possible moment. This problem is less severe in the case of Apollo, Bally and Yves Rocher as stocks of any particular line are of relatively low value.

The absence of a separate lump sum in the case of Bally, Austin Rover and Ford may be explained by noting that all sunk costs can be covered by the product sales involved in an initial sale of stock. In the case of Bally, the franchisee buys a large stock at wholesale prices. This creates a substantial part of the franchisor's income, covering sunk costs, avoidable costs and possibly contributing towards profit. There is no need for the franchisor to further protect himself. The car makers charge stocking fees which are not related to the full cost of the stock but to sunk costs like the interest payments which must be incurred on stock finance. Their behaviour in charging no additional lump sum is consistent with adding more general sunk costs to these charges. Apollo and Yves Rocher charge a separate lump sum because their franchisees' stocks are insufficiently valuable to cover all sunk costs, by this reasoning.

Summary and some conclusions

Lump sum franchise fees and sales royalties may be explained by noting that the franchisee will only pass over income to the franchisor if he is sure that this will

not remove the franchisor's incentive to perform his services. The model in this chapter defines this constraint very carefully and explains an observed link between lump sums and the franchisor's sunk costs. In doing this, attention is drawn to the importance of franchisors' and franchisees' expectations over business returns and each other's behaviour. These expectations are worthy of further study.

Product franchises use transfer pricing and stocking charges in their fee schedule. These devices correspond to sales royalties and lump sums.

Events may turn out differently from the predictions at the start of a franchise system, when the fee schedule is set. We should expect other contractual devices to allow the franchisor some scope for extracting economic rent from the retail end of the businesses and for supporting weak outlets. As we shall see, there is not evidence for the queues of franchisees or the bankruptcies which uniform fee schedules imply if demand and cost conditions are variable. This sort of reflection leads us to consider the wider aspects of the franchise contract. The fee schedule is a narrow aspect of the contract and may represent no more than a statement of average returns around which the full contract can operate.

CHAPTER 6

The Wider Franchise Contract

Franchisors and franchisees usually sign written franchise agreements covering some key aspects of their relationship. These areas include the fee schedule, operating standards, sale of the franchise, and termination of the relationship. The written agreement, or 'explicit contract', is much wider than just the fee schedule, which we have analysed in Chapters 4 and 5. At the same time, it tends to be incomplete; not every detail of all transactions can be specified in advance so that much is left to be decided as events unfold. The written agreement is augmented by business understandings which form an 'implicit contract' that also governs the franchise relationship.

This chapter analyses the wider franchise contract in terms of the economic functions of contractual devices which are found within the sample of franchises in this study. Franchise agreements are studied in terms of their common and unusual contractual clauses.¹ The fee schedule analysis is not repeated, however. Implicit contracts are then analysed for the franchises to see how incomplete agreements are augmented in practice. It emerges that some contractual devices interact with specialised assets in franchise systems; this link is fully investigated. Methods of contractual enforcement and performance monitoring are also of interest in what follows as they have implicit contract aspects. Finally, we look for evidence of contractual revision.

¹ A clause is a separable written requirement acting on at least one party.

Revision of the contract implies adaptation to the business environment. Chapters 4 and 5 have indicated a need for this in that observed fee schedules have been shown to fit a second-best model based on average performance. There is scope for individual franchise contracts to cope with non-average performance and to encourage moves towards efficiency, either in the sense of moving towards a first-best (joint profit maximising) situation, or in the sense of ensuring that production methods are least cost and that all profitable revenues are realised. The interest in adaptation leads us to assess whether franchise contracts are 'relational' (Goldberg, 1980, 338; Macneil 1978, 901; and Williamson, 1985, 72).

Macneil (1978) has proposed a three-way classification of contract. 'Classical' contract arises when parties attempt to cover all possible future contingencies by contractual instruments such as written agreements or oral undertakings. In classical contract there is a presumption that written agreements always dominate informal ones and that courts are used to interpret the contract. According to Macneil (1978, 864) in this arrangement the consequences of non-performance are predictable and are not open ended. The economic counterpart of the classical contract is the impersonal market transaction where only the specified details of an exchange are important. In particular, the identities of parties are not important (Williamson, 1985, 69).

Macneil observes that litigation through the courts can be an inflexible means of governing long-term contracts. In an uncertain world, complete specification of a contract is prohibitively costly. Parties therefore try to create contracts which preserve trading by setting guidelines to precise performance which must be interpreted at a later date. Macneil calls this

'neoclassical' contract and notes that alternative governance structures to the use of courts may have advantages here. In particular, an arbitrator may be chosen at the start of the contract whose job is to settle disputes that may arise over areas which require interpretation. A simple advantage for arbitration is that the arbitrator may be chosen as someone with professional knowledge relevant to the contract: for example, if property values may change in a property lease, the parties may accept the valuation of a Chartered Surveyor. Williamson (1975, 70) points out that recognition of real-world complexity, incomplete contracts and the need for confidence in settlement machinery characterises neoclassical contract.

The third classification is that of 'relational' contract, which refers to the interest in adaptation noted above. A 'relational contract is one where adjustment processes become transaction specific. The duration and complexity of a contract reaches such a height that continual adjustment is necessary. According to Macneil (1978, 901) a 'minisociety' grows up around the contractual relationship which uses principles to guide adaptation which are drawn from the long term value of the economic relationship between parties. Thus, it may not just be the value of a specific area of transaction which governs performance but also the value of other possible connections which may emerge. According to Williamson (1985, 72) in contrast with neoclassical contract, a written agreement may play a small role in defining the reference point for a relational contract. Outside intervention either by courts or arbitrators is not central to this type of contract; instead individuals engage in their own monitoring of each other's performance and possess their own sanctions.

Relational contracting is the essence of the economics of organisation for Goldberg (1980, 339) who uses the franchise contract as an illustration of the relational form. It is clear that the relational contract can contain explicit and implicit aspects, and that it exists in a dynamic context. That context is one where individuals grapple with problems of incentive alignment as circumstances change and is very different from agency theory, where ex-ante contracts are designed with perfect knowledge of the probability distributions affecting relevant behaviour. In a relational setting, the ex-ante contract may bear little resemblance to what unfolds.

The franchise agreement

All 19 franchise systems in this study use (or intend to use) a written franchise agreement. In the case of Bally, the franchise system initially ran for one year without an agreement but this is very unusual. In the cases of the vehicle franchises, makers use a number of agreements covering dealership standards, sale or return of stock and stock-financing arrangements; these are identified in the case studies. Some franchisors use property-lease agreements in addition to basic franchise agreements. For simplicity these sets of agreements are treated as the franchise agreement in the economic analysis which follows.

In studying the written agreement, the sample of 19 franchises produces only 15 useful cases as four franchisors did not provide copies of their franchise agreements. This figure represents three refusals as Barstock has not yet (1987) designed an agreement.

It should also be noted that all except one of these agreements are registered with the Office of Fair Trading (OFT) in accordance with restrictive trade practices

legislation. Generally, the OFT takes a benign view of franchise contracts. There are no cases where franchisors and franchisees have been prosecuted for restrictive practices. The franchisor's incentive to register is that his agreement is unenforceable if it is not registered under the Restrictive Trade Practices Act, as one franchisor in the sample found to his cost. The OFT has a longer-term concern with clauses in agreements that limit the franchisee's competition with his old network if he leaves.² Another area of concern is the granting of sales territories to franchisees. However, exclusive sales territories are usually avoided by legal draughtsmen, who instead allocate the use of a trade-mark or the right to advertise in an area.

Common clauses

It is sometimes suggested that franchise agreements are all very similar as they tend to originate from the same law firms. In fact there is considerable variety between different agreements over the manner in which particular clauses are applied and over the clauses themselves. The variety reflects the product, brand, or specialised-input nature of the franchises. Nevertheless, some clauses are broadly common to the reduced sample of 15 agreements. These are now briefly analysed.

In all cases except Mobiletuning (Case D) the franchisor sets maximum prices for the product or service. Where a franchisor has tied-in sales, price reductions by the franchisee do not affect the transfer price of the products although they may induce payments under a bonus scheme if they increase sales. Where a royalty is paid, this is normally calculated as a percentage of actual sales turnover and is not based on list prices for the

² Stated in a letter to the author.

final service or product. In the royalty case, the franchisor must believe that price reductions will only be undertaken if they increase total revenue. Mr. Rowntree states in the case of Mobiletuning that he worries that franchisees might compete on price alone when left to their own devices; he therefore has limited their scope for this by setting minimum labour charges in his agreement.

A term for the franchise agreement is usually set. This ranges from the modal value of five years up to 26 years. Two important exceptions are Austin Rover and Ford: the car makers confer open-ended franchises on dealers. Renewal clauses almost always equal the initial term. The exceptions to this are the vehicle cases, where renewal is not relevant, Bally, which confers open-ended renewal, and Avis and Budget, who automatically extend their five-year terms by continuous annual extensions. Renewal of the term is always conditional on the franchisee's (or franchisor's) satisfactory performance. Renewal dates give a definite point at which an unsuccessful franchisee could be required to abandon his specific investment in a franchised outlet. The franchisee may also freely leave at these points without fear of any court action.

It is most unusual to require franchisees to pay a further lump-sum franchise fee at renewal dates. Only Garage Door (Case B) and Nationwide (Case I) do this. The practice can amount to a crude attempt at the extraction of economic rent given that franchisees have made a specific investment in building up their local business goodwill. This is most likely to be the position with Nationwide, where franchisees describe their negative incentive for paying the fee. They feel that they are buying continued access to their client lists. If they leave the network then they cannot trade in their old

areas for one year. The Nationwide franchisor has experienced difficulties in retaining franchisees and these may well reflect a failure to convince them that the second lump-sum still leaves an incentive for franchisor performance. In the case of Garage Door, the low fee of £2,500 may only cover administrative costs after allowance has been made for inflation; this depends on the costs of training franchisees. All franchises charge modest fees of 5% to 10% of the current initial fee for renewal and these cover administration costs.

Renewal fees provide support for the conclusion of Chapter 5 that lump sums do not extract economic rent in any significant fashion. Unless the franchisor were able to take a very long term view and were certain of renewal, we should expect rent extraction to be repeated. It seems rather that sunk costs limit early rent transfer once and that much lower additional sunk costs limit it for any continuation of the contract. Nationwide is a definite exception to this (an outlier) and experiences difficulties in retaining franchisees.

All 15 agreements contain clauses which exhort franchisees and franchisors to perform their services to their best endeavours. For the franchisee this means that he is required to make his best efforts over sales and over the quality of his services. For the franchisor, the requirement is for adequate levels of training, support and consultation. These areas refer to the quality of effort and are left incomplete in the explicit contract. An example may be given from the Midas agreement:

'The franchisee shall diligently carry on the said business at the said premises and shall use its best endeavours to promote and extend the said business...'

(Midas Franchise Agreement, 1986, 12)

This is a typical incomplete clause which exhorts the franchisee to consummate performance. It is left for subsequent events to show whether performance is judged adequate with no guidelines being set in the agreement. In particular, we do not observe the minimum standards predicted by agency theory (Mathewson and Winter, 1985).

All franchisors in the sample providing agreements require the franchisee to have no similar business interests. As we shall see in a later chapter, franchisors seek commitment to the business on the part of their franchisees. The clause can be simply interpreted as a means of ensuring this.

The 15 franchise agreements all allow franchisees to sell their businesses subject to the franchisor's approval of any intending purchaser. The vehicle franchises are just about included in this as they do not stop a garage from selling out to anyone but reserve the right not to transfer the franchise. The main purpose of this type of clause is that the franchisor is not forced into a long-term relationship with someone not of his choosing. Unlike classical contract, the franchise agreement is not impersonal.

The franchisor's right of approval of any prospective purchaser implicitly defines a right to control the ownership of the franchised outlet. Sometimes this is supported by a clause which specifically requires changes in the ownership of the franchisee's business to be subject to approval. Avis, Young's, Austin Rover and Ford all use such a clause. Garage Door prohibits partnerships because of a poor experience with one (this suggests relational aspects). These clauses imply that franchisors wish to contract selectively with individuals with whom they believe their systems will work.

The franchisee always indemnifies the franchisor against employer-liability and public-liability risks. Sometimes specific insurance policies and values are required; for example, Young's requires £500,000 worth of such insurance to be taken by the franchisees. This type of clause simply locates these risks clearly with the franchisor.

All franchise systems in the sample require franchisees to report sales to the franchisor, usually at weekly or monthly intervals, and to pass over audited accounts at intervals of usually one year. Clauses which support these arrangements simply define one aspect of the franchisor's monitoring system.

Commonly absent clauses

Among the sample of 15 franchise agreements some clauses stand out as being rarely used. The first of these covers the franchisee's right to give notice and quite the franchise contract. Only Austin Rover, Ford, Avis, Budget and Mobiletuning confer this right on franchisees. The car makers allow franchisees to resign with two-years notice. Avis and Budget both accept six-months notice. Mobiletuning only requires three-months notice. In each case, voluntary exit of this kind is subject to the same consequences as any other form of contract termination. The clause reflects the wish of these franchisors to provide an explicit exit route for franchisees who become unhappy with their situation. We should note that there is always an implicit route as franchisees can contrive an exit by performing so badly that the franchisor terminates the agreement.

Only Ford and Nationwide write penalty clauses into their agreements. Nationwide charges franchisees a deposit equal to 13-weeks royalty. This sum is just less

than £2,000 (1987) and may be drawn on if the franchisee falls into arrears with his payments. Ford applies modest penalties of £500 per vehicle and £100 in the case of parts if a dealer exhibits products outside of his area of responsibility. In the case of sales at prices above list prices, Ford fines dealers one-half of a vehicle price and three times a part price. These are the only cases of explicit hostage taking (or bond posting) throughout the entire 15 agreements. The rules governing lump sum payments (Chapter 5) apply here also: the franchisee must be confident that the franchisor will not abscond with the hostage. At the same time the franchisor needs to know that the hostage is valued by the franchisee. Although these hostages are valued by the franchisor - they are not ugly princesses (Williamson, 1985, 177) - the sums are small in relation to the continuing benefits from the contract. Nevertheless, the general drift of franchising theory is correct in assuming that explicit hostages are not usual (Rubin, 1978; Klein, 1980; and Mathewson and Winter, 1985).

