Vertical Contracts in Petrol Retailing: Incentives and Risk-Sharing

Cristina Caffarra

Oxford Institute for Energy Studies

WPM 19

September 1994

VERTICAL CONTRACTS IN PETROL RETAILING

INCENTIVES AND RISK-SHARING

Cristina Caffarra

WPM19
Oxford Institute for Energy Studies
September 1994

The contents of this paper
are the author's sole responsibility.
They do not necessarily represent the
views of the Oxford Institute for
Energy Studies or any of its Members.

ISBN 0 948061 84 7

Cristina Caffarra joined the Oxford Institute for Energy Studies (O.I.E.S.) as a Researcher in 1988, working mainly on issues concerning the upstream oil industry. This paper is a revised version of a Chapter from her D.Phil. thesis at Oxford University ("Strategic vs. Agency Considerations in Vertical Contracting: The Case of Petrol Retailing"), undertaken with support from the O.I.E.S., the ENI-Enrico Mattei Foundation (Italy), and the Institute of Petroleum (London).

CONTENTS

1 INTRODUCTION

The purpose of this paper is to describe the variety of contractual arrangements in use in the UK petrol industry between refiners/wholesalers and their retailers, and assess their respective properties. Section 2 provides a brief account of the evolution of the main contractual options over the past two decades, emphasizing both the process of dynamic change which characterized the choice of organizational form, and the competitive element in such choice (i.e. how changes by one company appear to set in motion a process of adjustment by other companies). The profile of each contract type is outlined in Section 3; the main differences across types are then analysed in an agency framework, with emphasis on the risk-sharing and the incentive properties – explicit and implicit – of each format (Sections 4 and 5).

On the basis of these structural differences, contracts may be separated out in a menu of discrete choices. The variety of contract types raises the issue of the assignment rule(s) being followed, i.e. what criteria are being adopted for matching station (and retailer) with contract type; and in particular, how the wholesaler's choice of contract type may be related to the fundamental characteristics of the local market, the asset and the retailer.

2 CONTRACT VARIETY: EVOLUTION OF CONTRACTUAL AGREEMENTS IN THE UK PETROL INDUSTRY

Of approximately 17,000 filling stations in the UK at end-1993, almost two-thirds were owned and operated by independent dealers. Their share of the total has been declining over time (90 per cent of all stations were independent in the early 1960s); in addition, their relative weight is far less substantial when fuel volumes (rather than site numbers) are considered: excluding the expanding network of the major greengrocers/supermarkets (which is estimated to account for around 15 per cent of total UK fuel sales, with only 650 sites), the independent dealers' share of petrol sales is only around 35 per cent. In contrast, against the background of a shrinking total population of stations (a phenomenon which is common to other mature petrol markets), wholesaler-owned sites have grown both in relative numbers and volumes.[1] By implication, while the throughput at UK filling stations has increased on average (from around half a million litres per annum (mla) in 1970, to over 1.7 mla today), the increase has been much more significant at wholesaler-owned sites (the average throughput is less than 1 mla for independents, but in excess of 2.5 mla for wholesaler-owned sites).

In terms of organizational structure, six main categories of agreements exist in the UK petrol industry between petrol companies and retailers, specifying the terms for the retail sale of petrol and the day-to-day operation of the filling station. In addition to exclusive supply agreements with *independent dealers*, who own their site and sell the wholesaler's branded petrol (normally also carrying out a number of ancillary activities), four different contractual agreements exist at present for the management of wholesaler-owned outlets. These are *tenancies, licences, commission agencies*, and *direct management* by the wholesaler's own personnel. *Franchising* has been recently added to the list, applying both to independent and company-owned sites.

This menu of contractual options currently available to petrol companies is the result of a continuing evolution. A feature of this process is the competitive nature of many of the changes recorded across companies: the introduction of a new contract form by any one company always tends to spur imitators, while 'lateral' changes across the menu (i.e. 'switching' a group of retailers into a different existing contract) also appear to set in motion responses and adjustments. The period 1991–3, for example, was a rather turbulent phase, with a new contract being introduced (the franchise) and several adjustments being made by companies to their contract-mix.

Taking a broader view, within the category of contracts at wholesaler-owned sites, tenancies were until the late 1970s the dominant form of agreement. Around that time the licence was introduced, and a number of petrol companies switched to this contract; among the reasons offered to the Monopolies and Mergers Commission (MMC) in its latest investigation of the industry (*The Supply of Petrol*, Cm 972, 1990), was the view that when it came to 'improving facilities and standards', 'some tenants were obstructive . . . [They] would seek lower costs by operating for shorter hours and having less staff'

[1] There are currently around sixty petrol wholesalers in the UK, of which only a limited number (twelve) are also domestic refiners. However these are also the largest wholesalers, with over 70 per cent of all UK sites displaying their brand name.

(pp. 368-9). And 'the benefits provided to a tenant under the terms of a lease, often of 21 years' duration, largely prevented the company from implementing its strategies in the short and medium term, and the right of renewal granted under the Landlord and Tenant Act could effectively deny the company the opportunity to realize the potential of its property in the longer term' (p. 371). The attraction of the licence was thus in the greater degree of control it allowed the wholesaler on the site, an important aspect of which was that it made it legally easier to remove unsatisfactory retailers.[2]

The introduction of the licence was in fact just the first stage of a more general search by UK petrol wholesalers for greater influence on the operation of their outlets. From the mid-1970s through the 1980s, there was a significant shift in the form of agreement favoured at company-owned sites, with tenancies being gradually replaced by contractual arrangements giving the wholesaler a tighter control over the asset: licences, commission agencies and direct management by salaried employees. The wholesalers presented these changes as being very much in the interest of the consumer and the retailer, emphasizing to the MMC that 'the consumer sought better facilities than were provided by the small forecourt, workshops and kiosk operations of the 1970s', and there was therefore 'a need to keep pace with the market and with technology' (p.368); in addition, they stressed the 'need to spell out to the retailer, in the interests of clarity, more about his (and the wholesaler's) obligations, given the requirement for better standards' (p.36). Of course, a key reason for the changes must be identified in the growing rent being generated by all the activities carried out at the site, and the companies' wish to improve the extraction of such rent – at a time when tenants resisted rent reviews and other changes, and the number of cases referred to arbitration or to the courts was increasing. Hence the switch by many wholesalers to commission agreements – where the retailer's duties are spelled out thoroughly, and the security of tenure is nil (i.e. the retailer has no right to renewal of the contract, and can be removed at very short notice); or to licences, which also carry no right to renew, are non-assignable and only recognize the retailer a negligible compensation on site vacation (i.e. no recovery of goodwill). In addition, the MMC acknowledges that while the early licences were less detailed, they have become 'more extensive and comprehensive over time' (p. 36), a process which often disguises a tightening of terms: schedules of agreement now specify terms with great precision, on matters ranging from the pricing of petrol to standards of operation, merchandising and choice of supplier for the forecourt shop.

The trend towards tighter wholesaler's control over all aspects of the station's operations has been strengthened by the recent introduction of the franchise format in the menu of available contracts. 'Business-format' franchise generally carries a substantial burden of rules, standards and business practices carefully laid down by the franchisor, implying a comprehensive degree of control exercised by the latter. While separate franchise agreements for the operation of the forecourt shop were already often used in petrol retailing, the franchise formula was introduced for the first time to fuel retailing by Shell, in 1991. This has been followed, in some measure, by Texaco (which switched a portion of its tenanted sites to franchises in 1992), and BP (which in 1993 introduced a 'franchise-type' agreement to replace its licence).

---

[2] However, some wholesalers were able to step around this undesirable feature of the tenancy by cutting out of their agreements the tenant's protection offered by the Landlord and Tenant Act (Sect.24-28). See Section 3.

Petrol wholesalers appear thus involved in a process of continuous adjustment of the contract types/terms they offer, frequently switching from one type to another, or slightly modifying terms within a chosen type. Among recent examples, Jet gradually converted its commission-operated sites to licences (1991–2), while Fina turned the bulk of its directly managed sites into commission agencies (1992). To some extent, many of the recent changes (since 1990) appear to have been inspired by a re-assessment of the importance of the retailer's direct effort, and an attempt to (cautiously) redress the balance – while retaining control of the overall commercial strategy. Renewed emphasis has been put on the importance of the retailer's effort and 'entrepreneurial skills' for the operation of the filling station, reflected in some companies' switch to contracts with higher marginal rewards for the sale of fuels. Significantly, the new franchise contract has a five- instead of three-year duration (renewable), and includes at least partial recognition of the retailer's goodwill. While companies remain intent on setting rigorous standards for their networks, and thus defend their large investments in advertising and promotion from hold-ups on the part of the retailers, some companies have opted for contracts of longer duration, and higher-powered economic incentives for retailer's effort. The emphasis is explicitly put on the potential for these contracts to attract more 'entrepreneurial' retailers, suggesting an implicit screening function is also performed.[3]

Against the background of these broad (cyclical) trends, a key stylized fact is that at any one point in time, all petrol companies tend to operate their retail network under a *mixture* of different contract types. Though evolving over time, the organizational structure of the retail function always preserves some variety (The table in the Appendix presents a 'snapshot' of the incidence of various contract forms for a number of major UK brands). The persistence of variety, and the adjustments continually being made by individual companies to the menu of contracts and/or to their terms, strongly suggests that delegation of the retail function, and its specific format (contract type/terms), is an important decision variable for the refiner/wholesaler.

