

Ownership rights and incentives in franchising

Nancy A. Lutz¹

*Department of Economics, Virginia Polytechnic Institute and State University, Blacksburg, VA
24061-0316, USA*

Abstract

I focus on the incentive effects of asset ownership in franchising. Franchise contracts give a manager ownership of some local assets; the franchisor owns other assets, notably the trademark. Under double moral hazard, the allocation of ownership affects the incentives of both the franchisor and the franchisee. I compare franchising with company-ownership of all assets. Franchising the local unit gives the manager strong incentives, but gives the central firm weak incentives. Franchising may be the preferred organizational form when the local manager's effort has a relatively small effect on the unit's current profit, but a large effect on the unit's future profit.

JEL classification: L22; M31

Keywords: Franchising; Double moral hazard; Vertical integration

1. Introduction

Chain businesses are an increasingly important part of American service and retailing industries. Such chains are made up of (possibly widely dispersed) individual stores using a single trademark and business method. The

¹ I thank Doug Eckel, Francine Lafontaine, Scott Masten, David Scheffman, and the participants at the Conference on Franchise Contracting, Organization, and Regulation for helpful comments. Thanks to Merton Peck, Joe Tracy, and especially Nancy Gallini for many discussions of the ideas modelled here.

profits obtained at any individual location depend on many factors. The trademark and business method used are crucial; so too are the day-to-day decisions made by the manager of the local outlet. The management of such a retail chain poses formidable challenges for coordination and control. Retail chains employ various organizational forms to meet this challenge; among these are franchising and vertical integration.

Franchising is a complicated organizational form involving a central party (the franchisor) and many individuals (the franchisees) who own and operate local stores using the franchisor's trademarks and business methods.² The relationship between a franchisor and a franchisee is governed by the terms of the franchise contract. These contracts generally have several main clauses covering payments, obligations of both the franchisor and franchisee, and ownership of the various assets employed in the chain. In a 'typical' franchise, the franchisor owns the trademark and the franchisee owns the local unit and its equipment. The franchisor retains extensive rights of control over the franchisee's activities.

An alternative to franchising is vertical integration. In this case, the central firm owns not just the trademark and business methods, but also the local assets; the central firm 'company-owns' the local unit. The manager of the local store is employed by the central firm. The relationship between the central firm and the local manager is then governed by an employment contract. This contract specifies the payments to be made to the employee and gives the employer the right to specify the methods and practices to be followed by the employee.

Franchising and company ownership might be distinguished on at least three grounds. This paper presents an explanation for franchising that highlights the importance of asset ownership. Franchisees generally own most of the assets used at the local unit, while employee managers do not own any of these assets. I argue that asset ownership plays a fundamental role in determining the local manager's incentives under any feasible compensation scheme. The optimal compensation scheme for the local manager depends on who owns the local assets; different ownership schemes and their associated compensation schemes imply different outcomes.

² I use the term 'franchising' to refer to what the U.S. Department of Commerce defines as business format franchising. 'Business format franchising is characterized by an ongoing business relationship between franchisor and franchisee that includes not only the product, service, and trademark, but the entire business format itself – a marketing strategy and plan, operating manuals and standards, quality control, and continuing two-way communications.' (USDOC (1986) p. 3). The second form of franchising, 'product and trade name franchising', is used only in automobile retailing, gasoline retailing, and soft-drink bottling. This second form is marked by a less extensive relationship between a product manufacturer and independent retailers who buy the product for resale.