Wimpy, Bally, Midas and Ford stand apart in specifying minimum opening hours for franchisees. Even then it is only Wimpy that sets core hours (between 10 am and 11 pm) in the agreement. Midas reserves the right to set hours as deemed fit. Bally and Ford franchisees are required to keep the shop or showroom hours which are normal for a locality. Except in Wimpy's case, the clause is incomplete. It may be that setting hours is a part of Wimpy's standardisation of service. However, most franchisors do not find such limits necessary.

In only one case are specialised assets provided by the franchisor for franchisees. Bally provides the shop sign, interior racks and display items free of charge. In this case, the items remain the property of Bally. We should note that franchisees still make considerable

specialised investments. The general lack of support over specialised assets indicates that quasi vertical integration is not common in franchise industries (Klein, Crawford and Alchion, 1978; and Monteverde and Teece, 1982). The franchise contracts are in themselves capable of supporting specialised investments. The devices which constitute this support are discussed further below.

Clauses which differ

Franchise agreements often impose controls on the franchisee's input, on the tenure of his business property, and on the territory within which the franchised outlet operates. In addition, it is common to control the conditions governing sale of the franchise or termination of the agreement. Advertising requirements are also often set for either the franchisor or franchisee. These are all examples of areas of agreements which differ in terms of the regularity with which they are controlled and in terms of the clauses which are used. Finally, some agreements provide for compulsory arbitration over the interpretation of at least some clauses.

We begin in this section by considering the controls which are sometimes imposed on inputs into the franchised business. First, tied-in sales have been considered earlier and it has been found that these are used to maintain standards for brand franchisors but may be used to transfer economic rent by product franchisors. The other inputs which may be controlled are the displays and fittings used by a franchisee, his personnel, and equity participation in the franchised business. The differences between franchises over these inputs is shown in Table 6.1.

Table 6.1 Input control

	<u>Displays</u>	<u>Design</u>	<u>Personnel</u>
Apollo	Charged	Strong	All
Garage Door	Charged	Yes	None
Avis	Charged	Yes	Management
Mobiletuning	Insignificant	Yes	All
Young's	Charged	Strong	None
Wimpy	Charged	Strong	None
Pizza Express	Charged	Strong	None
Bally	Free	Strong	Management
Nationwide	None	None	All
Budget	Charged	Yes	Management
Midas	Charged	Yes	None
Olivers	Charged	Strong	Management
Quikframe	Charged	Strong	None
Austin Rover	Charged	Yes	None
Ford	Charged	Yes	None

Let us first consider display materials. We have already considered the provision of some of these items by Bally. With the exceptions of Nationwide and Mobiletuning, the remaining franchisors provide display items at cost for their franchisees. In all cases except Nationwide, display items are subject to control by the franchisor in the agreement. This is to preserve a corporate appearance for all outlets. Having set controls, it is most economical for the franchisor to establish a centralised supply line for display items in order to realise buying economies.

We may treat these inputs analogously to tied-in sales. They are not important in the case of Mobiletuning, where the business is van based, nor in the case of Nationwide where no high-street premises are required. In all other cases, displays are important and require control to maintain uniform standards. The question then is how to charge for the items. Generally,

it is best to charge directly for the displays unless there is some other transaction already occurring which could carry the charge without distorting inputs. In the case of some of the product franchises, where the business is based virtually entirely on an irreplaceable product, the cost of displays may be absorbed by product transfer prices without causing problems, providing the franchisor imposes physical controls on the amount of display material that an outlet may receive. We observe Bally using this option. The preference of other product franchisors for direct charging probably reflects a reluctance to set restrictions on display-material consumption; it may be thought better to face franchisees with true costs and then let them choose a locally optimal amount of display.

Further input control occurs over the design of premises. This goes further than the control of displays. Most high-street franchises in Table 6.1 have a strong form of this control. Apollo, Young's, Wimpy, Pizza Express, Bally, Olivers and Quikframe take the right to direct the design of the franchisee's business premises, and often have their own design departments. The other franchises, except for Nationwide, subject the franchisee's designs to the franchisor's approval before they may be used. Office design is not considered relevant to the Nationwide brand image. In general, it appears that design control becomes stronger the closer that a business is located to the high street. Such an increase in control involves more administrative dealings with architects, local authorities and the franchisee. The franchisor is more likely to accept this increase in transaction costs if detailed standardisation of premises is more highly valued in a business. Since many brand franchises are high-street based, there is a slight tendency in Table 6.1 for stronger design control in these.

A final area of input control concerns the franchisee's personnel. Avis, Bally, Budget and Olivers take the right to control the franchisee's appointment of any senior manager. Apollo, Young's and Nationwide have the right to veto the appointment of any of the franchisee's employees. This control of personnel is difficult to explain. It may be that franchisors fear that inexperience would lead new franchisees to make poor appointments. Equally, franchisors may fear that franchisees could chisel on service quality and free ride on the brand name by hiring cheap low-quality employees. Finally, the franchise relationship may adjust over time and the franchisor needs to be confident that personnel will be adaptable under changed circumstances. The difficulty with these explanations is that we find that over half of the franchisors in Table 6.1 do not use personnel control. The control may well be superfluous; franchisees are probably led by their own business interests to make similar appointments to the ones that would be made by the franchisor. Support for this view is found in the case studies: only one franchisee (S₁) reports that his franchisor occasionally expresses dissatisfaction with personnel. Generally, personnel issues seem not to arise.

All of the above areas of input control, just like tied-in sales, operate to keep the franchisee's standards of operation within limits set for the franchise system. Franchisors necessarily incur monitoring costs in policing the standards which are set.

A further area in which agreements show variety in their clauses concerns advertising, either by the franchisor or by the franchisee. This area also refers to input control but it is of great importance in most franchise relationships and is worthy of separate

treatment. Table 6.2 shows the range of advertising requirements in the sample of 15 agreements provided for this study.

Table 6.2 Advertising requirements

	<u>Franchisor</u>	<u>Franchisee</u>
Apollo	Exhortation	Discretionary minimum
Garage Door	Exhortation	Exhortation
Avis	Exhortation	2.5% of turnover
Mobiletuning	Yellow Pages	Exhortation
Young's	30% of royalty	3% of turnover
Wimpy	Exhortation	Exhortation
Pizza Express	Exhortation	Exhortation
Bally	Exhortation	Exhortation
Nationwide	Exhortation	Yellow Pages
Budget	25% or royalty	2% of turnover
Midas	Exhortation	Exhortation
Olivers	Exhortation	Exhortation
Quikframe	Exhortation	Yellow Pages
Austin Rover	Exhortation	Exhortation
Ford	Exhortation	Exhortation

From Table 6.2 it can be seen that only six franchisors set minimum standards of local advertising for franchisees. In all other cases, the agreement is incomplete and merely exhorts the franchisee to undertake 'appropriate' amounts of local advertising. In two of the six cases (Quikframe and Nationwide) the minimum requirement is minor and consists only of a duty to insert a Yellow Pages advertisement.

The picture is similar when we look at franchisors. Except in three cases, the agreement is again incomplete and merely exhorts. In one of the cases (Mobiletuning) the requirement on the franchisor just consists of a duty to pay for franchisees' Yellow Pages inserts. Young's uses a 10% sales royalty so that the requirement to spend

30% of the franchise royalty on central advertising matches the franchisee's duty to spend 3% of his turnover on local promotion. Budget also uses a 10% royalty so that 25% of royalty income, which is the minimum to be spent centrally on advertising, just outweighs the franchisee's duty to spend at least 2% of turnover on local advertising. It is unusual to constrain the parties in this way. Only Avis constrains the franchisee without doing the same to the franchisor.

The most likely explanation of incompleteness over advertising requirements is that the value of such promotion is dependent upon demand and cost conditions as these unfold. This is not a problem if we contemplate the possibility that advertising in excess of some minimum level may be required on the part of the franchisor or franchisee. In that case, if advertising looks worth while we simply undertake it. A problem arises if minimum levels are set too high, in which case profitability will be eroded. Rather than risk this, both franchisors and franchisees may prefer to adjust advertising over time. Uncertainty together with bounded rationality indicates that this part of the contract be left implicit.

There is evidence in the case studies that the contract over advertising tends to evolve over time and that it has relational aspects. First, three of the franchises set an advertising levy in their agreements. Garage Door apparently requires a contribution from the franchisee of 4% of his turnover, whereas Pizza Express requires 0.5%. These levies are not taken in practice. Olivers only takes 1% of a 2% levy. We may also note that E₂ tells us that Young's does not enforce the requirement that franchisees spend 3% of their turnover on local advertising. This all suggests that flexibility in determining the advertising budget is found to be

desirable to avoid depressing franchisees' profits. We return to implicit contractual matters shortly.

Table 6.3 analyses details of property-lease arrangements which either result from the agreement or from an attached property agreement that the franchisee is required to use as part of his implicit contract. As the terms of lease arrangements are in an agreement of some sort it is convenient to deal with them in this section.

Table 6.3 Lease control

	<u>Must take</u> <u>from</u> <u>franchisor</u>	<u>May take</u> <u>through</u> <u>franchisor</u>	<u>Property</u> <u>rent</u>
Apollo	Yes		Variation
Garage Door			
Avis			
Mobiletuning			
Young's		Yes	Not issue
Wimpy		Yes	Not issue
Pizza Express	Yes		Not issue
Bally	Yes		Fixed
Nationwide		Yes	Not issue
Budget			
Midas	Yes		Variation
Olivers	Yes		Fixed
Quikframe			
Austin Rover			
Ford			

NOTE: blanks indicate that lease control is not used

The table shows that is is only just more usual for a franchisor to be involved in his franchisee's property

leases than for the franchisee to be quite independent in this respect. Five of the franchisors insist that franchisees take leases using the franchisor as head lessee. Under these arrangements a property owner leases property to the franchisor who in turn sublets it to the franchisee. In three more cases (Young's, Wimpy and Nationwide) the franchisee may choose to use the franchisor as head lessor if he so wishes. It is often true that high-street landlords are reluctant to let shops to small businesses but will deal with companies who are the size of some of the franchisors.³ Nationwide does not use high-street property but the franchisor sometimes holds leases dating from the time when an office may have been directly operated.

Franchisors typically argue that head-lease arrangements help the franchisee to secure high-street property when this would otherwise be difficult for him. This point is made by franchisors throughout the cases. However, this cannot be a reason for insisting that franchisees take subleases. The argument that the franchisor is just offering a service over the lease can be accepted in cases like Young's, Wimpy and Nationwide, where the franchisee is free to choose whether or not to use it.

Why then does a franchisor oblige franchisees to take subleases from him? One possibility is that by controlling the franchisee's property, the franchisor obtains an additional instrument through which to extract economic rent from the retail end of his business. The franchisor may add an administration charge to the property rent. Along with the rent, the additional charge may be varied (usually at five-year intervals). Thus a

³ The landlords are worried about business failure giving their site a bad name and depressing its value.

franchisor could alter his extraction of economic rent at these times. This could be a useful facility; as we saw in Chapter 5, uniform royalty charges may be insensitive to variations in local cost and demand conditions. There is case evidence for the occasional use of property charges in this way. A₁ and A₂ state that property charges have varied as Apollo has detected changes in their profitability. In particular, when A₁ showed that profitability was poor on a site the lease charge fell. K₂ relates a similar experience.

However, the experience of these franchisees must be seen in context. Of the 15 franchisors who provided agreements for this part of the study, seven have nothing whatsoever to do with franchisee's tenure of their property, yet they all face similar profit-variability issues. Of the eight with lease connections, the property rent which may be passed on is either fixed in the agreement or never becomes an issue in six cases. Franchisees' reports that property charges are at normal commercial levels have surprise value; we should expect them to have suspicions about this. However, in most cases they experience no anxiety concerning their vulnerability over the costs of their property.

It is safe to conclude that whilst lease control may sometimes lead to attempts at extracting economic rent generally it does not. We must therefore look for some other function that lease control performs.

Franchisors frequently mention that control of the lease enables them to retain the goodwill attached to a site in the event that a franchisee leaves the network. Mr Phillips of Midas, for example, cites this as the reason why Midas introduced lease control: a number of outlets had been lost when franchisees failed to renew

their agreements.⁴ With lease control the franchisor can ensure that the franchisee hands over the business premises upon leaving the network or having his agreement terminated. Even in cases where the lease does not automatically revert to the franchisor upon termination (as with Wimpy) limitations normally imposed on the franchisee's freedom of operation after termination would cause him to prefer to surrender any unexpired lease to the franchisor.

Over the life of a franchised outlet, both the franchisor and the franchisee invest in its development. We have seen in Chapters 4 and 5 that the franchisor may recover his sunk investment in an outlet through the lump sum franchise fee. This sunk investment really refers to expenditures incurred at the start of business. Over time, the franchisor's advertising and development of the brand contributes to building the local reputation of an outlet (the 'goodwill'). The franchisee similarly makes an initial sunk investment and subsequently builds up his local business reputation. The practical effect of goodwill is that customer recall of the proficiency of the business will lead to repeat sales and to recommendation.

There is no simple reason why the franchisor should expect to retain all goodwill if the franchisee leaves the network. Even in the case where the franchisee's and franchisor's contributions to local goodwill are not separate, the franchisor could use penalty clauses in the agreement to obtain compensation for losses that he incurs if he has to start up another franchisee and wait for goodwill to again build up. As long as compensation only covered these costs there would be no incentive for the franchisor to contrive the franchisee's exit by finding

⁴ The copy of the Midas agreement in the author's possession is too old to contain the new arrangement.

spurious grounds for terminating the agreement. In the case where the national and local contributions to local goodwill are separable then the franchisor would only require his search costs in finding another franchisee as compensation; he would then continue to benefit as before from his contribution to local goodwill. In the separable case, it is worth noting that problems of measuring goodwill do not cause difficulties in finding appropriate compensation.