[3] Consider some of the comments which have accompanied some recent contracts changes (from the trade press): 'the new franchise is aimed at ambitious self-starters with proven business skills' (Texaco); 'the new (licence) agreement is intended to replace the commission agent arrangement...Operating on a commission basis was restricting our ability to expand...It was clear that if our retailers were more interested in the business and if their earnings were open-ended, then that would be better for both' (Jet); 'the local entrepreneur knows the local market better...we will lose some control over our sites, but if the operator is committed, it will be to our advantage' (Fina); 'one of the most serious grievances of all licensees is that they do not have enough security of tenure, and this reduces their incentive to invest personally and to expend effort in the site' (BP).

3 STRUCTURAL DIFFERENCES ACROSS CONTRACT TYPES IN UK PETROL RETAILING

This Section provides a description of the contract types in use between UK petrol wholesalers and retailers. Information on contract structure was gathered in a series of meetings and discussions with the Petrol Retailers Association (which allowed us to survey some of its members' actual contracts), and a number of individuals in the marketing/development departments of some major petrol companies (Shell, BP, Texaco, Q8, Jet, and Burmah). They provided valuable insight into the organization of the industry.

An important feature emerging from our contract survey is that while companies adopting the same contract type might choose different terms, there is evidence that the variation of terms across companies (within each type) is less significant than the difference across types. In other words, each contract form can be fairly generalized across companies. The process of imitation is helped by the detailed information generally available to both retailers and competitors on the specific terms offered by a company: new contract terms are fully reported in the trade press and in company advertisements; and the competition across wholesalers for attracting independent dealers ensures that the terms of the various offers are well publicized. The main structural differences across types were found to fall broadly into two categories: on the one hand, differences in the retailer's legal status, his security of tenure in the contract, and in the initial investment required; on the other hand, differences in the structure of the retailer's compensation. As argued in Section 5, these two categories identify – respectively – the implicit and explicit incentive provisions incorporated in the contract.

Legal Status, Duration, and Initial Capital Requirement (Bond)

Independent dealers (IDs) own the station they operate, and sell branded petrol on the basis of an exclusive supply agreement with a refiner/wholesaler. The agreement gives the wholesaler the right to be the sole supplier of fuel to the station, in return for an undertaking to supply at a specified wholesale price (see next section). The exclusive nature of the agreements with IDs (and the general principle of 'solus ties' in selling petrol) was fully accepted by the MMC investigation of 1965, and more recently by the EC.[4] The 1990 MMC Report did not make any objection: 'we received no representations against the principle of the solus tie, . . . [and] we saw no reason to dissent from the earlier assessment that they are not in principle against the public interest' (p. 291). As for the length of such agreements (normally five years), the MMC considered arguments for both shorter and longer maximum periods than five years. One wholesaler maintained that 'a longer-term agreement gives the supplier the increased security that may be needed if he is to underwrite a major investment on a site.' On the other hand, the MMC was aware that '[a] shorter maximum term would increase the proportion of sites coming up for renewal each year . . . and therefore provide greater scope for competition between wholesalers, though at an extra cost to them' (p. 292). At the same time, the MMC also noted that not all the 12,000 solus dealers in the UK have agreements for the full five years, and a

[4] Regulation 1984/83, providing a block-exemption for petrol retailing from the application of Article 85 of the Treaty of Rome.

significant number of dealers prefer a shorter tie; indeed, some of the larger retailers prefer to negotiate single year contracts, while making their own arrangements for financing site developments. The MMC decided that there were 'no compelling arguments for changes in either direction' (p. 292).

A *tenancy* agreement (filling station lease), is the granting of exclusive possession of premises owned by the wholesaler, incorporating a fuel-supply agreement in addition to conventional tenancy arrangements (term, initial rent and regular review provisions, equipment, premises to be leased, payment of rental, repairs, etc.) The purchase and sale of the landlord's petrol is a condition of the tenancy, as in all other types of contracts for the operation of company-owned sites. In terms of security of tenure, there is an important distinction between tenants who have the protection of the Landlord and Tenant Act 1954 (LTA, Sect. 24-28), and those who do not. These provisions give tenants a statutory right to renewal of their tenancy upon the expiration of their contract (unless they are in substantial breach of obligations, or the landlord requires the premises for redevelopment). Where LTA-protection is not granted (about one-third of all tenancies, according to the MMC), the important aspect of tenure security is lost, and from this point of view the contract has much in common with a licence.[5] When this protection is ensured, however, the tenancy has quite distinct incentive properties, as emphasized by those who use it: though the contract's relative importance has been somewhat declining in recent years, companies using tenancies told the MMC they 'believed this was appropriate for their portfolio of sites . . . A tenant's greater security of tenure attracts a more resourceful businessman whose stake in the business encouraged him to build it up, to the corresponding benefit of the wholesaler. This is particularly the case when making a success of other activities (such as a car workshop or used car sales) on the site was necessary for the site as a whole to be profitable' (p. 37). Thus wholesalers appear to favour tenancies for medium-small sized sites, where ancillary activities (e.g. car repair/sales) are an important feature of the retailer's job.

A *licence* is legally the granting of permission to carry out a business activity (sale of petrol) on premises owned by the wholesaler. Like the tenancy, the licence permits the retailer to occupy premises for the purpose of retailing petrol; the crucial difference is that the licence does not incorporate any of the renewal rights conferred by LTA. In fact, the normal duration of a licence is three years, with no guarantee that the contract will be renewed, and with provisions for early termination. This is a very controversial point in the relationship between retailers and suppliers, as retailers have argued that short duration and low security of tenure dampen their incentives. In addition, licensees have maintained that the contract gives the wholesaler too much control over the operation of the site: not only does it contain full details on notice, termination, margins, fees, payment terms, division of responsibility for site costs between the parties, and ownership and risk related to the fuel; but also, the licence specifies a series of detailed rules on how the wholesaler wants retailing to be carried out, spelling in great detail the obligations of the licensee in the conduct of business (wearing the wholesaler's uniform, opening hours, operating standards, etc.). Also, most agreements have clauses allowing the wholesaler to change operating standards and opening hours unilaterally and

[5] In this case, 'although there is a legal distinction between a tenancy . . . and a licence, in commercial terms they are similar in the rights of occupation and security of tenure thereby conferred' (MMC Report, p. 35).

at short notice, with failure to comply considered breach of the agreement and immediate ground for termination.

A further frequent complaint on the part of licensees is that the contract implies no recognition of goodwill, though the initial capital requirement may be quite significant. Licensees, like tenants, take ownership of the petrol they sell once it is delivered to them, usually on terms that require them to finance at least one load; in addition, they must buy the shop stock. This initial working-capital requirement may therefore be quite substantial;[6] against this, because he has no formal stake in the business, when the contract expires the licensee does not benefit from any increase his activity might have generated in the value of the station. It is partly to answer this criticism that some wholesalers have introduced franchising in the contract menu.

Commission agents have a fundamentally different status from licensees or tenants. In commission agency the retailer is not acting as a principal, but sells petrol which belongs to the wholesaler on his behalf, receiving a commission per unit of petrol sold. The ownership of the petrol does not pass to him at any point, but it remains with the wholesaler who also sets the pump price. This implies the retailer does not carry risk of changes in the wet-stock worth; and also, the initial capital requirement is considerably lower. Commission agency agreements provide for the retailer to operate the petrol station as the wholesaler requires; typically, they can be terminated at very short notice (thirty days), and are concerned with the agent's duties, the division of responsibility for site costs, payment of commission and termination. The agent is self-employed, and sometimes operates the site shop under a separate franchise agreement or lease, providing his own shop stock.

The *franchise* is the most important development to have been introduced in the menu of petrol retailing contracts in the last few years – though the licence still remains the most common mode of operation for wholesaler-owned sites. The wide debate this has fuelled among companies and in the trade press, on the relative merits of different contracts, indirectly confirmed the importance of contract choice for brand competition. Relative to licences and tenancies, the key structural differences in the franchise agreement are a (rather substantial) initial capital investment on the part of the retailer, in conformity with the principle of 'bond posting' in franchise; a longer duration (five years initially, plus an option to renew the contract for a further five years); and assignability of the contract (allowing the retailer to cash-in on goodwill he created). The initial bond (or 'membership fee') that franchisees are required to post has a different nature from the capital requirements for ordinary licensees/tenants, which – as we have seen – are essentially working capital. As is standard in 'business-format' franchise, franchisees are required to provide an initial bond in addition to the working capital requirement; this feature has been explicitly used by the petrol companies for screening purposes.

The role of contract duration (and tenure security) for incentives will be discussed in Section 5.

---

[6] At the time of its investigation in 1988, the MMC calculated that the (average) initial investment for a licensee with a shop on the site was of the order of £40,000 (against an income in the region of £23,000 per annum).