Franchising might also be distinguished from company-ownership for two additional reasons. First, an employment relationship may give the central firm more extensive rights to specify the actions taken by a local manager than are possible under franchising. However, observed franchise contracts typically give the franchisor nearly unrestricted authority to specify how a franchisee should conduct his business. While a franchisee clearly retains some freedom of action, this freedom depends less on the terms of the contract than on the difficulties inherent in monitoring a local manager. Franchisees are able to choose their actions because it is difficult (or expensive) for the central firm to observe all their actions and verify these actions to a third party. There is little reason to think that the ownership form itself directly affects the observability of the local manager's actions. Therefore, it seems reasonable to conclude that the central firm's ability to directly control the local manager's actions does not differ significantly between a franchisee and an employee manager.³ As other authors have concluded, franchisees are effectively no less subject to the central firm's control than are the managers of company-owned stores. Gillian Hadfield claims that 'With respect to control, franchising relationships are closest to the employment end of the continuum: the franchisor typically exercises significant amounts of relatively unrestricted decisionmaking authority.' (Hadfield (1990) p. 932) Paul Rubin has noted that '... the franchisor has almost complete control over the franchisee –in fact, the relationship is almost that of a firm and an employee.' (Rubin (1978) p. 225)

Another possible way to distinguish franchising from company-ownership is in the form of the local manager's compensation. Under a typical franchise contract, the local manager retains the local profit, making a fixed payment and a variable royalty payment to the franchisor. Under company-ownership, the local manager receives a wage from the central firm. It is frequently assumed in the literature that this wage is fixed. Thus franchising, according to this definition, is fundamentally a profit-or revenue-sharing arrangement between the central firm and the local manager.⁴

There is however no obvious reason why the employee manager of a company-owned unit must be paid a flat wage. As with other employees, incentive payments are possible. While not all incentive payments take the form of profit-sharing, there is no *ex ante* reason to rule out incentive payments that are conditioned on unit profits. There is limited direct evi-

³ While employee managers are frequently the subject of more frequent monitoring by the central firm, this monitoring is an endogenous decision. Nothing in the typical franchise contract restricts the franchisor from monitoring a franchisee as closely or as intensively.

⁴ See Klein and Saft (1985) and Mathewson and Winter (1985) for examples of this definition. See Minkler and Park (1992), Minkler (1992), and Schmidt (1993) for examples of the assumption that the wage paid to employee managers is necessarily fixed.

dence about the actual use of incentive payments for the managers of company-owned outlets of retail chains. Krueger (1991) claims that ‘The long-standing practice in the [fast-food] industry has been to pay managers of company-owned restaurants a fixed salary that does not directly depend on their performance or their restaurant’s profitability.’ (p. 78), but provides no evidence for this claim. In fact, there is some evidence that incentive payments have been used by several fast-food chains, and that these payments are linked to unit profits, revenues, or costs. McDonald’s, for example, in the late 1960’s used a compensation plan that tied the base salary of a unit manager to a measure of direct effort, but tied the manager’s bonus to unit profits.⁵ The chain also gave serious consideration during the early 1970’s to a plan that would make a manager’s compensation dependent almost entirely on unit profits, with up to 20% of the profit going to the manager (Sasser and Pettway, 1974.) Au Bon Pain, which competes in the fast-food market with a chain of entirely company-owned outlets, adopted its ‘partner-manager program’ in 1987. The program is a compensation plan under which managers receive a base salary of \$500 per week with a 35% share of the unit’s incremental profit; an associate manager receives \$400 per week and a 15% share of profits. The program was immediately successful in increasing per-store profits (Hallowell, 1993). In 1988, RKO Warner Video adopted a compensation plan for its chain of company-owned video rental stores that tied manager’s bonuses to store revenues (Shimer, 1993). There is therefore little reason to conclude on either theoretical or empirical grounds that the compensation offered to employee managers necessarily takes the form of a fixed wage; identifying franchising with share contracts is unconvincing.

Several observers have concluded that even when incentive payments are used the compensation of employee managers is necessarily less incentive-based than that received by franchisees.⁶ While the stylized facts may support this conclusion, if we are to accept a compensation-based definition of franchising this conclusion should be derived from the fundamental features of the employment and franchising relationships. Employment contracts would be necessarily less incentive-based than franchise contracts if the employer cannot require the employee to pay a fixed fee at the beginning of their relationship, while franchisees could be required to make such payments. This would mean that the central firm could not simultaneously make an employee manager the residual claimant to the firm’s entire profit stream and at the same time extract all rents from the manager. In contrast, the

⁵ The measures of direct effort employed were actually the unit’s scores on McDonald’s standards for quality, service, and cleanliness (QSC). See Sasser and Pettway (1974).