Goodwill retention can be a reason for lease control in the following sense. The control sets the costs of establishing the brand name on a site in the area equal to zero. This may be useful where these costs are normally high and possibly difficult to calculate. The control does not remove the costs of finding a new franchisee. Nor does it retain local goodwill for the franchisor which is properly regarded as belonging to the franchisee.

It is possible to distinguish between the franchisors in Table 6.3 on the basis of their re-establishment costs. With the exception of Nationwide, whose property interest does not appear to have resulted from conscious policy, the franchisors using lease control are those likely to have high costs in searching for new sites. It is this transaction cost which appears to determine the use of lease control. Pizza Express, Apollo, Bally and Olivers all require high-street sites. These are difficult to acquire and take time to adapt. Midas has the kind of fast-fit vehicle maintenance operation for which it is not easy to obtain planning permission. Young's and Wimpy generally find that their franchisees need sublease arrangements so that they may benefit from this type of cost protection without setting a contractual clause. However, Avis, Budget, Garage Door and Quikframe are businesses which need quickly commissioned offices or workshops located away from the high street. Mobiletuning

is a van-based business so that disruption costs consist simply of search costs for a new franchisee, where these are not altered by lease control (which is anyway infeasible). The car makers also just face the search costs of finding another garage wishing to take the franchise. It does look as though lease control is adopted in situations where it will reduce the franchisor's transaction costs of re-establishing the franchised outlet. These are principally the result of high-street location problems. If franchisors did not protect themselves they would create appropriable quasi-rents for franchisees.

This line of reasoning suggests that Klein (1980, 359) may be questioned over his suggestion that lease control is adopted where the franchisor wishes to be able to impose capital losses on the franchisee should the franchisee cheat in the performance of his services. Most shop-based or office-based franchises require leasehold (ie, property) improvements to be made by the franchisee. If he is forced to give up the lease, this part of his wealth is lost by him. We return to Klein's argument in Chapter 7 which covers specialised asset creation in franchises. Williamson (1985, 180) follows Klein. For the moment, we may note that protecting the franchisor from disruption costs is compatible with Klein's hostage view of lease control: ~~both~~ functions could be performed by the device.

Franchise agreements vary in the conditions which are imposed upon the franchisee when he either leaves the network or has his agreement terminated. Table 6.4 shows some of these conditions for the 15 franchise agreements under analysis in this section. We start by considering the less frequently used clauses.

Table 6.4 Termination conditions

	<u>Professional limitation (years)</u>	<u>Telephones</u>	<u>1st refusal</u>	<u>Asset repurchase</u>
Apollo	2	No	No	May
Garage Door	2	Yes	Yes	No
Avis	1	No	No	No
Mobiletuning	2	No	No	No
Young's	1	Yes	Yes	May
Wimpy	No	No	No	May
Pizza Express	1	No	Yes	May
Bally	1	Yes	Yes	Must
Nationwide	1	Yes	No	May
Budget	0.5	Yes	Yes	No
Midas	2	Yes	Yes	Must
Olivers	2	Yes	Yes	Must
Quikframe	1	Yes	Yes	May
Austin Rover	No	No	No	Must
Ford	No	No	No	Must

A number of the franchisors require that franchisees should transfer telephone lines to the franchisor or a replacement franchisee in the event of exit or termination. This represents an attempt at preserving the franchisor's goodwill in the area which may be important in businesses like vehicle rental where there are many telephone enquiries from customers. The restriction probably does very little for franchises like Bally.

It may also be seen from the Table 6.4 that eight of the franchisors take a right of first refusal on any offer to sell the franchised outlet to a third party. This gives the franchisor the certainty of being able to buy control of the franchised outlet without bidding beyond a reservation (market) price that the franchisee may have.

Of more interest are the undertakings by 10 franchisors that they either may or must repurchase some of the assets of the business if the agreement is terminated. The strong version, where the franchisor must repurchase, generally arises in cases with high capital investment by the franchisee. In addition, it usually only applies if the franchisor terminates the agreement. It appears that this type of clause may explain how franchise contracts can support some highly specialised investments on occasions. We will return to the issue of specialised assets in Chapter 7.

All franchisors except for Wimpy, Austin Rover and Ford impose professional-limitation clauses on their franchisees. For periods of between six months and two years, ex-franchisees must not compete in the same line of work within their old franchise territories. This does not really act to preserve local goodwill for the franchisor. Rather it deprives the franchisee of the benefit of his local goodwill. Thus, the investment the franchisee has made in building up his business becomes worthless to him upon exit or termination. This clearly constitutes an hostage in the sense of Klein (1980). The franchisor can impose this (intangible) capital loss upon the franchisee if the franchisee contrives an exit, gives notice to quit, or has his agreement terminated because his performance deteriorates. The franchisee retains this wealth and can sell it, if he transfers the business with the franchisor's consent, providing his performance is satisfactory. It is worth noting that the hostage is an ugly princess in Williamson's terms (1985, 177). The franchisee's local goodwill is not transferable to the franchisor who therefore values it at zero and has no incentive to opportunistically appropriate it. As long as the franchisor remains in the district on a site, he will retain his own goodwill (albeit disrupted).

It can be noticed that in the three cases (Wimpy, Austin Rover and Ford) where the goodwill hostage is not posted, it would pose major problems in franchisee recruitment. Wimpy specifically aims its recruitment at investors who may well need to feel able to sell their investment and open something else in the district. Vehicle dealers have a lot of human capital in the form of trading and maintenance skills when they take on a franchise; it would be foolish of them to jeopardise the possibilities for its continued investment in the motor trade to which it is specific.

Finally in this section, we may briefly comment on the definition of territories and on an aspect of the settlement of disputes.

Territories are defined in 10 of the franchise systems. Apollo, Garage Door and Mobiletuning confer the right to use a trade mark in an area. Avis and Budget award the right to use their systems. Young's, Bally and Nationwide undertake to locate no other outlets within the district surrounding a franchisee. The car dealers have marketing areas in which only they can advertise or have showrooms but within which anyone can sell. In the other cases, implicit understandings arise over territory. The point is that territories are weakly defined as non-exclusive sales areas.

Only seven franchise systems, out of the 15 providing agreements, have provisions for compulsory arbitration in their explicit contracts. These are Apollo, Garage Door, Pizza Express, Bally, Budget, Midas and Olivers. This tells us that in at least some cases there is an intention to use neoclassical contractual governance (Macneil, 1978). We return to dispute resolution issues after the next section on the implicit contract.

The implicit contract

Apart from their written agreements, or explicit contracts, franchise relationships are also characterised by a number of business understandings that develop over time. These rely on the long-term value of the relationship to the parties for their enforcement and are regarded in this study as implicit aspects of the contract between franchisor and franchisee. In this section, we restore Computerland, Yves Rocher and Cure 30 to the sample of franchises; although no written agreement was provided for analysis, it proved possible to obtain data on the implicit contract. Barstock is still omitted as it has not yet appointed franchisees. In all of the 18 cases at least some important aspects of the contract can be shown to be governed by understandings.

Table 6.5 shows the important aspects of implicit contracts for the franchises. We have already considered some aspects covering advertising and leases; these details are not repeated in this section. The allocation of training costs is not covered in the table as it is straightforward: the franchisor normally pays for training excluding the franchisee's travel and subsistence costs, and this results from an understanding if it is not in the written agreement. The four headings in Table 6.5 select major aspects of the contract which are often left implicit.

Transfer prices for products or other transactions within the franchise are not normally specified within the agreement, or, if they are, they are subject to revision. In most cases it is the price of products which must be determined as events unfold. In the case of the vehicle-rental companies (Budget and Avis) the prices to be used for vehicle transfers within one-way rental schemes must be determined. In the case of Nationwide, surveillance

Table 6.5 Implicit Contracts

	<u>Other</u> <u>sales</u>	<u>Transfer</u> <u>prices</u>	<u>Territory</u>	<u>Financial</u> <u>support</u>
Apollo	Poles	Products		
Garage Door		Products		
Avis		One way		
Mobiletuning	Servicing	Products		
Young's	Accessories	Products		Buy back
Wimpy		Products	1st refusal	
Pizza Express		Products	1st refusal	
Bally	Hose	Products		Buy back
Nationwide		Equipment		
Budget		One way		
Midas	Clutches	Products	Location	
Olivers		Products	Location	Royalty relief
Quikframe		Products	Location	
Cure 30	Building	Products		
Computerland		Products		Buy back
Yves Rocher	Jewels	Products		Discret- ionary
Austin Rover		Products	Relaxation	
Ford		Products	Relaxation	

NOTE: blanks indicate that the implicit aspect was
not observed

and similar equipment loans are made between franchisees at prices which are not set in the agreement. In all cases, the franchisor is trusted to set prices which franchisees will find acceptable.

The principles governing transfer pricing have already been discussed in Chapter 4. Product franchisors like Austin Rover, Ford, Bally, Yves Rocher and Apollo may use mark ups to provide the franchisor's income as long as

use mark ups to provide the franchisor's income as long as the product remains competitive in the market place. The other franchisors must supply products at cost or risk putting franchisees at a cost disadvantage. The difference is due to the possibility of input substitution in the brand and specialised-input cases. Deviation from these principles is checked by potential losses of revenue to the franchisor. These would occur through the loss of current sales and through difficulties in the recruitment of new franchisees if their profits were likely to be depressed.

In eight of the franchises, the franchisor accepts that franchisees regularly break their franchise agreements and have sidelines to the main business. Thus, we find Apollo franchisees selling curtain poles, Mobiletuning franchisees offering full-service work, and Cure 30 people picking up flood-damage work. Young's franchisees sell accessories such as bouquets, those with Bally sell ladies' tights, Midas outlets undertake clutch work, and Yves Rocher franchisees often sell costume jewellery. In all of these cases, the franchisor receives no income from sales occurring under his brand name. Yet he chooses not to enforce the agreement which rules out such sidelines.

Mr. Rowntree of Mobiletuning explains why this is so when he states that it is not worth his while to police minor infringements of the agreement. Whilst it is of value to him to preserve his brand image without dilution by other services, the benefit is not large enough to warrant the monitoring costs that are implied. Franchisees come to understand that they do not push their luck, however, by developing such sales to the point where it becomes worth while for the franchisor to act. This is because they do not wish to disrupt their long term relationship with the franchisor.

Some franchisors award territories to franchisees as part of the implicit contract. In part, this may be due to the worry that explicit contractual details may cause problems with restrictive trade practices legislation. In fact, both sorts of agreement could be problematic. The important observation is that the implicit version appears not to cause problems within franchise systems, unwritten contracts are respected over territories. /

Midas, Olivers and Quikframe have understandings that they will not locate additional branches in a manner likely to threaten the livelihood of existing franchisees. Wimpy and Pizza Express give existing franchisees first refusal on a new outlet which is likely to jeopardise their incomes. It may well be that first refusal opportunities are used to make new locations acceptable in other cases as well. New outlets may cause problems when they are desirable from the viewpoint of aggregate profits but lower the profits of an existing branch. A merger solves the difficulty. The franchisor wishes to ensure that he does not develop a reputation for misleading franchisees over their likely profits or else his recruitment of new franchisees may suffer.

Ford and Austin Rover have special undertakings over territories. They will move from the usually rigid positions over territory in the case where a remote rural area cannot support a dealer who sells one brand. This was dramatically illustrated by one of the Ford franchisees interviewed in this study, who proclaimed that he had seen a combined Ford and Austin Rover showroom in a remote part of Wales. The two companies are great rivals normally. In such instances, some representation is valued by the maker and the remoteness of the location makes it unlikely that the shared facilities will compromise the brand image of the wider network.

Six franchisors (Young's, Wimpy, Olivers, Bally, Computerland and Yves Rocher) make it clear in the case studies that they offer some financial support to weak franchisees. Young's has bought shops where franchisees have virtually failed: as in the case of the Newcastle Pronuptia shop. Wimpy stands ready to do the same, although Mr. McGlashan states that this has not been necessary. Olivers applies discretionary sales royalty relief if profitability projections do not materialise. Yves Rocher gives discretionary product discounts if profits are poor for the franchisee. Bally and Computerland have also bought out weak franchisees. In all cases, the worry is that failure will harm the franchisor's reputation and make his recruitment of new franchisees more difficult.

The behaviour of the franchisors is consistent with the analysis of fee schedules in Chapter 5. This suggested that uniform charges could leave some franchises with profits and some with losses. The franchisors will try to correct the losses if they are worried about reputation and recruitment, which is what we observe. We can also note that we discussed the use of lease charges earlier, where some attempts to remove or improve profits can be detected. It is important to realise that this type of information is very difficult to obtain as franchisors may not wish to advertise their practices for fear of encouraging strategic behaviour among franchisees. Yet there are at least these attempts to correct some of the effects of applying uniform charges. It appears to be commoner to correct for underachievement of projections.

Revision of agreements

In a number of the case studies there is evidence of revision of the explicit contract. Sometimes implicit

aspects of the contract arise due to mistakes made or rigidities created drafting the agreement. It may, for example, prove impractical to enforce franchise territories based on sales. This was the case for Nationwide, where franchisees were required to deal with clients only from their specified territory. This proved awkward and costly as it meant that clients could not choose agents located away from their residences but in areas in which the work is required. An understanding arose whereby the territory came to be seen as an advertising area. Eventually, Nationwide turned the understanding into a clause of a revised version of the agreement.

It is possible to move implicit aspects of the contract into the agreement if there is no advantage in leaving them imprecisely defined. This does not happen with such 'things' as the franchisor's support of weaker franchisees. In this case, the franchisor does not wish to encourage franchisees to misrepresent profits to gain support and may wish to distinguish between franchisees on qualitative grounds. There is a difference between two franchisees with the same profits where one maintains the franchisor's standards better. The franchisor will know when this arises but may not be able to fully specify the circumstances in the agreement. Interestingly, Yves Rocher has dropped a system where initial support for franchisees was set in the agreement and has adopted an implicit aspect in this area.

A further example of an implicit aspect which moves into the agreement occurs in the case of Midas. Initially, a need to control leases which is perceived by the franchisor became the subject of an understanding. However, future agreements will contain clauses on this. In a number of cases, drafting mistakes are fairly rapidly corrected in the agreement without the need to formulate

stop-gap implicit aspects of the contract. This can be done as new franchisees join, although it leads to differences between those of different vintage. It is clear, either from the comments of franchisors, or from handwritten corrections or revision codes on the copy of an agreement supplied for this study, that Bally, Nationwide, Austin Rover and Ford have all revised details in this way. Also, both Cure 30 and Computerland have altered their fees.