Structure of the Retailer's Compensation

In most cases, petrol retailers derive their income from a number of complementary activities carried out on the site: from the sale of petrol, to the sale of miscellaneous goods in the forecourt shop, to the supply of repair services. While they are seen by wholesalers as an integral part of their offers to retailers, in fact these ancillary activities are normally covered by a separate contract, running parallel to the contract for the fuel. The present analysis is restricted to contracts affecting fuel sales only.

In terms of the retailer's compensation for the sale of fuel, the different categories of contracts differ from one another in a number of aspects: in the 'typical size of the margin' recognized to the retailer (or, equivalently, in the 'typical range of the wholesaler price'); in the size/sign of the 'fixed fee'; and in the share of the site's 'operating costs' the retailer is directly responsible for.

On the first of these aspects, it must be emphasized that the definition of both wholesale and retail prices – and of the margin, as the difference between the two – is rather complex in the UK petrol industry. First, it is standard practice to calculate the wholesale price 'backwards', as the difference between an official 'scheduled price' (changed periodically, by public announcement) and a discount/ rebate which differs by category of contract (and sometimes has an individually-negotiated component). The final retail price is set in principle by the retailer, on the basis of this invoice price; however, contracts at wholesaler-owned sites also tend to specify that the retailer should not charge more than a 'maximum recommended retail price', defined as the sum of the invoice price and an official 'recommended margin'. This margin is also publicly announced by the wholesaler for all retailers in a particular class, and changed from time to time.

This type of arrangement implies at least a danger that the retailer's freedom to set the price may be *de facto* restricted in some circumstances. Indeed, the charge of implicit Resale Price Maintenance is often levied to the petrol companies: it was hinted at, for example, in the evidence given to the Trade and Industry Select Committee by the Petrol Retailers Association (*Petrol Retailing in the UK*, HC 384, 1988) (pp. 8, 14, 19), and explicitly accepted by the same Committee ('The effect of a combination of a price ceiling and a fairly small profit margin on petrol sales, which typifies the price mechanism used by all the major oil companies' led the then Director General of Fair Trading, Sir Gordon Borrie, to accept that *de facto* RPM was taking place; pp. x, xi). However, the MMC investigation which followed (1990) cleared the companies of this charge, arguing that factors such as competition from independent dealers and hypermarkets are sufficient to ensure that 'each wholesaler's control over prices is far from complete' (p. 301).[7]

In commission agency, the commission on the sale of fuel is derived from an estimate of the operator's costs (essentially wages) and of the expected volumes at the site: total estimated labour costs at the

[7] The practice of setting a margin, rather than explicitly defining a wholesale price, was also considered justified in an industry where exogenous cost shocks are frequent (e.g. changes in the dollar price of oil, or in the £/$ exchange rate): by setting the official margin, the wholesaler commits himself to a (minimum) relative distance between retail and wholesale price. Hence setting the margin is equivalent to specifying a *rule* for calculating the wholesale price, without defining it explicitly.

site (including an 'average' salary for the agent himself) are divided by the expected fuel sales, and the result is the commission paid to the agent. Of course, without appropriate safeguards such an arrangement would put all the risk on the operator; contracts therefore establish that if the agent fails to meet sales targets (on reasonable grounds), the wholesaler will supplement his income with a lump-sum transfer up to the forecast level. In case volumes exceed expectations, the agent retains at least part of the additional return – wherein lies the incentive aspect of the contract.

Contracts with independent dealers also imply a unit margin for the sale of fuel, generally calculated as the company's official scheduled price less an agreed rebate. Rebates are specified in the contract, and may vary slightly depending on the retailer's individual circumstances and bargaining power (for example they may incorporate repayment of capital sums provided by the wholesaler for site development). Dealers' margins are significantly larger than any found at wholesaler-owned sites: the MMC found that in 1988 the net wholesale price (after rebates and discounts) to dealers was on average about 1.5 pence per litre (ppl) below that of tenants/licensees, and the gap is thought to have widened in recent years to around 2 ppl. As we shall see in Section 5, this lower net price was explained by the companies to the MMC partly on grounds that retailers at wholesaler-owned sites are more risk-averse, and therefore the wholesaler will have to recoup the return on his investment in the site through a higher wholesale price mark-up, rather than through a fixed fee.

In addition, the comparatively higher margins offered by wholesalers to IDs reflect an element of 'competitive bidding' among them to attract more IDs to their brand. There is evidence that "as many as one quarter of (independent sites) may change supplier at the expiry of a solus tie" (MMC Report, p. 292); the previous MMC investigation (*Petrol. A Report on the Supply of Petrol in the United Kingdom by Wholesale*, 1979, p.19) also reported that (for a particular wholesaler) 28.4 per cent of ID sites supplied had been 'lost to other wholesalers', while 22.8 per cent had been 'gained' over a five-year period. Contract terms are a crucial element in the wholesaler's 'competition for agents': since supplying these sites increase market share without increasing overheads, 'competition among wholesalers to sign up independent retailers is strong, and this has been reflected in larger rebates. Wholesalers put considerable effort into both signing up new prospects and retaining contracts with their existing dealers. Negotiations for a new tie can start 18 months before the old one expires' (p.292).[8]

---

[8] The important role of independent dealers as a relatively inexpensive way to develop the retail network (and as such, an important complement to the wholesaler-owned network) means that wholesalers try to better their competitors' offers in order to attract retailers to their brand. In addition to the rebate, some of the solus contracts may include some form of financial assistance from the wholesaler, such as a grant or a loan on favourable terms for the redevelopment/improvement of the site, or the supply of equipment. Dealers have some degree of choice in the balance between rebate and other financial assistance. Indeed, dealers can choose from a fairly large range of possible options in the contract with their supplier:
– a large agreed rebate on the (wholesale) schedule price. The final size of the rebate is a matter for negotiation between the parties at the time the solus contract is set up. According to the MMC, 'the level will depend on the particular circumstances of the retail outlet, and on the degree of competition amongst wholesalers to sign up that dealer'. These 'rebate-only' agreements are about 30 per cent of the total;
– a rebate and some form of financial assistance (in this case, the rebate will be lower). Any combination of soft loans, subsidized loans, or capital grants are possible;
– sometimes, agreements might include an advance rebate: i.e. each year, the dealer could receive upfront a lump-sum

The end-result of this rather baroque structure is that the actual margin (per litre) realized by independent dealers tends to be the difference between scheduled price and rebates; for all tenancies, licences and franchises the actual margin will generally be the official margin, plus any additional rebates (and less any throughput royalties). Commission agents have a small margin calculated as the ratio of site labour costs to expected volumes.

Though no systematic data were available, discussions with petrol companies and surveys of the trade press tend to support the hypothesis that there is a 'typical range' for the margin in each agreement. That is, margins for the same type of agreement tend to fall within fairly narrow bands, across different companies; while for different contract types, the typical margins are different enough to justify a *ranking* of contracts with respect to this feature. In particular, margins are lowest for commission agencies and increase step-wise through licences, tenancies and franchises to independent dealers: recent information puts commissions in the range of 0.7–0.8 ppl; licences offer margins in the 1–2 ppl range; margins for tenants are normally larger, in the range of 2–3 ppl, as are margins for franchisees; while margins for dealers may be as much as 4–5 ppl (depending on whether the dealer is or is not in receipt of financial assistance from the wholesaler). This ordering, pictured in Figure 1 below, was found to be accurate for the contracts we examined directly, as well as being the consensus view in the industry (as ascertained in interviews with petrol companies, and the retailers' association).

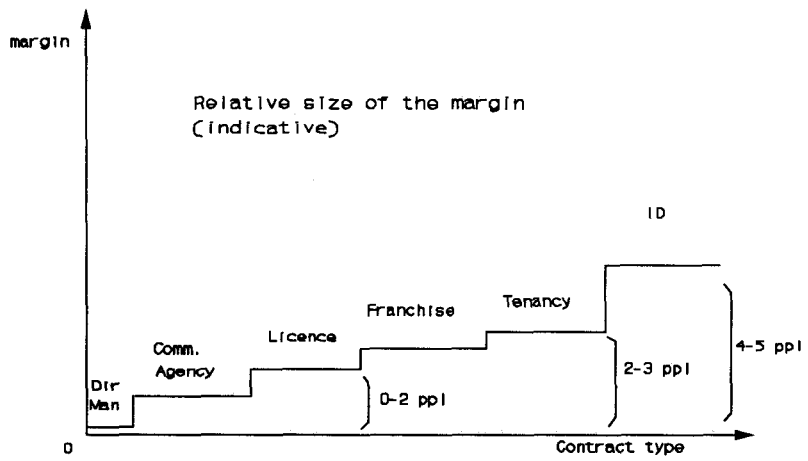

Figure 1

The MMC Report also confirmed this, describing a similar ranking for wholesale prices across contract types; in particular, the wholesaler's average 'net realizations/proceeds' for different categories of retailers were found to rank as follows:

---

payment which is equal to the total (or an agreed fraction) of the year's rebate on the expected volume (often there is a claw-back clause attached: if at year-end the volume-target is not reached, then the wholesaler claims back a proportion of the advance lump-sum payment).