⁶ See Norton (1988 and Norton (1989) and Lafontaine (1992b).

franchisor can extract all rents from the franchisee in the form of a fixed franchise fee.⁷

There are two reasons why it may be impossible to require that employees pay a fixed fee. First, a large fee gives the central firm an incentive to act opportunistically, terminating the relationship early and confiscating the fee. Second, a manager may be unable to raise enough capital to make a large fixed payment. However, any argument that employee managers cannot post bonds while franchisee managers can pay fixed fees then must rest on an exogenous distinction between franchisee managers and employee managers. Either employers will act opportunistically while franchisors will not, or only employees, not franchisees, face capital constraints.⁸ Without some careful justification of these assumptions, we cannot conclude that franchise contracts are necessarily more incentive based than employment contracts, and cannot identify franchising with a compensation scheme for the local manager.

In this paper, I use an ownership-based definition of franchising to examine the incentive problems faced in a single store in a retail chain. I assume that the profits of the local unit depend on actions taken by the central firm, as well as on actions taken by the local manager, in two ways. First, these actions affect the unit's current profit. Second, the actions have continuing effects on the unit's long term stream of profits. I also assume that all contracts, employment and franchise contracts alike, are incomplete in two important dimensions. The contracts cannot specify the actions to be taken by the two parties. Also, the contracts cannot credibly commit the two parties to a division of future surplus that is not supported by ownership rights. The first dimension of contractual incompleteness means that the local unit is subject to all the problems inherent in any situation of double moral hazard. Because forcing contracts are impossible, both the central firm and the local manager choose their actions to maximize their own return under their contract. Contract terms that give strong incentives to the central firm give weak incentives to the local manager, and vice versa. Both franchise and employment contracts are affected by this moral hazard problem.

The second dimension of contractual incompleteness is a dynamic one. I assume that it is impossible for any contract to specify exactly how the entire stream of profits is to be split between the central firm and the manager into an indefinite future. Because current effort affects future profit, each party's incentives depend in part on the expected division of future profits. These

⁷ See Norton (1988 and 1989) and Brickley and Dark (1987) for examples of this argument.

⁸ In fact, Kaufmann and Lafontaine (1994) find that at least some McDonald's franchisees face significant capital constraints. I will discuss the impact of such constraints in more detail in the conclusion.

expectations, in turn, are determined by the ownership of the various assets used in production. By owning the local unit, a franchisee has a credible claim to at least part of that unit's future profit stream. Employee managers, who own no assets, have no such claim. Thus the manager's incentives under any given division of current period profits are stronger under franchising than under company-ownership.

At the same time, ownership form affects the central firm's incentives. When the unit is company-owned, the central firm expects to receive all of the unit's future profits. When the unit is franchised, it will have to share those profits with the franchisee. Thus the central firm's incentives are stronger under company-ownership than under franchising.

Using this ownership definition of franchising, I reach several additional conclusions. First, I find that the first-best outcome is not achievable in a company-owned outlet; under certain conditions it is achievable in a franchised unit. Second, in any given outlet the manager retains a smaller fraction of current profits if the outlet is franchised than if it is company-owned. Third, the ownership form for any individual unit depends on the relative impact of the manager's action and the central firm's action on current profit, as well as the relative impact of these actions on future profit. When the manager's action has a relatively small impact on current profit, and a relatively large impact on future profit, franchising may be the preferred organizational form. Fourth, because the ownership form is determined by the outlet's characteristics, it is possible that we will indeed observe that the manager retains a smaller fraction of current profit in company-owned outlets than in franchised outlets. Thus while I do not assume that employees necessarily receive compensation that is less incentive-based than that received by franchisees, observations that this pattern is true in fact are consistent with my model.