Occasionally, there is evidence in the cases that franchisors are willing to tailor their agreements to suit particular business conditions. Young's permits E₃ to operate a fractional franchise with tailored conditions. Garage Door waives a requirement to impose a manager on franchisees if performance is poor in the agreement with B₃. Tailoring of agreements shows sensitivity to details of particular business sites.

The case studies also reveal changes in implicit aspects of the contract. In particular, the Computerland franchisees have successfully moved the franchisor over to greater reliance on product sales for his income. Transfer prices are governed by market discipline as franchisees will not purchase at prices above market ones. This arrangement was voluntarily selected by franchisors and franchisees as a means of motivating the franchisor to efficiency in his buying operations. Computerland's franchise agreement uses what Williamson calls an 'high-powered incentive' at this point (1985, 76): the franchisor's income is linked closely to market rewards.

A further example of revision of implicit contracts arises with Olivers. Over time it has come to be accepted that franchisees may move away from company supply lines if there are good reasons put forward for this, according to Mr. Allan.

We may conclude in this section that franchise contracts are not static items which are written, once and for all, in the spirit of agency analysis. Rather, they are subject to selective revisions by franchisors and franchisees. These revisions typically recognise constraints which may have been unperceived at the start of contracting or seek to avoid some transaction cost. The revisions move franchise participants towards more preferred positions. They illustrate Coase's (1960) theorem on bargaining: where negotiations do not run into high transaction costs they will move economic agents in the direction of greater efficiency.

The governance of franchise relationships

We began this chapter with a discussion of Macneil's (1974, 1978) division of contract into classical, neoclassical and relational types. Each of these forms has implications for the manner in which contracts are defined, including the precision with which this is done, and for the methods of enforcement and monitoring used by contract participants. The term 'governance' is used to describe enforcement and monitoring by Williamson (1979). We now analyse the governance aspects of franchise contracts. Barstock is again excluded from comparisons as it has not yet signed agreements with franchisees.

Table 6.6 shows details of monitoring and enforcement for the 18 franchise contracts under analysis. Before considering this, we should note that we have already discovered that franchisors rarely impose penalty clauses on franchisees and that all franchisees are required to submit regular sales reports. Table 6.6 shows a number of further regularities.

Table 6.6 Enforcement and monitoring

	Inspection	Franchisor outlets	Quotes contract	Termination
Apollo	Yes	No	No	'Few'
Garage Door	Yes	One	Rare	No
Avis	Monthly	Yes	No	No
Mobiletuning	No	One	Rare	2
Young's	6 weeks	Yes	Rare	6
Wimpy	Monthly	Yes	No	No
Pizza Express	Yes	Yes	Rare	1
Bally	Yes	Yes	No	No
Nationwide	No	Yes	Rare	No
Budget	Yes	No	Rare	No
Midas	Yes	Yes	Rare	No
Olivers	Weekly	Yes	Rare	1
Quikframe	No	One	Franchisees	No
Cure 30	6 months	One	Rare	1
Computerland	Rare	No	No	No
Yves Rocher	Yes	No	Rare	1
Austin Rover	Yes	No	Rare	'Few'
Ford	Yes	DDS	Rare	'Few'

NOTE: Table 6.6 is continued overleaf

Table 6.6 continued Enforcement and monitoring

	Voluntary exits	Buy out
Apollo	No	Yes
Garage Door	No	No
Avis	No	No
Mobiletuning	Yes	No
Young's	Yes	Yes
Wimpy	No	No
Pizza Express	No	No
Bally	No	Yes
Nationwide	Yes	No
Budget	Yes	Yes
Midas	Yes	No
Olivers	No	Yes
Quikframe	Yes	No
Cure 30	No	No
Computerland	No	Yes
Yves Rocher	No	Yes
Austin Rover	Yes	No
Ford	Yes	No

Many of the franchisors make periodic visits to franchisees' premises to assess the quality of the local service. Mobiletuning, Nationwide and Quikframe do not do this. Quikframe has lost its franchisees perhaps through lack of support for them. Nationwide has intensely private relationships between investigators and clients which makes inspection extremely difficult. Mobiletuning has regular training contacts with franchisees which make inspections unnecessary. Apart from these special cases, inspections are regularly undertaken at intervals ranging from one week to six months.

Inspections operate partly to reveal whether standards are adequate in the franchisor's opinion, to give advice to the franchisee and to give the franchisee an opportunity to raise any queries he may have with the franchisor. Franchisees in the cases always talk of this type of meeting having predominantly consultancy characteristics. The franchisor does not appear to act in an heavy-handed manner over inspection and physical monitoring. Instead, the inspection acts as the principal point of regular contact between franchisor and franchisee.

This view of inspections is reinforced by the methods used. Only three franchisors spot check franchisees. Budget admits to having used mystery-buyer checks in which they send an inspector to pose as a customer at a franchised outlet. In addition, Budget allowed franchisees to spot check one another following requests from the franchisees for this facility; the terms of this check are specified. The Budget franchisees are concerned that poor quality neighbouring outlets will confer reputation loss upon them; J₁ is worried about a neighbour's standards, for example. It is unusual for franchisees to check each other in this way. The practice is reminiscent of the behaviour of Chinese bargees, who used to hire their own overseers according to Cheung (1983). Bally also admits to mystery-buyer techniques. Wimpy may spot check on franchisees according to the franchise agreement. In the remaining cases, appointments are made for the franchisor's visits which increases their consultancy orientation.

In 13 of the cases in Table 6.6, the franchisor runs his own outlets. In the cases of Garage Door and Quikframe, the franchisor runs one outlet and is effectively a small businessman with franchising interests. Nationwide is in much the same position,

although Mr. Withers runs as Associated Detectives in some areas. Mobiletuning and Cure 30 operate in their 'home' areas as a cost and systems check. Ford operates the Dealer Development Scheme (DDS in the table) in order to establish a new dealer. Wimpy owns 'flagship' operations to promote its franchise. The other six have taken decisions to operate company and franchised shops. We shall return to the reasons for franchising later.

For the moment, we may note that most of the franchisors have direct operating experience which enables them to assess franchisees' performance against their own standards. This type of close-grained knowledge of the industry in which they operate may explain why monitoring is often quite a casual affair: franchisors are able to pick up quality signals with great efficiency; It is significant that in the case of the car makers, where the franchisor's experience is not specialised in retailing, franchisees are subjected to dealership audits. According to S, such audits occur in the case of Ford every three years and consist of an extensive check of dealership operations. Most of the business-format franchisors refer to their own experience as of considerable importance in monitoring their network.

Table 6.6 also contains information on how contracts are enforced (by franchisors in the main). Franchisors vary over the extent to which they cite the agreement in their dealings with franchisees. There is a danger that the incidence of this may be misinterpreted. A total of 12 franchisors report citing the agreement. In addition, the Quikframe franchisees quoted their agreement to the franchisor. Generally, when quoting the contract occurs it is in extreme cases where all else has failed. Thus, Mr. Rowntree tells us (Mobiletuning, Case D) that if it came to quoting the agreement he would rather find a means of ending a franchise relationship amicably. In the case

studies, franchisors admit to either very few or to no instances of citing of clauses. As with Mr. Rowntree, citation is viewed as a sign that the relationship has come to the end of its useful life.

Franchisors generally regard the prospect of regular clause citation with horror and argue that they aim not to be heavy handed in their interpretation of the agreement. With very few exceptions, franchisees experience and expect to continue to experience a business-like rather than legalistic relationship. This is fully consistent with Macauley's (1963) empirical study of US businessmen: litigation is relatively rare and a legalistic orientation is believed to be harmful to good 'management relations'. Macauley was often told that business problems can always be solved 'if the lawyers ... are kept out' (Macauley, 1963, 61).

Franchisees quoted their agreements in the case of Quikframe, arguing that support was not adequate and that the agreement could not be enforced as it was not registered with the Office of Fair Trading. This is consistent with the argument above: citation occurred when the relationship was finished and it just remained to sort out a formal severing of relations. Only one other franchisee in the entire study contemplated the use of legal sanctions: P₂ thought there to be no necessary conflict between, for example, taking the franchisor to court and continuing to deal with him. One franchisee (P₁) did not hire a lawyer to check over his agreement prior to signing, but preferred to base his judgement on personal interaction with the franchisor's staff.

The weak status of the written agreement in regular dealings within franchises is confirmed by the use of termination clauses. In general, these are rarely used by franchisors. In total, nine franchisors had never

terminated an agreement. Of the remainder, franchisors admit to very infrequent use of their right to terminate. They have a right to terminate at will if there is dissatisfaction with the franchisee's performance. Termination is normally used only for extreme failures such as non-payment of the franchise fee; although even here there can be amazing tolerance as in the case of Budget. Franchisors devote efforts to screening franchisees at recruitment to avoid later problems. Termination is costly in terms of lost income and damage to the franchisor's reputation.

Franchisors also admit to a few cases of voluntary exit by franchisees. Table 6.6 shows that, in the case studies, eight franchise systems had lost franchisees in this way. Mostly there were small numbers of exits, although Quikframe has lost all its franchisees. Voluntary exits occur through non-renewal of agreements, notice to quit under the rare agreements which allow this, or through claims that the franchisor has broken the agreement. The litigious route was taken by Quikframe franchisees and by Young's franchisees at the time when the franchisor failed. Generally, franchisees stay with their agreements if they are of benefit to them, even if they harbour odd grievances.

Table 6.6 also shows that a number of franchisors have either bought out franchisees (or encouraged them to sell to a third party) in situations where termination might otherwise be contemplated. In the cases, franchisors express preferences for this type of solution to serious disputes where the franchise relationship is judged to have broken down.

Franchisors regularly state in the cases that they value their reputations for fair dealing very highly. They appear to worry about discouraging franchisee

recruitment by creating an impression of harassing outlets. This is a reasonable worry as an outsider would find it difficult to judge whether a franchisor's grievances were justified in some case. An outsider may judge regular attention to the detail of agreements as reflecting poor management, particularly as quotation of contract appears to be rare in the economy at large.

It is instructive to consider the nature of disputes which respondents mentioned in the case studies. These are shown in Table 6.7, which shows that there are grounds for more use of the agreement than is observed: it is not as if franchises always run in a problem-free manner.

Table 6.7 shows that both franchisors and franchisees have disagreements over standards of operation. Sometimes dissatisfaction can be quite fundamental, as in the case of Nationwide franchisees who often feel that they receive minimal levels of consultancy support. Bally franchisees believe they are given unrealistic sales targets, to give another example. Franchisors are dissatisfied with franchisees' level of sales levels of sales in a number of cases, or with such things as attempts to design premises in an unapproved manner. Failure to pay fees by the franchisee and occasional efforts at purchasing non-approved products also cause disputes with franchisors. Otherwise, franchisors and franchisees report 'gripes' (minor disputes over such things as the style of a current promotion).

Generally these disputes are resolved in the course of the franchise relationship and are not allowed to disrupt a long-term association which is of value of both parties. Sometimes something is not ideal, as when Mobiletuning franchisees do not always achieve their

Table 6.7 Disputes*

	Unapproved purchases	Performance		Fees	Gripes
		Franchisee	Franchisor		
Apollo	Yes			Yes	
Garage- Door		Yes			
Avis					Yes
Mobile- tuning		Yes			
Young's	Yes	Yes	Yes	Yes	
Wimpy		Yes			
Pizza- Express		Yes			Yes
Bally		Yes	Yes		Yes
Nation- wide		Yes		Yes	
Budget		Yes			
Midas			Yes	Yes	Yes
Olivers					Yes
Quikframe			Yes		
Cure 30		Yes		Yes	
Computer- land			Yes		
Yves- Rocher	Yes		Yes	Yes	
Austin- Rover			Yes	Yes	
Ford	Yes				

*Note: blanks indicate that an area of dispute is not apparent in a case

minimum number of monthly engine tunes. Nevertheless, their continued trading is of value to Mr. Rowntree. It is notable that there is a correlation between non-payment of fees and the franchisor exercising his right to terminate: this failure ends long-term benefits for the franchisor unless it is temporary.

Summary and some conclusions

In this chapter, the franchise contract has been shown to contain explicit and implicit aspects which adapt over time. Writing an agreement is not a barrier to subsequently devising a better way of ordering an element of the franchise relationship. Given that uncertainty exists and that contracting parties may have only bounded rationality it is preferable to leave some aspects of the contract implicit or subject to revision as events unfold. Some aspects of contract, where the franchisor requires to use his discretion, are best left implicit from his viewpoint.

The governance of the franchise contract rests upon the long-term value of the relationship and tends to eschew the use of legal sanction. Earlier findings of Macauley (1963) are supported over the avoidance of close quotation of agreements. This practice is not just bad business etiquette but is perceived as positively harmful to efficient operation due to reputation effects on the franchisor. Even where relationships have broken down, franchisors often avoid termination of agreements by buying back franchised outlets and refranchising them. Governance is bilateral: it does not involve third parties like courts or arbitrators.

The explicit contract acts as an initial statement of the contractual relationship which is then augmented. Together with the bilateral governance within the

franchise system, this makes the franchise relationship relational in the terms of Macneil (1978) as is asserted by Goldberg (1980). It should be noted that whilst the governance relies on the long-term value of a market relationship it is not characterised by the continual exercise of market sanctions. We do not have a case of market governance as this is defined by Williamson (1985, 73) where the term refers to the exclusive use of instant market sanctions over poor performance. /

An earlier question has now been answered. Lease control and implicit support for weak franchisees give some relief from the impact of uniform fee schedules which we identify in Chapter 5. This relief does not appear to be extensive. There is a case for additional research into this area.