The NRP data for 1988 outlined [below] show how the average net revenues from sales to retailers relate to those from other sales . . .

Wholesale petrol sales: summary of net realizations (NRP) for 1988	
Type of sale	*General ranges of NRP (ppl)*
Export sales	6.3 to 7.2
Wholesale sales to:	
Other wholesalers	7.2 to 8.8
Hypermarkets	7.5 to 9.4
Independent retailers	8.2 to 10.0
Tenants/licensees	9.3 to 10.7
Wholesaler-managed	9.3 to 12.5

. . . All the principal wholesalers achieved higher average NRP on petrol deliveries to retail sites that they own, than they do from sales to independently-owned sites. The average difference in 1988 amounted to about 1.2 ppl. . . . We also noted that wholesalers' sales of petrol through retail sites that they own generated higher average NRP if the sites were directly managed than if the sites were operated by tenants or licensees. The difference in 1988 amounted to about 0.7 ppl' (pp. 85–6).

The MMC evidence lends support to our working assumption that the ranking of margins described in Figure 1 is the mirror image of the ranking of wholesale prices. Thus – depending on the circumstances – we will formally identify a contract as a pair $\{m, F\}$ (i.e. a combination of margin and fixed fee), but also a pair $\{w, F\}$ (wholesale price and fixed fee).

The level of the periodic fixed fee is the other systematic difference across contract types. This may be paid quarterly, or twice-yearly, or annually, generally in advance. Tenants pay an annual lease fee, which is linked to expected volumes and is normally fixed for the duration of the contract. Some licensees and franchisees also pay an *ex ante* periodical fee based on expected volumes. Many licensees and commission agents receive instead a 'base wage' (i.e. a lump-sum transfer from the wholesaler). After adjusting for volumes, there appears to be a correspondence (at company owned sites) between the margin and the size of the fee – i.e. retailers tend to pay a larger periodic fixed fee, the larger the size of the margin they receive. For example, tenants (who receive the largest margin of all CO-sites) tend to pay a large annual fee (to include a fixed rent for the use of all site facilities), and licensees a smaller one; indeed in many cases licensees receive a base-wage from the company (formally a negative fee), as do commission agents.[9]

[9] A typical tenancy contract we were able to examine read: 'The Tenant shall pay the Landlord in addition to the minimum annual throughput rent of £ X, a fixed rent of £ Y for each period of twelve months during the continuance of this lease, payable by equal quarterly instalments in advance.' The sums 'X' and 'Y' are site-specific, and would be set separately for each individual site. A licensee will also pay a fee for the use of special facilities, buildings and equipment not related to the sale of petrol (e.g. carwash), but for the sale of petrol it is often the case that '[the Company] shall make a payment of a Monthly Operating Cost allowance to the Licensee of the sum £ Z'. Hence the reward to the licensee may be split between a fixed contribution from the company, and a (smaller) unit margin (in this particular case, the margin was 1.1 pence per litre, as opposed to 2.3 ppl for a typical tenanted site).

Finally, station operators under different contract types are responsible for different categories of retail costs (according to the standard terms of the contracts). Licensees pay the wholesaler a contribution to central advertising and promotions (usually a fixed royalty per litre), and are directly responsible for wages and miscellaneous expenditures (these may include insurance, power, rates, etc). In addition to these, tenants also pay maintenance/repair costs, as leases make the occupier of the premises liable for ordinary repairing costs. This important qualitative difference in the nature of the risk taken by the operator is reflected in the average unit margin recognized to the tenant on fuel sales, which is slightly larger than for licences. Commission agents are only responsible for staff wages, and not for other costs.

A summary of the main differences across contract types is given in Table 1.

While the fixed fee and the sharing of retail costs are important structural differences across contract types, the mentioned positive relationship between margin and fixed fee justifies at least in part our concentration on the size of the margin as the key discriminant between types. The agency analysis which follows is indeed based on the implicit assumption that all the structural differences among the main contractual arrangements may be entirely 'captured' by the 'typical size of the margin'.

Table 1

Main Economic Differences between Contract Types

	Dir.Man.	Comm.Ag.	Licence	Tenancy	Franchise	ID
Duration/ Security of tenure	-	1–3 years/ 30 days' notice	3 years/ no right to renewal	3 years/ renewal only with LTA protection	5 years/ renewable for 5–10 more yrs	5 years –
Structure of the retailer's compensation						
a. Base-wage received	Yes	Yes	sometimes (Esso, BP)	No	No	No
b. Margin per unit sold	No	Small (e.g. 0.7ppl)	Slightly larger (1–1.4 ppl)	Larger	Larger	Larger
c. *Annual fee paid*	No	No	Yes	Yes	Yes	No
d. Operating costs borne by the retailer	-	Staff wages	Wages,promotion general retail costs	As licence + maintenance costs	As tenant	All
Initial investment	-	None for fuel	Working capital	Working capital	Working capital + initial bond	Full

4 PETROL RETAILING IN AN AGENCY FRAMEWORK

The 'Agency Content' of Petrol Retailing

Moral hazard, which is the source of the classic principal/agent problem, arises when two fundamental conditions are fulfilled: (a) that the outcome of the activity delegated by the principal is affected by the agent's 'action', and (b) that such action is not directly (or costlessly) observable by the principal. In petrol wholesaling, the adoption of a principal/agent perspective for the analysis of vertical contracts is justified by the fact that both these conditions for the existence of moral hazard are satisfied. First, fuel sales at filling stations also depend on the retailer's effort, as well as on price and general demand conditions (which are uncertain *ex-ante*). Though the provision of effort may not appear as important for selling petrol as for other goods, the retailer's choice of effort is expected to affect in some measure volumes sold at the station, and thereby the wholesaler's profits. Most importantly, this is very much the perception of the petrol companies which design the contracts: in their concern with shifting attention away from prices onto other competitive elements, they have been putting increasing emphasis on brand image and high level of service (true product differentiation is notoriously difficult to achieve for petrol). The view espoused by the petrol companies is that a clean, well maintained petrol station with a variety of ancillary activities (shop, garage, etc.) is more likely to attract customers both for the fuel and the other services on offer, and this might generate positive externalities between them; great importance is therefore attached to the provision of effort, and the attainment of the desired standards by the retailers. In other words, retailers' effort is certainly an element of the wholesalers' objective-function. The other important aspect is that such effort cannot be directly/costlessly observed by the wholesaler at each station, which satisfies the second of the above conditions for moral hazard.

Contracts as Solutions to the Agency Problem

If effort were costlessly observable at each station, it would be possible in principle to achieve a first-best solution for the delegation problem: in a 'perfect control' situation (defined by Katz (1991) as the case where 'the P's [principal's] information is detailed enough and the contract space rich enough that he can order the agent to play any strategy the P desires'), the wholesaler could write the contract conditional on the retailer's effort. He would then set a forcing contract, where the retailer is instructed to provide the optimal amount of effort,[10] and is paid the equivalent of his reservation wage for that effort. If both parties were risk-averse, risk could also be shared optimally in relation to their respective coefficients of absolute risk-aversion; if only the retailer is risk-averse, his optimal compensation would be the fixed wage corresponding to the optimal effort ('the P can order the agent to choose any given strategy at a cost that leaves agent with no surplus in any state of the world' Katz, ibid.). If the retailer did not comply, since effort is observable he would be immediately found out and penalized (fined) for it, i.e. paid less than the reservation wage. In this full-insurance, forcing-

[10] Effort should be optimal when its marginal utility for the wholesaler (i.e. the marginal product of effort in terms of sales, times the wholesaler's marginal profit for each unit sold) is equal (in expected terms) to the retailer's marginal cost of effort.

contract case, there is no question of agent moral hazard.

However, while prices and quantities are perfectly observable at each filling station (conspicuously so, with prices prominently displayed on pole-signs), in the case of effort the limitations on the information available to the wholesaler make perfect control and forcing contracts impossible. Effort might also have dimensions which are difficult to quantify (e.g. customer relatioships), and the range of activities among which the retailer is to apportion his effort is rather diverse. The implication is that it is too costly for the petrol companies to set up a full monitoring structure, while providing incentives is relatively cheaper. Retailer's effort towards the attainment of higher service standards is therefore pursued by means of explicit incentive contracts (and to a lesser extent by monitoring, depending on relative costs).

Of course, if the retailer were risk-neutral the problem of moral hazard could still find a first-best solution through 'perfect delegation': with symmetric information at the time of contracting (no danger of adverse selection), and no risk of moral hazard on the part of the wholesaler (e.g. failure to make the necessary investments to keep the value of the brand, etc.), then residual claimancy would be the optimal solution. In this case, the retailer would pay for the fuel at marginal cost, and the wholesaler would be able to charge a fixed fee extracting the entire donwstream profit (this is equivalent to selling the asset to the retailer for a fixed fee, and letting him make the optimal effort choice). In the presence of retailer's risk aversion, however, there is no feasible contract yielding a first-best solution to the agency problem of inducing the retailer to provide the optimal effort.