There is a fairly large literature on the incentive effects of franchising. Much of this literature has focused on the effects of franchisee moral hazard. Rubin (1978) first suggested that the terms of a franchise contract, as well as the choice of franchising or company-ownership, can be explained by considerations of double moral hazard. Lal (1990) presents a model of double moral hazard in franchising with a focus on deriving the terms of the optimal franchise contract. Bhattacharyya and Lafontaine (1992) use a double moral hazard model to explain several stylized facts about franchise contracts: the simple payment rules used, and the uniformity across different franchisees of the same franchisor.⁹ There is also extensive empirical support for the

⁹ Double moral hazard is also seen in a wide variety of other economic situations. Holmstrom (1982) analyzes the double moral hazard problem inherent in team production. See Cooper and Ross (1988) and Dybvig and Lutz (1993) for models of dynamic moral hazard applied to optimal contract terms for manufacturer warranties.

importance of moral hazard in explaining franchising contracts. Brickley and Dark (1987), Norton (1988), Krueger (1991), Brickley et al. (1991), Sen (1993) and Lafontaine (1992a) all present empirical evidence suggesting that there is local manager moral hazard. Lafontaine (1992a) and Sen (1993) also find support for the presence of central firm moral hazard.

Three papers in the franchising literature have pointed to the importance of ownership in determining incentives under employment as well as franchise contracts. Brickley, Dark, and Weisbach (1991) and Gallini and Lutz (1992) both note that because the franchisee owns the local unit he has the right to sell the unit when he wishes to leave active management of the unit.¹⁰ Employee managers, in contrast, have no such rights of sale. If the sale price reflects the unit's future profit stream, then franchising provides better incentives for the unit manager. Brickley, Dark, and Weisbach (1991) do not present a formal model of this idea. Gallini and Lutz (1992) do have a formal model, but do not consider the effect of franchisor moral hazard. Under their assumptions, franchising is always the more efficient organizational form.

Mathewson and Winter (1994) consider the incentive effect of territorial restrictions in franchise contracts. An exclusive territory clause gives veto power over expansion within the territory to the franchisee; such a clause gives the franchisee ownership of a particular kind of local asset. If the franchisee has an exclusive territory, then the franchisor must share the incremental profit from a second unit with the franchisee in order to obtain the franchisee's approval for expansion. In a model with double moral hazard, they find that the franchisee will 'own' the local territory when franchisee effort has a large effect on profit. There are two main distinctions between my work and theirs. First, they focus on the ownership of one particular asset, and thus do not consider the choice between franchising and company-ownership of a local unit. Second, they do not consider the possibility that the contract between the franchisor and franchisee may include a sharing rule as well as an assignment of territorial rights.

Alternative theories of franchising that do not rely on incentive advantages include Caves and Murphy (1976), who argue that franchising gives the central firm access to low-cost capital. This capital-cost theory has been criticized (see Rubin (1978)); it is difficult to justify why a franchisee faces lower capital costs than the (more diversified) franchisor. This assumption can be justified by assuming incentive problems, but then franchising is fundamentally profitable for incentive reasons, just as in the incentive theory. Gentry (1993) develops an alternative capital market theory. He argues that

¹⁰ This right is frequently restricted by the terms of the franchise contract, but is rarely eliminated.

franchising may be used to raise capital, in preference to an equity issue, because new equity dilutes the claim of the firm's current shareholders on the firm's intangible assets, its trademark.