We have also identified an hostage, which is an ugly princess in Williamson's (1985) terms. Franchisees are extensively subject to professional-limitation clauses which give franchisors the ability to isolate the franchisees' local business goodwill if the agreement ends. This happens as the franchisee is prevented from competing in his old area for a period of time. However, lease control acts to minimise the transaction costs faced by franchisors in maintaining their brand presence in an area if the agreement ends. We consider whether lease control also has hostage aspects as argued by Klein (1980) in the next chapter.

CHAPTER 7

Specialised Assets and Hostages

In the last chapter we observed that professional-limitation clauses act as hostages put up by the franchisee. We also observed that lease control minimises the disruption costs of termination for the franchisor but questioned whether the control has hostage properties as alleged by Klein (1980). In this chapter, we analyse asset specificity within franchise relationships and are able to develop these observations.

Initial investments and specificity

Table 7.1 shows some very rough estimates of the initial investments required to start as a franchisee in 18 of the franchise systems, along with details of sunk expenditures which are incurred. Barstock is again excluded from comparisons as it is not yet operating. The estimates come from cost projections supplied by franchisors and from discussions with franchisees. Some crude assumptions have been made: for example, we calculate the second-hand value of equipment as one half of its original cost. Setting-up costs normally cover obtaining and adapting premises, purchasing tools and other equipment, and advertising the launch of the new business. Some of these expenditures become unrecoverable or 'sunk' once they are made. Sunk costs refer to assets which are specialised and where at least some of their value cannot be transferred out of the business. That part of the franchisee's wealth remains intact if he continues to operate his business or if he sells it as a going concern.¹ The wealth is lost if the business fails for one reason or another.

¹ The specialised assets are still subject to physical depreciation.

Table 7.1 Initial investments and sunk costs (£'000)

	Initial Investment	Lump Sum	Sunk Costs
Apollo	20+	4	21
Garage Door	10	2.5	3
Avis	30	15	30
Mobiletuning	10-16	3	4-6
Young's	45-78	10-15	22-45
Wimpy	500	10	450
Pizza Express	200	12.5	160
Bally	120*	No	70
Nationwide	3.5	7.5	9
Budget	30	20	35
Midas	100	10	60
Olivers	300	10	280
Quikframe	Not known	6	6+
Cure 30	10	16	6
Computerland	230	13-20	60
Yves Rocher	54*	7.5	31
Austin Rover	Not known	No*	8+
Ford	Not known	No*	8+

***Note:** franchisees pay stock charges in these figures

A good example of a specialised asset is the practice of leasehold improvement. Almost all high-street shops take their premises on leases. When a shop is first taken it must be altered to suit its purpose by the retailer. Generally, he moves walls, fits a shopfront and designs

the premises to accommodate his business. If his business were to fail he would have the option of selling the unexpired portion of the lease (with the landlord's consent) or setting up another business himself on the site. The landlord gives no compensation for the cost of the adaptations. The value of the improvements is generally very low away from their original purpose as retail requirements are normally for highly specialised premises. Asset specificity is created as soon as premises are refitted. The analysis is similar if the retailer owns his own premises, except that no landlord's consent is required for change of use.

The figure for initial investment in Table 7.1 does not include the initial lump-sum franchise fee which is shown in a separate column. The lump sum plus an estimate of the non-recoverable part of the initial investment gives the sunk cost in the third column of the table.

It is noticeable that in all cases where we have data, except for Cure 30 and Nationwide, the initial investment is well in excess of the lump-sum payment. Both exceptions arise because the franchisor provides much of the equipment required by the franchisee in return for the initial fee. In the other cases, the lump sum is often a small percentage of the total start-up costs of a franchise. Wimpy is an extreme example where the lump sum is less than 2% of total costs (i.e. of the lump sum and initial investment combined).

It is also true that the lump sum does not fully account for sunk costs in each franchise, except in the case of Cure 30. Franchisees could retrieve the second-hand value of some equipment in the case of Cure 30. In all cases, including Cure 30 there still would be sunk costs if there were no lump sum charged. Thus, the lump sum is not required to create sunk costs if it is of some

value in the franchise relationship to do this. In the analysis of Chapter 5, the lump sum is explained in terms of the franchisor's sunk costs. We did, however, state that the lump sum could also be a commitment device acting on the franchisee as an explicit hostage with ex-post bonding characteristics. If during the course of the contract the franchisee is seen to cheat in some way, then the bond is sacrificed as the agreement is terminated. The lump sum may act in this way but this is incidental as it cannot be primarily explained by such a role.

Klein (1980, 359) specifically argues that the lump sum as an hostage will equal the expected value to the franchisee of cheating. In Chapter 5, we showed that an observed correlation between the lump sum and the franchisor's sunk costs could be explained by expectations over profitability in a model where the franchisor controls cheating by choosing quality standards and monitoring levels. This suggests that the franchisor's sunk costs are the largest sum that may safely be passed to him at the start of a contract whatever the reason for the transfer, from the franchisee's viewpoint. From the franchisor's point of view, his sunk costs must be recovered early on or he creates an appropriable quasi rent as defined by Klein, Crawford and Alchian (1978). Since lump sums are well explained on that basis and since they are usually a small part of sunk costs for the franchisee we may conclude that any ex-post bonding function is incidental. The franchisor's need for protection from opportunism over his sunk costs, rather than the creation of a penalty to suppress cheating over quality on the franchisee's part, binds on the lump sum.

Lease control and specific assets

In Chapter 6, we observed that some franchisors control the leases of franchisees, or offer them

assistance in obtaining leases. We argued that this control, could minimise the costs of re-establishing a branch in an area if the franchisee decides to leave the network. The incidence of lease control suggests that it is important where franchisors face high transaction costs attached to obtaining high-street sites. Klein has argued that lease control enables franchisors to render leasehold improvements worthless to the franchisee so that his sunk costs may be increased:

'..... the franchisor can require the franchisee to move and thereby impose a capital loss on him up to the amount of his initial non-salvageable investment. Hence a form of collateral to deter cheating is created'. (Klein, 1980, 359).

As with the lump sum, we may show that the main purpose of lease control is not to act as a discipline on the franchisee. This does not rule out its incidental operation in this manner.

The problem with Klein's view is that so few franchisors use strong forms of lease control. In the sample studied here, only five have the facility to insist on acting as superior lessee to the franchisee: these are Apollo, Pizza Express, Bally, Midas and Olivers. Yet, many of the franchisors face similar problems in controlling franchisee incentives. Why do not Avis and Budget, in particular, make this rule over leases? Both of these have franchisees who must work from a set location. It has already been argued that Wimpy and Young's may be able often to rely on the franchisee needing this practice so that they do not insist on it.

However, when we look down the column headed 'Sunk Costs' in Table 7.1, we see that these can be high proportions of the initial investment regardless of whether or not lease control operates. To take the

example of Avis, a franchisee rents a site on which he fits out offices using Avis's designs. If he were to decide to leave the system, these designs would be largely useless to him even without professional-limitation clauses that stop him from trading in his old area. Wimpy franchisees have no professional limitation-clauses yet if they wished to leave their network there would be little they could do with the property other than sell the unexpired lease to Wimpy or to new franchisees. The counter-service restaurant is designed in an highly specialised way with most items being trade marked. With franchises, leasehold improvement and much equipment becomes highly specialised by virtue of trade markings. Franchisors do not appear to need the controls on leases suggested by Klein (1980) to create sunk-cost penalties.

We should also note that other controls which the franchisor has at his disposal indicate that lease control is not needed to support asset specificity. Professional-limitation clauses, in particular, mean that franchisees cannot continue in the same line of business within their old franchise area. This renders leasehold improvements worthless outside of continued operation as a franchisee or sale of the business as a going concern with the franchisor's approval.

We must be careful to show that we are not double counting hostages at this point. Franchisees set up businesses which are greatly affected by Williamson's (1985, 61) fundamental transformation. Before they sign franchise agreements their assets could be used to support transactions with a number of different franchisors or to start independent businesses. Once they sign, they become locked into transactions that are supported by highly specific assets. The franchisee regards his expenditure and efforts as an investment in building his local business. Thus, his assets are the local goodwill he

builds up plus the tangible capital committed to the franchised outlet. At any time the value of these is affected by expected returns, by depreciation and by details of the franchise relationship. The trade marking of equipment and premises, and professional-limitation clauses restrict the market for these assets to buyers approved by the franchisor. Professional-limitation operates on immoveable physical assets as well as upon goodwill.

Hostages in franchise contracts

Franchisees appear to post hostages which are implicit and which equal the entire value of sunk investment in their businesses. The sunk investment consists of trade-marked equipment and designs, most leasehold improvements, built-up goodwill, and lump sum franchise fees. The sunk costs shown in Table 7.1 are only an estimate of the value of these hostages as they are just historic cost of things like leasehold improvements and trade-marked equipment. What is really needed is an estimate of the franchisee's expected income flow if he remains within the network or sells to an approved buyer compared with income if the agreement is terminated.

The franchisor does not post hostages in contrast to the franchisee. His sunk costs are recovered when these are attributable to individual franchisees.

The franchisee posts an hostage whereby he makes the value of his assets subject to the franchisor's right to terminate the franchise agreement, and subject to his own efforts within the relationship. It may be enquired whether the franchisee thereby creates an appropriable quasi rent for the franchisor. Specifically, are there situations in which the franchisor could contrive a

contractual breach by the franchisee which would result in the franchisee having to accept a below-market price for his business? In practice, the danger might be that franchisors would use this as a bargaining lever to revise franchise fees upwards as in the parallel situations analysed by Klein, Crawford and Alchian (1978).

Generally it is not possible for the franchisor to contrive such a situation. The franchise agreement (or perhaps a separate property agreement) usually gives a method of independently valuing assets such as trade-marked equipment and unexpired portions of any lease which the franchisor may or must purchase if he terminates the agreement. Table 7.2 gives a breakdown of asset transfer within the 15 franchises for which we have agreements. Table 7.2 expands on Table 6.4 in Chapter 6. All franchisors except for Garage Door, Avis, Mobiletuning and Budget give the franchisor the right to purchase the franchisee's assets at termination. In five of these cases, arbitration clauses in the agreement govern the transfer of assets. In two of these cases (Bally and Olivers) definite scales are laid down for the price of transfers. In two further cases (Austin Rover and Ford) scales are laid down as parts and vehicles are sold back at list prices. In yet two further cases (Garage Door and Budget) compulsory arbitration clauses could anyway govern asset transfer even though no right to buy is taken by the franchisor. This leaves Young's, Wimpy, and Quikframe as franchisors who undertake to buy at 'fair market value' (as the Wimpy agreement states). Fair market value could be defined in the courts. Nationwide has the right to take over any property lease at termination, if the landlord agrees, but this would be at the landlord's

price. Avis and Mobiletuning are alone in having no conditions governing the franchisor's purchase of assets.²

Table 7.2 Asset transfer

	Compulsory Arbitration	Product transfer	Equipment transfer	Lease transfer
Apollo	Yes	Yes	No	Yes
Garage Door	Yes	No	No	No
Avis	No	No	No	No
Mobiletuning	No	No	No	No
Young's	No	Yes, FM*	No	Yes, FM
Wimpy	No	No	Yes, FM	Yes, FM
Pizza Express	Yes	No	Yes, FM	Yes
Bally	Yes	Yes, P*	No	Yes, P
Nationwide	No	No	No	Yes, P
Budget	Yes	No	No	No
Midas	Yes	No	Yes	Yes
Olivers	Yes	No	Yes, P	Yes, P
Quikframe	No	Yes, P	Yes, FM	No
Austin Rover	No	Yes, P	Yes, FM	No
Ford	No	Yes, P	Yes, FM	No

* Note: FM = fair market prices; P = set price/s

In all cases, we would expect post-contract opportunism on the part of the franchisor to be limited by the losses he would suffer if he were to terminate the agreement. Thus, in the cases of Avis and Garage Door, in

² Must-buy clauses in some agreements (e.g. Bally) never cover all assets (or they could erode hostage taking).

particular, the possible gain in obtaining either cheap assets or increased royalty payments as a result of successful hold up is most likely outweighed by longer-term losses. These losses are either in the form of damage to the franchisor's reputation or in the simple form of the loss of a particular franchisee over the longer term as specific assets depreciate.

Thus, the hostages posted by franchisees have the characteristics of ugly princesses. The franchisor can impose capital losses upon the franchisee but will not do so opportunistically because he will normally be made no better off by attempting to contrive a termination. This is to say that the hostage is not valued in itself by the franchisor.

The hostage is likely to be given up in situations where the franchisor may be led to believe that a site is not viable and he may not choose to buy specific assets. The franchisee had better be sure that he is as able as he claims when applying for a franchise, if he is aware of the possibility. Franchisees often point out in the cases that they are aware of having much of their wealth tied up in this way in their business (for example, H₂). This suggests that the hostage may have ex-ante screening functions as well as ex-post bonding characteristics: it helps franchisors to select franchisees who are most likely to succeed. We return to this issue in the next chapter.

The other situation where the hostage may be taken is if the franchisee cheats on standards of operation. In this case, the franchisor may choose not to buy at least some assets as a simple penalty. It may even happen that he does this when it is against his immediate trading interests if the existence of the rule enables him to

economise on other enforcement devices like exhortation, or makes other devices more effective.

Summary and conclusions

The franchise relationship is characterised by unilateral hostage posting. Franchisees stand to lose substantial parts of their initial investment if they fail to perform consummately within their networks. The sunk cost involved is larger than would occur if the franchisee operated independently, mainly due to the trade marking of equipment and property fittings and professional-limitation clauses. The franchise relationship appears to increase asset specificity at the retail end of the business compared with independent retailing.

Franchisees are protected from opportunistic seizure of hostages by the reputation of the franchisor and by valuation clauses attached in leases to some specialised fittings and to some stocks. These protections help to support the increase in specificity described above.

CHAPTER 8

Why Franchise?

In Chapter 2 it was argued that existing theories of franchising concentrate on explaining observed franchise relationships rather than predicting when they will arise. We may add that the observations are often somewhat casual. Rubin (1978) may be partly exempted from this criticism as he suggests that franchising will be chosen, in preference to running company-owned outlets, where distances are great and monitoring is costly in a retail chain; this is likely to be a special case, however. Mathewson and Winter (1985) are bad offenders in taking the franchise relationship as given and deriving necessary conditions for its principal feature, which they take to be profit sharing. For Rubin, monitoring becomes less costly as a result of profit sharing so that his theory is really subject to the same criticism which may be made of Mathewson and Winter's. The criticism is that the theories are about profit sharing and are not essentially about franchising.