The essence of the agency approach is that each contract, with its specific provisions, embodies a solution to the underlying agency problem, and in this way reveals the implicit terms of such problem. First, whenever the agent's compensation scheme is (even in part) performance-related, an incentive scheme is in place, which simultaneously defines the sharing of risk between the parties; the scheme will reflect the agent's risk-aversion (assuming the P is risk-neutral), as well as his ability or inclination for effort. Secondly, the contract's incentive provisions should also reflect the intensity of the principal's monitoring problem, which in turn might depend on the specific characteristics of the activity and/or of the asset. We thus expect the manufacturer (refiner/wholesaler) to optimally set the terms of the contract (i.e. so that the retailer will choose the 'optimal' course of action), on the basis of given retailer, asset, and market characteristics.

Contract Menus as Simplified Solutions to the Agency Problem

As the set of possible combinations of retailer/asset/market characteristics is potentially very large, the total transaction costs of individually negotiating a contract with each agent would be disproportionate. These costs should induce the manufacturers to narrow down their contract offers to a limited number of types, yielding a 'menu' of standardized options. In this framework, picking out a contract from a given (standard) menu, and offering it to the retailer(s) on a take-it-or-leave-it basis replaces the individual tailoring of contract terms to suit each specific situation. Both stem from the manufacturer's search for a solution to the agency problem, but using a menu drastically reduces the transaction costs

of contract negotiation because retailers 'self-select' on the basis of their characteristics.[11]

In the established menu of contracts found in UK petrol retailing, each format was seen to incorporate a different structure for the retailer's compensation – and therefore a different incentive mechanism. The general evidence that terms differ across types, but tend to be fairly homogeneous within each type, suggested that each type forms a separate category (i.e. different contract types tend not to be observationally equivalent); each separate category then represents the agreement which is 'best suited' at least for a certain range of circumstances, and will be offered when such circumstances arise. Thus, for given ranges of the underlying market/asset/retailer parameters, there should be one particular type of contract which is the wholesaler's preferred solution to his agency problem; this is the contract we expect to be adopted.

Of course, the parameters which ultimately determine the optimal incentive scheme are in a *continuum* (e.g. the retailer's profile, concisely expressed by his risk aversion and attitude to effort, can have potentially infinite values, as many as the possible combinations of these two parameters); while the contracts offered in the menu are a limited number of types. It is therefore to be expected that some degree of pooling will take place within each type. The fact that, for example, retailers with somewhat different individual profiles might be offered the same contract to operate a particular filling station is an inevitable expression of this pooling. The implication is that if no individual adjustment can be made on the lump-sum transfer, there is always a portion of the downstream rent the contract is not able to extract from the retailers which are 'better' than implied by the contract profile. This loss is however compensated, for the manufacturer, by saving on the transaction costs of negotiation.[12]

The purpose of the agency analysis in Section 5 is to identify the specific parameters of the underlying agency problem each contract implicitly solves.

---

[11] It is a standard result that firms can offer a menu of wage packages to compel heterogeneous potential employees to self-select according to their abilities; hence the form of contract can serve as a screening device to select retailers when their ability or motivation is difficult to ascertain. This is an argument extensively used for example in the literature on share cropping, which is justified as a contractual form to select able workers on the assumption that lower risk-aversion implies greater ability. (Hallegan (1978), Allen (1982)). The argument is that economic agents with low skills or motivation levels are less willing to accept a compensation package that makes their wealth largely dependent on residual income than are agents with higher skill or motivational levels. Sharecropping, like franchising, gives agents significant residual claims to localized operations.

[12] We may anticipate that any contract 'type' will be generally acceptable to 'better' retailers (e.g. less risk-averse, more able), because of the pockets of downstream rent the manufacturer cannot extract from them through the fee. There will be a point, however, at which they are better off switching to a different contract.

5 AGENCY ANALYSIS OF CONTRACTS IN PETROL RETAILING

The main differences identified in Section 3 across contract types were: the relative size of the unit margin recognized by the wholesaler; the size of the fixed fee (base-wage) paid (received) by the retailer; the size of the initial fee payable by the retailer (where applicable), and his initial capital commitment; and the retailer's security of tenure. In an agency perspective, these features define the incentive contents of each contract. The former two elements identify the *explicit* incentive provisions incorporated in the contract, and are discussed next; the latter may be interpreted as the *implicit* incentive provisions of the contract, and will be analysed in the following section.

Explicit Incentive/Risk-sharing Properties of the Contracts

Margins and Fees across Contract Types. As described in Section 3, petrol companies periodically announce their 'standard' margins (often differentiating between two broad categories of retailers: independent dealers, and licensees/tenants); these are reviewed on average every year, and changes are regularly reported in the trade press (which makes them public knowledge). These margins bear a greater or lesser resemblance to those actually realized by retailers, depending on the specific retailers' group. In the case of company-owned sites, though they are formally defined as 'maximum recommended retail margins', their true nature is that of 'minimum' margins the company commits itself to preserve for the retailer (if needed, by use of Selective Price Support). In this sense, the official margin incorporates an element of commitment on the part of the wholesaler to maintaining at least that reward for each litre sold. But the announcement of the official margin is only one aspect of the convoluted process of wholesale and retail price determination in the UK petrol industry. In the case of independent dealers, key features are the 'scheduled price' and various categories of rebates and allowances (often individually negotiated), which must be deducted from the scheduled price in order to arrive at the wholesale price. When all these elements are taken into account, petrol retailing margins for different contract types form a rather complex structure, which generates the ranking shown above in Figure 1.

In addition to the margin (wholesale price), the other systematic difference across contract types was found to be the level of the periodic fixed fee (paid quarterly, or twice-yearly, or annually, generally in advance). As described in Section 3, after adjusting for volumes, retailers at company-owned sites tend to pay larger periodic fees, the larger the size of the margin they receive.

Risk-sharing Properties of Wholesaler–Retailer Contracts. The positive relationship between fixed fee and margin (as the margin becomes smaller, the fee decreases and may become negative – i.e. a wage) is in line with the basic rules of risk-sharing in the presence of uncertainty and moral hazard: when the agent is risk-averse, and yet incentives for effort must be set because his effort is not observable, the variable component of his reward should be smaller. The impact of the agent's risk-aversion on his pay-structure is an extensively explored theme, for example in the literature on franchising with incomplete information: if both parties had perfect complete information about future demand and retail costs conditions, in the optimal scheme the wholesaler would charge the retailer a wholesale price $w = c$ (marginal production cost), while extracting all the downstream profit through

a fixed fee $F = \pi^r$ (no double marginalization). If, however, the parties have imperfect information on future demand and cost conditions, then a risk-averse retailer would not accept such a scheme: an *ex-ante* lump-sum payment calculated on the expected level of sales will turn into a loss if actual sales are below the expected level.[13] If however the franchisor sets a wholesale price mark-up above marginal cost, while reducing the franchise fee, the retailer receives at least some insurance. Hence, whenever the franchisor is less risk averse than the franchisee, the optimal wholesale price is determined as the marginal cost plus a mark-up, while the franchise fee is set at a lower level.

In the petrol retailing case, if the retailer were risk-neutral he could be charged for the petrol he sells at marginal cost (no double marginalization, and optimal for incentive purposes), while the downstream rent could be extracted through the fixed fee (making the retailer the residual claimant). If however the retailer is risk-averse, the variable component of his pay structure must be suitably reduced. Defining his absolute risk as the total variance of his income, for given exogenous variance of sales a reduction in risk may be achieved by lowering the unit margin (equivalent to introducing a wholesale-price markup), and lowering the fixed-fee, or offering a base-wage.[14] Hence in risk-sharing terms, the different margins (and corresponding fees) which are typical of each contract type imply different degrees of exposure of the retailer to income variation. Contracts with a lower unit margin (e.g. commission agency) do imply less risk for the retailer, who receives a base-wage; while contracts with a larger unit margin and a larger fee (e.g. tenancies) imply greater variance of income – greater risk.

In this framework, the simultaneous existence of different arrangements (risk-profiles) might be explained with the argument that retailers are heterogeneous in terms of their risk preferences; and different contracts (which may be summarized as different combinations of mean/variance of income) will be attractive to retailers with different attitudes to risk. The fact that he will lose more/gain more if volumes are below/above the expected level, implies that a retailer with a large margin must be less risk-averse – or he would not accept that contract. In practice, a commission agency will be accepted by a retailer with higher risk-aversion, while a tenancy, or an independent dealership contract, will only be attractive to less risk-averse retailers.

This aspect was emphasized to the MMC by the petrol companies, which argued that different contracts incorporate different methods for extracting the downstream surplus. In the case of company-owned sites, in addition to the retailer's surplus, this rent must include elements related to the ownership of the asset (e.g. a return on the capital tied up in the site, and the economic rent of land and buildings). The companies argued that because of some retailers' greater risk-aversion, they tend to receive their return on the capital invested in the site by charging a higher wholesale price to operators at company-owned sites, than to dealers. The MMC did not object to this argument:

---

[13] Caves and Murphy (1976), among many others, showed that under perfect information only a fixed fee would be used in franchising, but output royalties will be introduced in the presence of risk aversion on the part of the franchisee (the franchisor takes up at least part of the volume risk).

[14] With a linear schedule of the type $y^* = mx + F$, obtained by integrating the optimal slope of the retailer's compensation function dy^*/dx (see Rees, 1985), the constant of integration F can be equally positive or negative.

> The companies told us that, in the case of wholesaler-owned sites, it was customary for the costs associated with site ownership, together with a return on the invested capital, to be recovered at least in part through the level of the wholesale price charged to the site operator (p. 158).
>
> We were told that the differences in wholesale prices partly reflects the bargaining power of the independents, and are a recognition that (they) require a larger margin in order to earn a return on their own capital invested in the site. Conversely, a return on capital invested in wholesaler-owned sites is included in the higher wholesale price charged to those sites. An alternative to this approach would be for wholesalers to charge all types of retailers a similar price, and to obtain a return on the investment in the retail sites which they owned in the form of a fixed rental or licence fee. The wholesalers argued, however, that *this would shift too much of the business risk to the retailers who are less able to support it*, and as a result the development of their businesses would be inhibited. It seems to us that volume-related payments of this kind are one way in which different distributions of risk between the parties can be achieved and are not in principle objectionable (emphasis added, p. 300).

A Parallel: The Case of Beer. These considerations on the risk/reward profiles of the various contracts, and on the risk-sharing function they performed for retailers with heterogenous risk/income preferences, were prominent also in the MMC investigation into the beer industry (*The Supply of Beer*, HMSO, 1989). They are especially relevant as the beer industry is similar to petrol retailing in its institutional arrangements, and particularly in the existence of a number of different contracts for the relationship between brewer and publican.[15]

In a note on the 'Risk-sharing Aspects of Brewer/Retailer Contracts' submitted to the MMC, the Brewers' Society emphasized the nature of contracts as 'market arrangements' developed to take risk away from the 'undiversified' publican ('typically with a highly undiversified income position, and highly vulnerable to loss of trade') and 'spread (it) among capital market investors . . . via loan agreements, tenancy agreements, and the managed house system' (p. 455). Indeed the purpose of the contracts, and the very reason for their existence, was identified in the fact that they make possible different degrees of risk-sharing between brewers and publicans, each being best suited to publicans with a particular degree of risk aversion.

> Different contracts will appeal to individuals with different characteristics. For example, a wealthy and not very risk-averse person may prefer the form of implicit contract associated with a small free house. On the other hand, a management contract for a large house may better suit a person who is less wealthy or more risk-averse. And free trade loans and tenancy agreements offer a range of intermediate possibilities (p. 456).
>
> Now it may be true that there will always be a supply of people willing to take this risk, . . . but it would be wrong to think that there is an unlimited supply of [them] (p. 455).
>
> Because of the diversity of individual characteristics and circumstances, the diversity of contracts

[15] The types of contracts in use between brewers and publicans include *managed houses; tenancies;* and *agreements with free traders*, where the premises are owned by the publican, who buys the beer from a brewer under an exclusive supply agreement, and often receives a free or a soft loan from the brewer, with the aim of tying him for the future.

> available makes a positive contribution to economic efficiency and the public interest. In this it is no different from other forms of differentiated goods markets. Thus where individual tastes and incomes differ, it is generally desirable that, as here, a range of different offerings are available on the market (pp. 456–7).

The key difference among contract types is thus also identified in the size of the margin received by the publican (or, implicitly, the wholesale price level): different margins suit publicans with different degrees of risk-aversion, and the gradual decline of the margin from 'free-trade agreements' to 'tenanted houses' and 'managed houses' reflects the increasing risk-aversion of tenants and managers relative to free publicans: 'Given that volume sold is an uncertain variable – being partly influenced by market factors beyond the retailer's control – the flatter the reward/performance line the lower is the burden of risk borne by the retailer' (p. 456).

As in the petrol case, subsidized or soft loans from the brewer can also reduce the burden of risk for a free publican, because

> he is likely to face in return a higher wholesale price (the discount on list prices will generally be lower). This reduces overhead costs but raises unit variable costs, hence the free trader suffers less when trade is poor (the cost of his premises, analogous to rent, is lower) but gains less when trade is good (net revenue is lower because of the higher wholesale price than might otherwise apply). The risk faced by the free trader is therefore reduced (p. 455).[16]

Explicit Incentives for Effort in Petrol Retailing Contracts. In each contract, the intensity of the explicit incentive for retailer's effort is measured by his marginal reward on sales achieved through additional effort. Assuming a linear compensation scheme, of the type $y = mx(e) - F$ (where y is the retailer compensation, x is the outcome of his effort (sales), m is the unit margin, and F the fixed fee), the margin associated with the contract implicitly defines how additional volumes translate into extra return for the retailer ($dy(x(e))/dx(e)$). As higher levels of m bring the agent higher returns to increased effort, the margin measures the intensity of incentives for effort.

In providing effort, the retailer will compare his marginal benefit ($\frac{\partial y}{\partial x} \cdot \frac{\partial x}{\partial e}$, i.e. his return on the

[16] In line with the position of petrol wholesalers, the brewers also rejected any suggestion that contracts should be standardized: 'Enforced standardization would have the effect of restricting competition because it would artificially prevent brewers from offering types of contracts that best catered to the preferences and circumstances of particular classes of potential publicans. [It is hard to see how] prohibiting free-trade loans and tenancy agreements [contracts involving an intermediate degree of risk sharing] – and forcing contracts to the extremes of risk sharing associated with managed and free houses – . . . could be regarded either as a sign of deficient competition or as acting against the public interest. It is of course true that the existing mix of different types of contracts . . . can be expected to change as preferences and circumstances change. For example, if an increased supply of wealthier and less risk averse potential retailers is forthcoming, it is to be expected that the market will respond by producing a higher proportion of free-trade outlets or by moving to tenancy agreements with higher fixed rents and lower wholesale prices than would otherwise have been the case . . . Indeed, it is one of the tests of an effectively competitive market that supply should so respond when demand changes . . . Thus, the proportion of free trade outlets has increased over the years and there is continuing experimentation with different types of tenancy agreements . . . [Of course] it is possible to conceive of circumstances where de-integrated industrial structures might produce a similarly rich spectrum of contracts for retailers. [But in practice these] have not emerged in the marketplace' (p. 457).

additional sales achieved by supplying one extra unit of effort), to the marginal cost to him of supplying such effort; effort will be provided until the marginal benefit exceeds the marginal cost. This is the incentive-compatibility condition, or feasibility condition, in any incentive contract. Taking this constraint into account (i.e. considering only contracts which are feasible), the wholesaler will choose the specific incentive scheme which is optimal from his point of view; that is, the incentive intensity which optimally trades off the wholesaler's benefit from extra retailer's effort, with the agency cost of providing him with the incentive to do so. These depend, as we shall see, both on the retailer's and the filling station's characteristics.

At its simplest, the incentive analysis of contracts focuses therefore on one key indicator, the 'typical size of the margin', as the discriminant factor among contract types. The margin is the concise expression of the insurance-incentives tradeoff, and ultimately represents the solution to the underlying agency problem. If a different 'typical margin' is what fundamentally separates types, the problem of contract assignment then becomes a problem of relating the typical margin for each contract to a number of explanatory factors. The same agency problem (delegating the retail function to an agent) will have different solutions depending on the parameters of the problem; thus different 'typical' margins (and different contracts) reflect differences in the following underlying parameters:

(a) the characteristics of the asset (attributes of the filling station);
(b) the individual retailer characteristics;
(c) and possibly also the company type (or corporate policy).

The objective of the investigation is therefore to relate each contract type to particular values of the underlying parameters which are likely to have generated it, and thereby formulate testable predictions on the determinants of contract choice. Before describing a simple model which brings in these factors in contract assignment, let us consider briefly the implicit incentive provisions of the contracts.

Implicit Incentives for Effort

'Explicit' incentives for effort are those created by design, making the agent's pay directly dependent on performance. In addition to these, petrol retailing contracts may also incorporate *implicit* incentives, which do not imply a direct link between current effort and current rewards: these are provided by the length/tenure security of each contract, and (where applicable) by the initial lump-sum payment 'posted' by the retailer at the start of the contract.[17]

It must be emphasized that the formal analysis of implicit incentives (requiring a dynamic approach), is *not* within the scope of this thesis; nor are the incentives implicitly provided by the structure of ownership (of the asset), i.e. by the allocation of residual returns. The focus of this thesis is strictly the design of optimal *explicit* incentive contracts for the retailer's choice of action, in a predominantly *static* perspective (i.e. where incentives are set for effort in the current period). However, since these

[17] And, to the extent that capital markets are imperfect, also the initial working-capital requirement of each contract.

elements are likely to be important features of the overall incentive structure incorporated in the contract, we briefly comment on them below.