In previous chapters, we have also taken the franchise relationship as given and have concentrated on explaining its key features in the UK. However, we now move over to consider the particular advantages that franchising has for the franchisor and franchisee compared with alternative organisational forms, such as full forward integration or independent retailing. Our analysis is based more on a comparative institutional approach in this chapter (Demsetz, 1969). We are able to predict some of the circumstances that will lead to the use of franchising.

Franchised and company-owned outlets

A good starting point to consider the choice of franchising, at least from the franchisor's side of the contract, is to analyse the operation of company-owned outlets by franchisors. As Table 8.1 shows, 14 franchisors operate their own branches directly in one way or another. However, the table also shows that only three franchisors have consciously chosen simultaneously to develop franchised and company-owned outlets in their network: these are Young's, Pizza Express and Bally. Each of these has a policy of considering whether a site should be franchised or directly operated.

Table 8.1 Franchisors with direct operations

	<u>Reason</u>	<u>Location</u>	<u>Policy</u>
Garage Door	Pilot	One site	Franchise
Avis	Market strategy	Concentrated	Franchise
Mobiletuning	Cost check	One site	Franchise
Young's	Investment	Dispersed	On merits
Wimpy	Flagships	Dispersed	Franchise
Pizza Express	Investment	Concentrated	On merits
Bally	Market strategy	Concentrated	On merits
Nationwide	Investment	Dispersed	Franchise
Midas	Investment	Concentrated	Franchise
Olivers	Investment	Dispersed	Franchise
Quikframe	Pilot	One site	Franchise
Cure 30	Cost check	One site	Franchise
Barstock	Pilot	One site	Franchise
Ford	Dealer Development	Dispersed	Franchise

We can immediately consider Rubin's (1978, 229) argument that franchisors will tend to operate units which are located near to them as travel-linked monitoring costs are lower for these. There is some support in Table 8.1 for Rubin's view. Bally and Pizza Express operate outlets of this kind. Pizza Express has company-owned restaurants in Central London. Bally's own shops are mainly in London's West End. As pointed out above, however, this is a partial explanation for the choice of franchising: Young's-owned shops are dispersed, and franchisors emphasise different reasons for directly operating outlets.

Table 8.1 tells us that locational concentration is of secondary importance in explaining the choice of whether to operate a site under company ownership. In seven of the cases, cost checking or piloting reasons are given for operating company-owned units (this includes Ford). It is cheaper to have such units concentrated unless dispersion is chosen as a promotional device as with Wimpy.

Generally, the franchisors in the sample have a current business strategy which revolves around franchising. The crucial question is why they have chosen this given their available alternatives. It is to this, along with a similar review of franchisees' choices, that we now turn.

The benefits of franchising to franchisors

Franchisors or their senior managers are able to be clear about the decision to franchise their businesses. In every case, there is a company philosophy over this. The principal reasons given in the case studies are summarised in Table 8.2.

Table 8.2 Franchisor benefits

	Monitoring Cost Savings	Franchisee Commitment	Quality Control	Honesty	Capital
Apollo	High	Yes	Yes	Yes	
Garage Door	Yes	Yes	Yes		Yes
Avis		High			Yes
Mobiletuning	High			High	
Young's		Yes		Yes	Yes
Wimpy	Yes	High		Yes	Yes
Pizza Express	Yes	Yes		Yes	Yes
Bally	Yes	Yes	Yes		
Nationwide	Yes			High	Yes
Budget		Yes			
Midas	Yes	Yes		Yes	Weak
Olivers	Yes	High			Weak
Quikframe	Yes	Yes			
Cure 30	High	High		Yes	
Barstock	Yes	Yes		Yes	Yes
Computerland	Yes	Yes			Weak
Yves Rocher	Yes	Yes			Weak
Austin Rover		Yes	Yes		Yes
Ford		Yes	Yes		

Note: blank spaces mean the benefit is
not perceived by the franchisor.

The franchisors frequently emphasise that monitoring cost savings arise from franchising. There are some striking illustrations of these. Mr. Rowntree operates

his network of 65 Mobiletuning franchisees from one small office in London. He and his partner do this virtually unhelped. This is in a business which is potentially open to fraudulent accounting by franchisees, or by employees if the system were not franchised. The franchisees are located all over Great Britain. Mr. Rowntree is satisfied that his network is not subject to widespread fraud.

Even where a franchisor does not recognise general monitoring-cost benefits explicitly in the table, they are recognised in more specific forms. Thus Avis, Young's, Austin Rover and Ford claim that they benefit from the commitment of franchisees who have their own capital tied up in their businesses. As we shall see, this benefit accrues as reduced monitoring costs.

We should also note that monitoring costs appear to be higher in company-owned outlets in the case studies of Volume 2. The Newcastle Pronuptia shop receives more visits from Young's management since it has been returned to company ownership. Wimpy units also receive more attention in this way, as does the bought-back Bally shop. Mrs. Fitz-Gibbon tells us that the need to temporarily run an outlet that has been purchased from a failing franchisee has used more of Computerland's managers' time than has been used running the entire franchised network.¹

Monitoring costs refer to the management and supervisory time and to the capital and operating costs of monitoring systems used in a franchise. Normally a franchisor monitors the quality of a franchisee's operations and the accuracy of the franchisee's notification of his performance. Column 1 of Table 8.2 tells us that franchisors perceive monitoring-cost savings

¹ Temporary ownership of units is excluded from Table 8.1.

in a general way compared with their alternative options. The particular alternative here is running a fully forward-integrated system of company-owned outlets.

Table 8.2 also shows that virtually all franchisors perceive that franchising enhances the commitment of local management. They believe that franchisees are more committed to the success of a local outlet than a paid manager would be. It is important to consider how this commitment is defined and also how it arises. Only Mr. Rowntree (Mobiletuning) and Mr. Withers failed to mention this benefit.

In all cases, franchisors were asked to consider whether they benefited from the entrepreneurial skills of franchisees. Entrepreneurship was defined as alertness to profitable business opportunities, which accords with Kirzner's (1973) views. The question was always placed in the context of comparing paid managers with franchisees, where the paid managers would receive a profit-linked bonus. Thus, the question was about franchising rather than profit sharing. The answer typically given by franchisors emphasises that franchisees would be more committed to their local business success due to having their own capital tied up in the outlet. The franchisors define commitment in terms of a quality sought in their local management: the franchisee is seen as more likely to possess and to exercise entrepreneurial skills.

This view of franchising as an organisational form which encourages local entrepreneurship is stated very clearly by Mr. Brewer of Avis (Case C). He speaks of Avis's wish to create a network of 'owner-driver' franchisees given the emphasis now placed on developing local markets. In discussions during his interview Mr. Brewer clarified the benefit from this. He described a situation in which a customer might be willing to pay only

£100 to hire a vehicle when the normal rate stood at £120. If the fleet has unused components, maximising contribution to profit or to fixed costs suggests renting out the vehicle for £100 if this covers variable costs. The franchisee will realise this as it is his capital invested in the business. The franchisee will automatically behave in a manner which is ideal from Avis's viewpoint.

I asked Mr. Brewer why Avis could not set a bonus scheme which would lead to a similar result; the answer to this question could have diminished the weight of commitment in the benefits accruing from franchising, if bonus schemes were easy to establish. This question represented an attempt to find evidence against the importance of commitment (negative evidence as Miles and Huberman (1984) put it).

Mr. Brewer pointed out that not all individuals are motivated by profits or possess entrepreneurial skills to the same extent. With franchising, Avis could be sure that profit-motivated individuals were managing the local branches. In general, they would take decisions, like the one described above, in a manner with which Avis would agree. Also, they would be more likely to vigorously pursue new business leads and to keep costs at a minimum. Avis's only real alternative would be to increase its monitoring of company-owned stations so that all these decisions were continually checked with central senior management. This would be very costly, and as Hayek (1945, 524) would agree, may deprive decision making of the advantages of a close knowledge of 'time and place' which the franchisee will have.

The franchisee is committed to making his business succeed and can be relied upon to run his outlet in a manner which the franchisor approves. This includes the

exercise of skills which the franchisee may have but which the franchisor lacks and therefore could not easily supervise anyway. Examples of such skills arise most noticeably in the vehicle dealerships, where franchisees have specialised used-car trading skills.

Commitment results from the franchisee having his wealth tied up in his business. As we saw in Chapter 7, this sunk investment acts as an hostage. From the analysis above, we may now conclude that the hostage has ex-ante screening properties (Williamson, 1985, 168). It allows the franchisor to select local managers who are more likely to be profit orientated and entrepreneurial. It was long ago pointed out by Scitovsky (1943, 354) that individuals differ in their profit orientation. Modern theories of entrepreneurship recognise that entrepreneurial alertness differs between individuals (Kirzner 1973; and Casson, 1982).

Five franchisors mentioned that a benefit to them from franchising is that quality control becomes easier. Four of these are product franchisors (Apollo, Bally, Austin Rover and Ford). In these cases, associated services such as product stocking and maintenance are important in promoting sales. Williamson (1985, 1549) argues that externalities in retail chains, of the horizontal or vertical kind discussed in Chapter 2, lead to pressures for forward integration. Externalities may then be at least partly internalised. However this is not a strong argument for franchising, rather than running company shops. We may note that franchisors do not emphasise direct quality control issues of this kind.

Many franchisors argue that honest reporting is encouraged by their franchise systems. Companies like Mobiletuning and Nationwide emphasise this benefit strongly. These are companies where there is much scope

for cash transactions carried out far from head office. Full forward integration gives employees scope to hide revenues altogether or to disguise them as costs. Franchising reduces the incentive to do this as a franchisee only gains his royalty payment, at most, from obscuring sales. In addition, since he has staked a great deal of his own wealth in his business, he has a lot to lose if he is detected cheating. Given professional-limitation clauses and the hostage nature of much of his investment, he is unlikely to practise high levels of cheating. In this instance the hostage acts as an ex-post bond (Williamson, 19875, 168) in reinforcing honest behaviour.

Finally, in Table 8.2, 12 franchisors state that they enjoy capital benefits from franchising, although four of these are clear that the benefits are weak. We have already acknowledged Rubin's (1978, 225) argument against capital-raising explanations of franchising. An obvious question is why do so many franchisors perceive this benefit?

One answer which may apply to smaller franchises is that a capital constraint operates for the franchisor. It may then be that, whereas banks will lend no more to the franchisor, the franchisee can borrow on his own account or has his own funds. The franchisee is prepared to invest in a branch as long as he can be personally responsible for the performance of the branch. On this view, there is a difference for the franchisee between investing in a share in the franchisor's business and starting up a franchised branch which is not considered by Rubin. If capital is rationed, he may not wish to buy shares given that other capital sources have refused to support the venture. By keeping control of his assets (subject to hostage conditions) the franchisee may be reassured into entering the business. However, most of

the franchisors in Table 8.2 would have no trouble raising capital (Barstock and Quikframe are the exceptions).

Rubin's logically correct argument that the franchisor can always improve on franchising as a capital-raising device by offering shares in all branches to investors may also be reconciled with the views of franchisors. Mr. Watson of Barstock states that franchising alters the nature of the investment which he will make in expanding his business. He will invest in specialised centralised services like brand promotion but not in such things as systems to monitor employees' performance. The capital advantage stems from concentrating investments in areas in which the franchisor has advantages. Mr. Field, for Austin Rover, says much the same thing. On this view capital is not brought into the franchisor's business by franchisees; rather, limits are defined for the different parts of the business.

The view that the franchisor invests in a different sort of business through franchising returns attention to the monitoring advantages of franchising. As Mrs. Darley states for Cure 30, she has been able to invest in an 'headache-free' business.

Two franchisors mention disadvantages which they have found to be attached to franchising. Mr. Dell of Pizza Express points out in his interview that he has had to develop a distinct style of management to cope with the independent status of franchisees. Mr. Peacock of Quikframe comments similarly over some aspects of the franchisee's entrepreneurship. Mr. Peacock believes that franchisees are more committed to local success than employed managers but thinks that their independence can cause problems over achieving uniformity of service within the network. Generally, franchisors do not perceive drawbacks to franchising compared with full forward

integration or selling products or services through independent retail outlets.

The benefits of franchising to franchisees

First, we should note that all franchisees, except in the case of Quikframe, state that they benefit from being able to trade under a national brand name. Most often the brand name benefit is direct as in the case of vehicle-rental, franchisees like J, who can give a clear view of the percentage of sales generated by the brand name. Sometimes the benefit is at least partly indirect as when a Wimpy franchisee obtains his lease using the franchisor's name to add to his credentials. It may well be that franchisees find it easier to raise finance because of a brand-name association. In some cases (Mobiletuning, Midas, and Computerland) the brand name is relatively weak; these are the three operational specialised-input franchises identified in Chapter 4. Table 8.3 summarises the perception of brand-name along with other benefits among the sample of franchisees.

An interesting result from the case studies is that few franchisees explicitly recognise benefits of risk reduction from trading in a franchised system. It is well known among bankers and business academics that new business failure rates are lower among franchised outlets compared with independent businesses (Stanworth, 1985). It is also important to note that risks are lowered by franchising; they are not borne by the franchisor for the franchisee. There is no evidence at all for the view of franchising which is sometimes put forward and which regards it as a risk-sharing contract between principal and agent (Ricketts, 1986, 241). My impression from interviewing is that franchisees regard themselves as less risk averse than regular employees: this is clear in the interview with Q, for example.

Table 8.3 Franchise benefits

	Brand Name	Risk Lower	Other Service
Apollo	Strong		
Garage Door	Yes		
Avis	Strong		Weak
Mobiletuning	Weak	D ₃	Strong
Young's	Yes		Weak
Wimpy	Strong		Yes
Pizza Express	Yes	G ₁	
Bally	Strong		
Nationwide	Yes		
Budget	Strong		Weak
Midas	Weak	K ₁	Yes
Olivers	Yes		Strong
Quikframe	No		Yes
Cure 30	Yes		Yes
Computerland	Weak		Yes
Yves Rocher	Yes		
Austin Rover	Yes		Yes
Ford	Strong		Yes

Note: Blank spaces indicate that no benefit is perceived by the franchisee

Apart from the brand name, franchisees tend to identify other specialised inputs from the franchisor as benefits. Thus, Avis and Budget franchisees place a relatively low value on vehicle supply lines. Mobiletuning franchisees highly value the training input.

The general rule appears to be that franchisors must offer some advantage over independent trading to franchisees. Often, franchisees note that this makes it possible for them to compete in non-price terms in their markets, as in the interview with A₂.

Franchisees more commonly note problems with franchise systems compared with franchisors. It is difficult to assess how serious these are: we would find employees criticising practices within their firms and independent retailers tend to have grievances with suppliers. C₁ finds that Avis's system for supporting vehicles on one-way rental is slow in paying costs which he incurs. H₃ found that Bally suppressed his window-dressing skills in the interests of uniformity. Both of these difficulties appear to flow from an history of management practices within company-owned networks, and are minor in nature. However, there are cases where franchisors have not delivered services which their franchisees expected: this may be said of Quikframe, Computerland and Nationwide. Quikframe has failed to develop, Computerland has changed the fee schedule, and Nationwide has settled into a pattern of low expectations on both sides.

A theory of franchising

It is now possible to state a theory of franchising based on hostage principles. Before we may find franchising there must be a tendency towards forward integration for the manufacturer of a product, owner of a brand name, or producer of some specialised service. This may arise in one of two ways: either there is an externality relationship that may be controlled by integration, or some specialised asset would lead to post-contract opportunism if open market transactions were to occur. Thus, we may draw on Williamson's externality

principle (1981, 1549) and Klein, Crawford and Alchian's idea of appropriable quasi rents (1978, 298) to establish the idea that the product, brand name or other specialised input will not be sold to just any retail outlet.

It is clear from Table 8.2 that product franchisors recognise the externality principle. However, this is not apparently so for brand franchisors. The difference between product and brand franchisors is that brand ones do not supply branded services which are embodied in a product. It is simply infeasible for them to supply the branded service through any independent retail outlet. The essence of the brand image is that there should be a set quality of service and this requires some form of obligational contract, which is a tendency towards integration. Because the externalities which would arise if just anyone could take on the brand name are so severe, we do not find brand franchisors seriously considering the option. Thus, for companies like Avis, it is a question of full integration or franchising.

Franchises which are based largely on different specialised inputs, face other problems. If training is the basis of the relationship, as in the case of Mobiletuning, a difficulty arises over appropriating the full benefits of this. Specifically, if Mobiletuning set up as a training school, individuals could buy training and then sell their acquired knowledge to others. This is stopped by integration of training and production. The externality at work here differs from Williamson's.

The other common specialised input is purchasing skill. A wholesaler may be able to create an economical bulk buying service for a small number of regular trading partners. The problem is that he must dedicate assets to providing the specialised links. This leaves him

vulnerable to post-contract opportunism, which may be stopped by integration.

We must establish a basis for integration between franchisor and franchisee, of at least a loose kind, before showing how they come to substitute the franchise relationship for full integration. This basis may be subject to variation as long as the tendency towards integration remains. The reason why this is required is simple: many of the benefits from franchising for the franchisor may be obtained by simple market transaction.

A franchisor emerges when the monitoring costs of full forward integration exceed those of franchising. The costs of monitoring the franchise system over quality standards, pursuit of business, honest reporting, and so forth are lowered when hostages may be used to screen for entrepreneurship in franchisees, and when they may be used to create ex-post bonds that penalise cheating. This implies that not all retail networks, nor all parts of networks are franchisable.

If there are economies of scale in monitoring, or if costs are reduced by simple geographical concentration then franchising may not occur. The manufacturer, brand owner or other specialised-input owner may make more profit by incurring these costs and operating the system himself. Rubin's (1978) views fit in here. The network may then run closer to some hypothetical first-best level of operation.

Also, monitoring costs may be lowered by factors other than scale or travel costs. Computerised sales recording, telephone call diverters (used by Nationwide) and video recording are all examples of technical changes which can lower the cost of monitoring a retail network. Companies may also learn by doing so that experience

enables them to monitor more effectively. It is probably a combination of scale economies and learning by doing which explains why some mature franchise chains begin to run company shops.

As we saw in Chapter 7, hostages are the value of the franchisees' sunk investments. These implicit hostages screen franchisees to ensure that they are profit-motivated entrepreneurs and they imply penalties for cheating. As a penalty mechanism, the hostage operates in the manner described by Klein (1980, 359) except that it is generally supported by professional-limitation clauses rather than by lease control. The penalty is much larger than the likely benefit to the franchisee from cheating. Opportunistic appropriation of the hostage by the franchisor is prevented by making it an ugly princess in Williamson's terms (1985, 176). Appropriation of the hostage is governed by price or arbitration clauses in the franchise agreement, or by the franchisor's reputation.

If franchisors have poor reputations or if contracts cannot be written to limit post-contract opportunism then franchising may be unduly costly relative to full forward integration. The same may be concluded if professional-limitation clauses are abandoned unless there are other screening and penalty devices open to franchisors. Other devices are open to the car makers, who do not use professional limitation. The dealer stakes his special tools, the specificity of which does not require contractual augmentation. Wimpy does not use professional limitation, but again assets are specialised all by themselves: £500,000 worth of Wimpy bar is not much use as anything else without considerable expenditure on adaptations.

The above analysis of the franchisor's use of screening and bonding aspects of hostages suggests caution

in the review of such things as professional-limitation clauses in franchise contracts. Competition authorities must be careful not to overlook efficiency properties of such clauses.

This far, we have an argument which points to monitoring costs in alternative organisations as the key factor determining organisational choice. The theory operationalises transaction-cost ideas in this area by acknowledging the screening and bonding properties of hostages. This is not quite the same thing as a theory of franchising. Specifically, we need to know that the franchisor will be motivated to perform his services efficiently and also that franchisees will be forthcoming to join the network. The first of these considerations may be satisfied by recalling from Chapter 5 that not too much of the franchisor's income should be paid at the start of the franchise contract. In the precise terms of Chapter 5, the franchisor's own sunk costs attached to a particular outlet may be paid as a lump sum to him. This ensures that it is always worth his while to perform his services provided forecasts are accurate. Clearly, forecasts were not accurate in the case of Quikframe.

The franchisor will recruit franchisees if they are better off in rather than out of the network. This means that they must perceive benefits like the ones listed in Table 8.3. Rubin's view (1978, 227) that joining the franchise network is an investment decision for the franchisee may be used at this point. However, we should look at this in slightly sophisticated terms. The franchisee joins if he is satisfied that he makes at least a normal profit and believes that he is safe from post-contract opportunism.

Summary and conclusions

Franchise systems analysed in this study generate monitoring-cost savings for franchisors, relative to full forward integration, which derive from the entrepreneurship and greater trustworthiness of franchisees relative to hired managers. Franchisees benefit from brand-name and other specialised-input advantages from franchising.

These observations may be developed into a theory of franchising which sees the organisational form in investment-decision terms but where contracting parties protect themselves from opportunism. Franchisors will aim to minimise their transaction costs. These are seen as resting on the screening and bonding properties of the franchisee's sunk investment, which acts as an implicit hostage.

Apparently 'unfair' contractual terms such as professional-limitation clauses, have efficiency aspects according to the analysis of this chapter. In some franchises, initial costs may be insufficiently sunk for the franchisee to provide a useful hostage unless the franchisor can control the franchisee's use of his assets.

Conclusions

This final section draws together conclusions which have already been reached, particularly in Chapters 5, 7 and 8, and adds some new insights. The main aim is to give an overview of the principal results of the study.

The fieldwork carried out on a UK franchising system shows that two areas of transaction costs emerge as important in franchise contracts. The first of these concerns the measurement of future benefits from the franchise relationship. The second concerns the monitoring costs which are faced by the franchisor. Both areas are affected by uncertainty, bounded rationality and a potential for opportunistic behaviour. Key aspects of the franchise contract are determined by considerations in these areas.

The analysis of preceding chapters does not use arguments based on the relative risk-preferences of franchisees and franchisors. This contrasts with some approaches to franchising that regard the contract as transferring risk from risk-averse franchisees to risk-neutral franchisors (Ricketts, 1986, 241). Interviews with franchisors and franchisees do not support this view of the franchise relationship. Franchisees often pride themselves over their risk taking. Viewing them as highly risk averse probably stems from regarding franchising in employment-contract terms. Whilst this study has an implied assumption of risk neutrality running through it, we do not need concepts of risk taste to explain the key features of franchising. The approach here views this matter rather as do Mathewson and Winter (1985). We follow Williamson's (1985, 388) advice to eschew arguments based on risk aversion as these tend to observe 'core efficiency purposes' of organisational forms.

The 19 studies summarised in Volume 2 of this thesis enable us to identify product, brand and specialised-input franchises. Of these the brand kind is most common. Each type makes a slightly varied use of a basic fee schedule consisting of a lump-sum initial franchise fee and a continuing royalty payment. Within a franchise, these payments are uniform and do not vary with site characteristics. Lump sums, which may be stock charges for product franchisors, are tied to the franchisor's sunk costs attached to a franchisee. Continuing payments, which are product mark-ups for product franchisors and fixed percentage royalties for brand and specialised-input franchisors, link the franchisor's rewards to continued performance of his services. The costed nature of the lump sum is explained by the franchisee's difficulties in predicting site performance coupled with the franchisor's need to protect himself from post-contract opportunism. Uniformity is explained by the franchisor's difficulties in predicting the franchisee's performance. The transaction costs here concern measurement for the main part.

Observed fee-schedule patterns can be demonstrated in a development of a model of Mathewson and Winter (1985). In this, we take a different view of the constraint required to prevent the franchisor from absconding with the lump sum. Also, franchisors and franchisees adopt normal-profit expectations of the performance of a particular site. An implication of the model is that franchisors may develop means of selectively assisting weak franchisees or removing economic rent. There is case-study evidence for selective intervention, with some bias towards the support of weak sites. In general, fee schedules economise on bounded rationality by simplifying problems for decision makers.

Fee schedules are only one area of the franchise relationship. Franchisors and franchisees use written agreements which are augmented by implicit contracts to specify at least the groundrules of their relationship. All franchises show implicit aspects of the contract covering such things as advertising levels, the price of transferred products, or the repurchase of a franchised outlet by the franchisor if projected sales do not materialise. The written agreement tends to be incomplete in areas like these, where flexibility is required as market conditions change, and where setting rules might leave the franchisor open to opportunism.

With the wider franchise contract, there is evidence that the written agreement and the implicit aspects of contract are revised over time without reference to outside bodies such as courts or arbitrators. The whole contract may be described as relational (Macneil, 1974 ; Goldberg, 1980; and Williamson, 1985). The case studies show that adjustments are subject to bilateral governance; the franchise relationship has a life of its own, so that its evolving value to the parties leads them to align their incentives over time. The transaction costs here mainly concern governance rather than measurement. There are few cases of litigation between franchisor and franchisee. The relational contract avoids the costs of such third-party intervention, especially where these refer to the disruption of business.

A specific governance issue faced by the franchisor arises over the monitoring of franchisee performance. A symmetrical problem for the franchisee is solved by ensuring that the franchisor always benefits from consummate performance of his services through the design of the fee schedule. Part of the franchisor's services is to monitor the franchise network to control horizontal and vertical externalities and to ensure that profitable

opportunities are pursued by the franchisee. An extra complication is that franchisees sometimes contribute specialised skills, like used-car trading, which franchisors may find difficult to supervise.

Monitoring costs are economised by the posting of implicit hostages by the franchisee. These take two principal forms. First, the lump-sum paid to the franchisor is a sunk investment from the franchisee's viewpoint; if his business were to fail then the sum would not be recoverable. The franchisee will never worry that the franchisor may abscond with the lump sum according to our analysis of fee schedules, as long as the sum is no greater than the sunk costs faced by the franchisor in connection with the particular franchised outlet.

In addition to the lump sum there are usually other investments made by franchisees at the start of their business which would be unrecoverable in the event of business failure. The trademarked equipment and leasehold improvements which the franchisee normally purchases are the main examples of these. In relation to straightforward business failure, the franchisee is much like any other small businessman except that the trademarking of his premises and equipment may make them less salvageable.

However, franchisors generally use post-termination, professional-limitation clauses in their agreements which intensify the sunk nature of the franchisee's initial investment. These prevent the franchisee from competing in his old franchise area for a period of time following the termination of the franchise relationship for any reason. Thus, the hostage of sunk investment is not only given up in the event of simple business failure but also if the franchisor terminates the agreement.

Franchisees are normally protected from opportunistically contrived cancellation of the contract by the franchisor through the use of asset-valuation procedures which are invoked by the agreement if the franchisor terminates. Thus, the franchisor is not able to obtain cheap assets or increases in the franchise fee because of professional limitation. The franchise contracts avoid creating appropriable quasi rents for the franchisor. There may well be a clue here to explain why written agreements are used at all, given the relational nature of the franchise contract. Written clauses can set unambiguous constraints, in an area like the post-termination valuation of assets which will reassure the franchisee. The benefits from creating precision then outweigh its costs (Macaulay, 1963). Given that hostages are of no direct value to the franchisor, they have the characteristics of ugly princesses (Williamson, 1985).

Hostages have ex-post bonding properties in the governance of the franchise relationship. The possibility of incurring a penalty implies that the franchisee will be less likely to cheat in some way. The expected benefit from cheating equals the probability of success multiplied by returns minus the probability of detection multiplied by the penalty. By increasing the penalty, the franchisor can reduce his direct monitoring of the franchisee rather as in models of crime and punishment (Becker, 1968) and as our data shows him to do.

Klein (1980) is correct in arguing that franchise contracts contain devices which intensify sunk costs for the franchisee. However, his chosen example of lease control is not relevant here. This is because professional limitation is a sufficient device on its own. Franchisors with high-street locations simply minimise their costs of terminating the franchise agreement or of losing the franchisee by controlling leases.