Tenure Security, and Contract Duration. An implicit incentive for effort may be provided by the retailer's fear of *early termination/non-renewal* of the contract (which is inversely related to his security of tenure). The argument is that when tenure security is low, the retailer knows that if he does not conform to the company's standards of operation the contract will not be renewed (or may even be terminated at short notice), and this should provide him with an additional incentive to perform. In particular, early termination (or non-renewal) implies for the retailer the loss of the future income stream associated with the contract; in deciding whether to provide effort or not, the retailer may be expected to implicitly compare his gain from shirking in the current period, to the present value of future income losses if the contract is terminated.[18] In this sense, the role of low security of tenure as an implicit incentive for retailers' effort has some analogies with the threat of termination in Shapiro and Stiglitz' (1984) analysis of 'efficiency wages'.[19]

Petrol-retailing contracts differ substantially in their tenure provisions: commission agents do not have any security (they need only be given thirty days notice of termination), licensees have limited security

---

[18] The notion that the 'fear of interruption' has incentive properties for the agent has been featured in the literature on franchising. Klein and Saft (1985) identify termination of the franchise contract as the most effective penalty in preventing free-riding between franchisees (i.e. preventing a franchisee from cheating on the group, by supplying a substandard product). 'Termination imposes an implied loss of a future quasi-rent stream. If the franchisee's estimate of the discounted present value of this future quasi-rent stream is greater than his estimate of the current short-run extra profits from cheating, then he will not cheat. The franchisor must create an arrangement in which the franchisee's expected discounted value of future rents is greater than the franchisee's expected short-run cheating potential. The franchisor must monitor franchisees and set the level of its policing expenditures so that, given his estimate of the cost saving to a franchisee supplying low-quality inputs and the implied expected time to detection and enforcement, the short-run profit a franchisee is expected to earn by cheating is below the value of the future quasi-rent.' In particular, Klein and Saft suggest that the franchisor may create quasi-rents by requiring a franchisee to invest in specific (not fully salvageable) production assets on which the franchisee is earning a normal rate of return but which, on termination, imply a capital cost penalty. If this penalty is larger than any short-run gain from cheating, the franchisee will not cheat. One possibility they suggest for the creation of such a specific asset is the requirement that the franchisee rents short-term from its franchisor (rather than own) the land on which the outlet is located. This lease arrangement implies that the franchisor can require the franchisee to move, and moving imposes a capital loss on the franchisee up to the amount of his initial non-salvageable investment. This creates a form of collateral to deter franchisee cheating.

[19] Assuming that (perhaps because of wealth limitations) the only effective means available to discipline an employee into providing effort (not shirking) is to terminate his employment (i.e. no explicit incentives), Shapiro and Stiglitz (1984) show that the threat of termination of employment may be an effective deterrent against cheating. This depends on what the employee stands to gain if cheating is undetected (g), how likely cheating is to go undetected (p), and the difference between the wage offered by the employer (w) and what other market opportunities the employee may have (w^0), i.e. the 'rent' $(w - w^0)$ that the employee earns from the job. In particular, they show that it will not be in the employee's interest to cheat if

$$g < p(w - w^0)N$$

i.e. if the gain from cheating is lower than the probability that cheating is detected, times the loss of income from being fired; this is adjusted by N, which is term expressing the long-term value of this relationship, including both the discount rate and the number of periods during which the worker is employed by the firm. The possibility of earning this rent makes the job and the prospect of being rehired in the future valuable to the employee and makes being fired an outcome to be avoided.

(three-year contract, three months notice), and so do tenants without LTA protection. Petrol wholesalers are certainly aware of the implicit incentives incorporated in reduced tenure, and made this view clear in the course of the MMC investigation. They complained that in the past, long leases had made it difficult for them to weed out the least competent operators; they had not provided an effective screening of retailers, and had not stimulated the provision of effort on their part. Many companies had in fact switched from long leases to licences and commission-style agreements, reducing the operators' security of tenure, precisely in the belief that much shorter terms and uncertainty on renewal would provide an incentive for effort.[20] As reported by the MMC,

> We asked BP what it thought of the suggestion that . . . licensees should be given greater security of tenure, so as better to be able to operate as they see fit. BP told us that it would in such circumstances have less control over the site in which it had invested, leading to reduced standards and impaired quality of service, and to the possibility of retailers charging more than would otherwise be the maximum price (p. 373).

Predictably, the retailers favour instead longer terms, on grounds that increased security would better justify their own investment in the site, and the degree of financial responsibility they undertake; in addition, they argue that the lack of proper compensation in case of non-renewal implies the appropriation by the wholesaler of the goodwill created by the retailer; and this is a deterrent to the provision of effort on his part (and possibly even a deterrent for the most able individuals to entering the industry). Recently, the argument that low tenure-security could attract less able retailers appears to have found more favour at least with some companies, as suggested by the introduction of a longer-term contract like the franchise (with possible recognition of goodwill). No evidence exists of the actual impact of greater or lower security; however a study of a similar problem in franchising could be taken to support the idea that reduced security is an effective way of providing incentives. This is described below.

A Parallel: Termination Laws in Franchising. Indeed the arguments for and against increased security of tenure present strong affinities with the US debate over termination laws in franchising. These are laws restricting the franchisor's ability to terminate the franchise contract only to 'good cause' (that is, cause that can be proved in court); such laws are present in fourteen states, and their introduction

---

[20] Esso, for example, told the MMC that it had moved from tenancies to licences because 'this allowed better adaptation to the changing market conditions' (p.368) and that it had been 'very successful in improving facilities and standards at licensed sites, whereas tenants were obstructive' (ibid.); BP, which had also switched to licensees in the 1970s, emphasized the undesirable effects of tenants' greater tenure security: 'Tenanted sites do not offer the advantages of licensed or managed sites in terms of optimizing site potential for petrol retailing. The benefits provided to a tenant under the terms of a lease, often of 21 years' duration, largely prevent the company from implementing its strategies in the short and medium term, and the rights to renewal enjoyed by the tenant can effectively deny the company the opportunity to realize the potential of its property in the longer term. Although the lease may specify standards of operation, an unsatisfactory lessee may nevertheless remain in possession of the premises. In contrast, a licence affords BP an effective degree of control over the capital invested in the wholesaler-owned forecourt facilities' (p.371). At the same time, Mobil felt that 'for most of [its] sites, a tenant was appropriate because the security of tenure attracts a better quality of operator who feels that he has a stake in the business and would build it up' (p.374). As for direct management and commission agency, the emphasis was very much on control issues. Of its network of directly-managed sites, BP said it offered the opportunity 'to gain first-hand experience of petrol and related retailing activities, and develop new marketing policies which could be applied across the network as a whole' (p.372). Texaco also regarded its directly-managed sites as the flagship of its retail chain.

is considered by others. The proponents argue that these laws would protect franchisees from opportunistic behaviour by franchisors, but franchisors are vehemently opposed on grounds that these laws increase the costs of controlling quality within a franchise system. This is because 'good cause provisions' require increased payments to the franchisee in case of termination unless good cause can be documented in court, which is expensive. (In addition, the 'termination laws' also usually restrict the ability of the franchisor not to renew existing contracts, and also place requirements on the minimum notice for contract terminations). Thus they potentially reduce the threat of termination, and without this threat franchisees have less of a deterrent to shirk. That is, franchisees have a greater incentive to free-ride, and the costs of controlling quality are likely to increase.

The implication from this argument is that 'good cause termination laws' should reduce the incidence of franchises: faced with an increased incentive to shirk by franchisees, franchisors should be inclined to reduce their proportion of franchised outlets in favour of company-owned outlets. Brickley, Dark and Weisback (1991), predict that if the laws increase the cost of franchising relative to company ownership, franchisors should shift marginal operations toward central ownership. They define 'marginal units' as those with non-repeat customers, where incentives to shirk on quality are highest, and test this prediction by comparing the incidence of franchises across states with different laws. Their empirical results, on a sample of US franchise industries, confirm that there is significantly less franchising in states *with* termination laws. This is interpreted as support for the hypothesis that termination laws increase the costs of franchising relative to company ownership, by making quality control among franchisees more expensive. This could explain the reduction in franchising.

Implicit and Explicit Incentives. It is also conceivable that the implicit incentive provided by the retailer's fear of termination might affect the intensity of the (optimal) *explicit* incentives designed by the wholesaler. Though we will not develop this theme formally, we suggest a possible intuition.

Consider a simplified setting with only two periods (current and future), and two possible contractual arrangements: one two-period contract, and two one-period contracts, with the probability of renewal at the end of the first period positively related to effort in that period.[21] The question is how the retailer's effort choice in the current period will be affected by the fact that such effort increases the probability of the contract being renewed; and therefore how this might affect the optimal explicit incentives.

Define the retailer's inter-temporal income (over the two periods) as

$$\Pi_1(e_1) + p(e_1)\Pi_2$$

where $p(e_1)$ is the probability of renewal, as a function of effort in the first period, $\Pi_1(e_1) = m_1x(e_1)$ is the retailer's gross income from sales in the first period, and $\Pi_2 = m_2(x_2)$ the retailer's income in the second period (with zero discount rate). Sales $x(e)$ may be defined as $x(e) = e + \theta$.

Consider first the case where there is no build-up of volumes through effort over time, i.e. effort in

[21] Of course, since effort is unobservable, the probability of renewal will depend on its signal, i.e. the actual sales volumes at the station.

the current period only increases current volumes, but the additional volume does not 'carry over' into the next period. This may be described as a case where there are 'no spillovers' of effort across periods, and in the above notation it implies $\Pi_2 = \Pi_2(e_2)$. In this case, the retailer's return to one unit of effort in period 1 is just m_1 in the case of one two-year contract, but it is $m_1 + dp/de_1\Pi_2$ in the case of two annual contracts (the marginal return to effort includes the marginal increase in the probability of being renewed, times the second period's income). Thus the retailer has a greater incentive to provide effort in period 1, if this increases the probability that he will appropriate some rent in the second period, and can be taken to describe the 'wholesaler's argument'.

Consider however the case where effort in the first period 'spills over' into the second, i.e. volumes in any given period are some function of the integral of past effort. In this case, the second-period return is more appropriately defined as $\Pi_2 = \Pi_2(e_1, e_2)$. The retailer's marginal return to effort in the second period is as follows:

– in the case of one two-year contract, where renewal is not an issue, the current choice of effort will take into account the fact that some of the volumes gained in the first year are also part of the second year's income (the retailer receives a margin m_1 in the first year, and m_2 in the second). Therefore his marginal return to effort in period 1 is

$$m_1 + \gamma m_2$$

(where γ is a parameter representing the degree of spillover).

– In the case of two annual contracts, the marginal return will be

$$m_1 + dp/de_1\Pi_2 + p\gamma m_2$$

because the return γm_2 will not be realized if the contract is not renewed. Unlike the previous case, where the comparison of the two contracts yields unambiguous results (the retailer's return to effort is greater with two one-year contracts, and therefore his incentive for effort will be greater), with intertemporal effort spillovers the result is ambiguous. Since it is not impossible that $m_2 > dp/de_1 + p\gamma m_2$, it is not certain that the retailer has a greater incentive for effort when the contract is shorter. The intuition is that if effort in the current period spills over into future periods, the retailer's effort choice is analogous to an investment decision (in effort), and he is less likely to make this investment if there is a probability that he will not benefit from it in future. This effect could outweigh the incentive to put in effort to increase the renewal probability. This appears to fit the retailer's side of the debate on the merits of shorter, less secure contracts. Indeed, it appears to lend support to the argument that longer contracts give retailers an incentive to 'build up the business', and to their complaints that their goodwill is not recognized.

The implication of this approach is that if the total effective incentives provided by a contract is the sum of its explicit and implicit provisions, it may also be the case that the implicit incentives 'feed back' into the explicit incentives for the current period. In the above notation, the retailer's Incentive-Compatibility condition for the case of two shorter contracts, and no effort spillovers, becomes

$$C'(e_1) = m_1 + dp/de_1\,\Pi_2$$

This could suggest, for example, that where implicit incentives are stronger (e.g. where the contract is shorter, or less secure), the explicit incentive the wholesaler must set to elicit the desired level of effort from the retailer might be lower. This is not contradicted by the stylized fact that in petrol retailing, contracts with lower unit margins also tend to have shorter duration, and lower security of

tenure (e.g. commission agency, licence). However the above case of effort spillovers also suggests that the wholesaler will have to weigh carefully the benefit to him of being able to set lower explicit incentives (lower agency cost), against the drawback of possibly attracting lower-quality retailers without the interest or the ability to develop the site.

Bonding

In franchising, franchisees are normally required to make an up-front lump-sum payment ('franchise fees'). Traditionally (e.g. Blair and Kaserman, 1982), the function of a fixed fee is the extraction of the downstream surplus on the part of the franchisor, though full extraction could only be possible in a complete information framework (in which case the franchise fee would equal the total profit of the retail outlet during the contractual period). With uncertainty about future demand and retail costs, and therefore on future profitability, the entry fee will be reduced and combined with an output royalty. Apart from this risk-sharing mechanism, other reasons have however been proposed for the combination of entry fee and output royalty which is typical of franchising (but also of other vertical contracts): these emphasize the function of fee and royalty in countering possible externalities which might arise in the relationship between franchisor and franchisee, or among franchisees (i.e. from vertical and horizontal externalities). For example, Norton (1988) argues that the existence of brand name capital makes the parent company (and other local outlets) vulnerable to quality debasing by a local outlet. Using local managers who make a large site-specific investment and post a large bond in the form of a franchise fee makes quality debasing less likely, because a franchisee has much more to lose upon termination than a local employee-manager. This aspect is also emphasized by Mathewson and Winter (1985): without some significant personal commitment, the franchisee has an incentive to reduce local quality and free ride on the national brand name, and appropriate any consequent rents by declaring that low demand has occurred. At the same time, the royalty provides some assurance that the franchisor will in fact provide the (costly) policing service necessary to maintain the value of the franchise, and it is therefore for the franchisee a guarantee against hold-ups on the part of the franchisor. The point is that the franchisee will not pay the present value of future monopoly profit without some guarantee of being able to collect such profits in future; without any safeguards, there is a risk of the franchisor 'cheating (i.e. selling more franchises once he has collected the first fee, and not maintaining the value of the franchise with appropriate investment, and monitoring).

In petrol retailing, a few contracts (notably the franchise, but also a few licences) do require the retailer to make an up-front payment. The role of this payment (in addition to any annual fees paid in connection to specific facilities/equipment provided, e.g. the shop or carwash)) is unmistakeably that of a 'bond' posted by the retailer, thereby providing an implicit incentive for his effort: the retailer will try to do especially well when a large portion of his wealth is irretrievably tied up in the site. In addition, the requirement that the retailer pays a significant *ex-ante* fee also provides the wholesaler with a screening device, complementing the screening role performed by the various working-capital requirements of the contracts. The argument used by the wholesalers is that the retailer's ability to provide this sum, either from his own wealth or borrowing from a bank, gives a signal of his entrepreneurial skills.

6 SUMMARY

In the UK petrol industry, the relationship between wholesalers and retailers may take the form of different vertical contractual agreements. These constitute a menu of established options (with only occasional innovations), from which the wholesaler chooses a particular 'type' for the operation of each particular station. The observation that wholesalers tend to adopt a mixture of contracts for the operation of their retail network suggests that the assignment of contracts is an exercise in matching the (heterogeneous) outlet characteristics with an appropriate incentive structure. In this approach, the menu of options – with heterogeneous retailers 'self-selecting' on the basis of contract offers – may be interpreted as a set of implicit 'solutions' to different (underlying) incentive problems (and a way of reducing the prohibitive costs of individual contract negotiations).

In this paper, we have provided a detailed description of the contract types in use in UK petrol retailing (direct management, commission agency, licence, tenancy, franchise, independent dealership). These were found to differ in three respects: the structure of the retailer's compensation (in particular, the typical size of the margin on fuel sales, the size of the fixed fee, and the retailing costs the station operator is responsible for); the duration of the contract, and the retailer's security of tenure at the site; and the initial capital investment the retailer might be required to make. The purpose was to identify their respective incentive/risk-sharing properties, and thereby the terms of the underlying agency problem which each contract implicitly addresses. In particular, the analysis distinguishes between the contract's *explicit* incentive provision, which directly relate current pay to current performance; and *implicit* incentives, which originate from more long-term (often dynamic) concern – such as tenure security. The focus is on the former: different contracts are found to be characterized by different typical margins, and correspondingly different fees, which imply different properties in risk-sharing/incentive terms. The reduction of contractual differences to different typical margins (and fees) is crucial for the models of contract design proposed in Sections 4 and 5; this mapping of the relative position of the contracts will be used to test the contract assignment rules defined by the models.

Appendix

Contract Use by Brand, 1993

Patterns of contract choice by selected UK brands

	Dir. Mng.	Comm. Agency	Lic.	Ten.	Franch.	Total CO	Indep. Dealers
Esso		100	900	10		1,010	1,246
Shell	164		607		245	1,016	1,224
BP	120		500			620	952
Texaco	120	70		400	12	602	797
Burmah		97	15	55		167	966
Jet	50	80	110			240	902
Mobil	211		58	256		525	293
Fina	80	90		90		260	362
Total	60	10	380			450	147

Source: collated from various trade journals and interview evidence

Recent developments in the types of contracts offered

Shell:	**Licence**	→	**Franchise (1991)**
Jet :	**Comm.Agency**	→	**Licence (1991)**
Texaco:	**Tenancy**	→	**Franchise (1992)**
Fina:	**Dir.Management**	→	**Comm. Agency (1992)**
BP:	**Licence**	→	**'Franchise-type' (1993)**
Burmah:	**Tenancy**	→	**Comm. Agency (1993)**

OXFORD INSTITUTE FOR ENERGY STUDIES
57 WOODSTOCK ROAD, OXFORD OX2 6FA ENGLAND
TELEPHONE (01865) 311377
FAX (01865) 310527
E-mail: publications@oxfordenergy.org
http://www.oxfordenergy.org