Predictions may be made of when franchising will be used instead of full forward integration or sales through independent retailers. First, a tendency towards integration must be present. Some organisational link between a manufacturer of a product, owner of a brand name or supplier of a specialised input and his retail outlets must give advantages in controlling externalities or opportunism within a retail network. A sufficient condition for franchising is then that there should be monitoring-cost savings relative to full integration. These may derive from the penalty versus monitoring trade off described above or from ex-ante screening properties of hostages.

Apart from controlling quality standards, franchisors also have a basic monitoring problem in that they need to supervise business decisions throughout the network. Cost saving and demand seeking requires encouragement, as does profit-orientated decision making. The fact that a franchisee has his own capital tied up in his business makes him more likely to be alert to profitable business opportunities. In case after case, franchisors state that they benefit mainly from the commitment of their franchisees through franchising. By this they appear to mean that their branches are more likely to be run with entrepreneurial alertness. It is as though a dilution of management control which may be expected to occur as a network grows (Williamson, 1967; and Calvo and Wellisz, 1978) is interrupted. The hierarchy starts all over again with the franchisee and this reduces the costs of monitoring decisions.

Thus franchising will be adopted whenever monitoring costs are reduced through the hostage posting which is analysed above, assuming that franchisees may be drawn into the system. Franchisees must be offered at least their next-best earnings by the franchise relationship.

This view of the basis of franchising moves us away from the traditional view which sees franchising in employment-contract terms (Caves and Murphy, 1976; Rubin, 1978; and Mathewson and Winter, 1985). In the traditional view, profit sharing is the driving feature of the franchise relationship. In this study, the relationship is driven by the bonding and screening properties of implicit hostages. Profit sharing is more a reflection of measurement difficulties at the start of the franchise (and some employment) contracts. This study increases our understanding of franchising by viewing it in terms of the theory of the firm.

As a final conclusion we may state that franchising is essentially a contractual relationship between specialised businesses which has monitoring-cost advantages flowing from the use of hostages. The relationship responds flexibly over time to changed cost, demand and information conditions. It appears to obey Williamson's dictum to economise on bounded rationality whilst safeguarding the relationship against opportunism.

REFERENCES

- Alchian, AA and Demsetz, H. (1972)
Production, information costs, and economic organisation,
American Economic Review, 62, 777.
- Alchian, AA (1984)
Specificity, specialization, and coalitions, Journal of
Economic Theory and Institutions, 140, 34-49.
- Akerlof, GA (1970)
The market for 'lemons': qualitative uncertainty and the
market mechanism, Quarterly Journal of Economics, 84, 488-
500.
- Ansoff, HI (1965)
Corporate Strategy, London, Penguin.
- Arrow, KJ (1962)
Economic welfare and the allocation of resources for
invention, in, National Bureau of Economic Research, The
Rate and Direction of Inventive Activity, 609-25,
Princeton.
- Arrow, KJ (1969)
The organization of economic activity: issues pertinent to
the choice of market versus nonmarket allocation, in The
Analysis and Evaluation of Public Expenditure: The PPB
System. Vol. 1 U.S. Joint Economic Committee, 91st
Congress, 1st session Washington D.C., U.S. Government
Printing Office, 59-73.
- Arrow, KJ (1985)
The economics of agency, in Pratt, JW and Zeckhauser, RJ
(eds.), Principals and Agents: The Structure of Business,
Boston, Harvard Business School Press.
- Barnard, C (1938)
The Functions of the Executive, Cambridge, Harvard
University Press
- Barzel, Y (1982)
Measurement cost and organization of markets, Journal of
Law and Economics, 25, 27-48.
- Baumol, WJ (1977)
Economic Theory and Operations Analysis, 4th Edition,
London, Prentice-Hall.
- Baumol, WJ (1986)
Williamson's The Economic Institutions of Capitalism, Rand
Journal of Economics, 17, 279-86.

- Baumol, WJ, Panzer, J and Willig, R (1982)
Contestable Markets, New York, Harcourt Brace Jovanovich.
- Becker, G (1968)
Crime and punishment: an economic approach, Journal of Political Economy, 76, 169-217.
- Buchanan, JM (1975)
A contractarian paradigm for applying economic theory, American Economic Review, 65, 225-30..
- Burgess, RG (1982)
Field Research: A Sourcebook and Field Manual, London, Allen and Unwin.
- Casson, M (1982)
The Entrepreneur: An Economic Theory, Oxford, Martin Robertson.
- Caves, RE and Murphy, WF (1976)
Franchising: firms, markets and intangible assets, Southern Economic Journal, 42, 572-86.
- Chandler, AD (1962)
Strategy and Structure, Cambridge, Massachusetts, MIT Press.
- Cheung, S (1983)
The contractual nature of the firm, Journal of Law and Economics, 26, 1-22.
- Coase, RH (1937)
The nature of the firm, Economica N.S., 4, 386-405.
- Coase, RH (1960)
The problem of social cost, Journal of Law and Economics, 3, 1-44.
- Coase, RH (1972)
Industrial organisation: a proposal for research, in VR Fuchs (ed.) Policy Issues and Research Opportunities in Industrial Organisation, New York, National Bureau of Economic Research, 59-73.
- Cohen, LR (1979)
The firm: a revised definition, Southern Economic Journal, 46, 580-590.
- Commons, JR (1934)
Institutional Economics, Madison, University of Wisconsin Press.
- Davis, OA and Whinston, AB (1962)
Externalities, Welfare and the theory of games, Journal of Political Economy, 70, 241-62.

Davis, OA and Whinston, AB (1966)
On externalities, information and the government-assisted invisible hand, *Economica*, 33, 303-18.

De Alessi, L (1980)
The economics of property rights: a review of the evidence, *Research in Law and Economics*, 2, 1-26.

Demsetz, H (1967)
Toward a theory of property rights, *American Economic Review*, 57, 347-59.

Demsetz, H (1969)
Information and efficiency: another viewpoint, *Journal of Law and Economics*, 12, 1-22.

Demsetz, H (1982)
Barriers to entry, *American Economic Review*, 72.

Dobb, M (1925)
Capitalist Enterprise and Social Progress, London, George Routledge and Sons Ltd.

Dobb, M (1928)
Russian Economic Development, New York, F.P. Dutton and Co.

Fama, EF (1980)
Agency problems and the theory of the firm, *Journal of Political Economy*, 88, 288-307.

Frey, JH (1983)
Survey Research by Telephone, Beverly Hills, Sage.

Glaser, BG and Strauss, AL (1967)
The Discovery of Grounded Theory: Strategies for Qualitative Research, New York, Aldine.

Goldberg, V (1980)
Relational exchange: economics and complex contracts, *American Behavioral Studies*, 23, 337-52.

Gravelle, H and Rees, R (1981)
Microeconomics, London, Longman.

Hayek FA (1945)
The use of knowledge in society, *American Economic Review*, 35, 519-30.

Holmstrom, B (1981)
Contractual models of the labour market, *American Economic Review*, 71, 308-13.

Jack, AB (1957)

The channels of distribution for an innovation: the sewing machine industry in America, 1860-1865, *Explorations in Entrepreneurial History*, 8, 113-41.

Jensen, M and Meckling, W (1976)

Theory of the firm: managerial behaviour, agency costs and capital structure, *Journal of Financial Economics*, 3, 305-60.

Kay, NM (1983)

Optimal size of firm as a problem in transaction costs and property rights, *Journal of Economic Studies*, 10, 29-41.

Kirk, J and Miller, ML (1986)

Reliability and Validity in Qualitative Research, Beverly Hills, Sage.

Kirzner, I (1973)

Competition and Entrepreneurship, Chicago, University of Chicago Press.

Klein, B (1980)

Transaction cost determinants of 'unfair' contractual arrangements, *American Economic Review*, 70, 356-62.

Klein, B (1983)

Contracting costs and residual claims: the separation of ownership and control, *Journal of Law and Economics*, 26, 367-74.

Klein, B, Crawford, RG and Alchian, AA (1978)

Vertical integration, appropriable rents and competitive contracting process, *Journal of Law and Economics*, 21, 297-326.

Klein, B and Leffler, KB (1981)

The role of market forces in assuring contractual performance, *Journal of Political Economy*, 89, 615-41.

Knight, FH (1921)

Risk, Uncertainty and Profit, Boston, Houghton Mifflin

Koo, AYC (1959)

A theoretical note on the dealer-manufacturer relationship in the automobile industry, *Quarterly Journal of Economics*, 73, 316-25.

Lakatos, I (1970)

Falsification and the methodology of scientific research programmes, in, Lakatos, I and Musgrave, A (eds.), *Criticism and the Growth of knowledge*, 91-195, Cambridge, Cambridge University Press.

Llewellyn, KN (1931)

What price contract? An essay in perspective, Yale Law Journal, 40, 704-51.

Lofland, J. (1971)

Analyzing Social Settings: A guide to Qualitative Observation and Analysis, Belmont, Wadsworth.

Macaulay, S (1963)

Non-contractual relations in business, American Sociological Review, 28, 55-70.

Macneil, IR (1974)

The many futures of contracts, Southern California Law Review, 47, 691-816.

Macneil, IR (1978)

Contracts: adjustments of long-term economic relations under classical, neoclassical and relational contract law, Northwestern University Law Review, 72, 854-906.

Marshall, A (1925)

Principles of Economics, 8th edition, New York, Macmillan.

Mathewson, GF and Winter, RA (1985)

The economics of franchise contracts, Journal of Law and Economics, 28, 503-26.

Mckee, M and West, EG (1981)

The theory of second best: a solution in search of a problem, Economic Enquiry, 19, 436-48.

McManus, JC (1975)

The costs of alternative economic organisations, Canadian Journal of Economics, 8, 334-50.

Mendelsohn, M (1985)

The Guide to Franchising, London, Pergamon.

Miles, MB and Huberman, AM (1984)

Qualitative Data Analysis, London, Sage.

Monteverde, K and Teece, D (1982)

Supplier switching costs and vertigal integration in the automobile industry, Journal of Economics, 13, 206-13.

Myerson, R (1979)

Incentive compatibility and the bargaining problem, Econometrica, 47, 61-73.

Ozanne, UB and Hunt, SD (1971)

The Economic Effects of Franchising, Washington, U.S. Government.

Porter, G and Livesay, HC (1971)
 Merchants and Manufacturers: Studies in the Changing
 Structure of Nineteenth Century Marketing, Baltimore, John
 Hopkins Press

Plott, CE (1966)
 Externalities and corrective taxes, *Economica*, NS, 33, 84-
 87.

Rees, R (1985)
 The theory of principal and agent, *Bulletin of Economic
 Research*, 37, 3-26 and 75-95.

Reid, GC (1986)
 Methodology and empirical issues in the application of
 field research techniques to the business enterprise,
 University of Edinburgh Department of Economics Discussion
 Paper.

Reid, GC (1987)
 Theories of Industrial Organisation, Oxford, Basil
 Blackwell

Richardson, GB (1972)
 The organization of industry, *Economic Journal*, 82, 883-
 96.

Ricketts, M (1986)
 The geometry of principal and agent: yet another use for
 the Edgeworth box, *Scottish Journal of Political Economy*,
 33, 228-249.

Riordan, MH and Williamson, OE (1985)
 Asset specificity and economic organization, *International
 Journal of Industrial Organization*, 3, 365-78.

Robbins L (1932)
 An essay in the Nature and Significance of Economic
 Science, London, Macmillan.

Robertson, DH (1930)
 Control of Industry, London, Nisbet.

Robinson, EAG (1934)
 The problem of management and the size of firms, *Economic
 Journal*, 44, 240-54.

Ross, SA (1973)
 The economic theory of agency: the principal's problem,
American Economic Review, 62, 134-9.

Rubin, PH (1978)
 The theory of the firm and the structure of the franchise
 contract, *Journal of Law and Economics*, 21, 223-233.

Scitovsky, T (1943)

A note on profit maximisation and its implications, Review of Economic Studies, 11, 57-60.

Shavell, S (1979)

Risk sharing and incentives in the principal and agent relationship, Bell Journal of Economics, 10, 55-73.

Simon, HA (1957)

Models of Man, New York, John Wiley and Sons.

Simon, HA (1961)

Administrative Behavior, New York, Macmillan. Original publication 1947.

Sloan, AP (1986)

My Years With General Motors, Harmondsworth, Penguin. First published 1963.

Smith, A (1776) (1925)

The Wealth of Nations, 4th edition, Edwin Cannan (ed.), Methuen and Co., London.

Stanworth, J (1985)

The Franchise Relationship, London, Franchise World.

Stekler, H. (1960)

Dealer-manufacturer relationships in the automobile industry: a comment, Quarterly Journal of Economics, 74, 330-3.

Teece, DJ (1976)

Vertical Integration and Divestiture in the U.S. Oil Industry, Stanford, California, The Stanford Institute for Energy Studies.

Vaughan, CL (1974)

Franchising, Lexington, DC Heath.

Von Neumann, J and Morganstern, O (1944)

Economics and the Theory of Games, London, Macmillan.

Williamson, OE (1967)

Hierarchical control and optimum firm size, Journal of Political Economy, 75, 123-38.

Williamson, OE (1971)

The vertical integration of production: market failure considerations, American Economic Review, 61, 112-23.

Williamson, OE (1975)

Markets and Hierarchies: Analysis and Antitrust Implications, New York, Free Press.

Williamson, OE (1979)

Transaction-cost economics: the governance of contractual relations, *Journal of Law and Economics*, 22, 3-61.

Williamson, OE (1981)

The modern corporation: origins, evolution, attributes, *Journal of Economic Literature*, 19, 1537-68.

Williamson, OE (1983)

Credible commitments: using hostages to support exchange, *American Economic Review*, 73, 519-40.

Williamson, OE (1985)

The Economic Institutions of Capitalism, New York, Free Press.

*

Williamson, OE Wachter, ML and Harris, JE (1975)

Understanding the employment relation: the analysis of idiosyncratic exchange, *Bell Journal of Economics*, 6, 250-80.

* WILLIAMSON, OE (1986)

ECONOMIC

ORGANISATION,

BRIGHTON, WHEATSHEAF.