The choice of whether to franchise or company-own a local outlet is one example of the broader issue of the scope of a firm's (vertical) activities.¹¹ Coase (1937) introduced the question of the boundaries of a firm into economics, and developed the importance of the transactions costs of writing contracts in explaining the boundary between a firm and the market. In assuming that contracts between the central firm and the unit manager are necessarily incomplete, I am assuming that the transaction costs of reaching a complete contract are extremely high. Grossman and Hart (1986) and Hart (1988) present a model of vertical integration that highlights the importance of asset ownership when contracts are necessarily incomplete because of high transaction costs. According to this view, given any incomplete contract, asset ownership allocates the residual rights of control, the rights to specify an asset's use under contingencies not covered in the contract. The allocation of these rights determines the outcome of any contract renegotiation that follows on the realization of an unforeseen contingency. By determining the expected outcome of these renegotiations, asset ownership has an important effect on incentives; the costs and benefits of ownership can then be understood in terms of these incentive effects. This paper applies the same general structure to the analysis of franchising. I focus on the ownership of two assets, the local unit and the trademark, and assume that the residual rights of control over the unit include the right to sell the unit and/or the right to operate the unit under a different trademark.

In the next section, I briefly review some of the key features commonly observed in franchise contracts. Section 3 then presents a simple model of franchising and company-ownership that incorporates those key features. Section 4 analyzes the model, focusing on incentives and the optimal contract terms under either type of ownership. Section 5 considers the conditions under which franchising is more profitable than company-ownership by means of two examples, and Section 6 concludes.

2. Characteristics of franchise contracts

Before presenting the model, it will be useful to survey in more detail some of the key features of franchise contracts. I will assume in my model that franchise as well as employment contracts are incomplete in ways that

¹¹ See Joskow (1985) and Perry (1989) for excellent surveys on the determinants of vertical integration.

leave scope for moral hazard on the part of both the local manager and the central firm. I also assume that the contracts are incomplete in allocating the entire stream of future unit profits. Franchisees, because of their ownership of local assets, capture more of the future profits from a local unit than do employee managers. Finally, I assume that the division of current period profits between the central firm and the local manager is governed by a simple linear rule. All of these assumptions are consistent with observed franchise contracts.

As Hadfield (1990) observes, we have relatively little information about full terms of franchise contracts. Some information, especially about the payment terms, is disclosed to potential franchisees under Federal Trade Commission regulations. Information from the UFOCs (Uniform Franchise Offering Circulars) is available to the general public through surveys published by several sources. However, these disclosure documents do not contain the full terms of the contracts. Scattered additional information is available on the terms of current franchise contracts. The most detailed survey of contract terms, Ozanne and Hunt (1971), examines terms used in franchise contracts during the late 1960's and early 1970's in the fast-food industry.

While the information is limited, several striking facts do emerge about the terms of franchise contracts.¹² First, these contracts do not completely specify the actions to be taken by both the franchisee and the franchisor. The franchisee's actions are often specified via a detailed operations manual. According to Ozanne and Hunt (1971), 61% of the contracts they surveyed specify that the franchisee will receive such a manual; 38% of the contracts give the franchisor the right to change the manual at any time. The contract generally requires that the franchise comply with the manual at all times: 16% of contracts had an explicit clause to that effect, and 76% of contracts required the franchisee to agree that any violation of the contract is grounds for termination of the franchise agreement.¹³ However, the franchisee still retains scope for moral hazard. Because the local unit is at some distance from the central firm, deviations from the operating manual will not be immediately detected. Also, the manual cannot specify all actions to be taken by the local manager; for example, the manual cannot tell the manager who to hire among several applicants for a job.

The franchise contract also contains terms specifying the actions that will be taken by the central firm. These actions can be divided into two categories: initial assistance provided to new franchisees and ongoing investments in the trademark and franchisee training. Ozanne and Hunt (1971) found

¹² Hadfield (1990) presents a more complete survey of the available evidence.

¹³ If not otherwise noted, all numbers in this section are drawn from Ozanne and Hunt (1971).

that 77% of contracts promised formal start-up training for franchisees. An alternative to formal training is on-the-job training; 34% of contracts promised on-the-job training for new franchisees. The franchisor also frequently promises assistance in opening the local unit (45% of contracts). This suggests that most contracts obligate the franchisor to provide initial assistance to the franchisee. However, Ozanne and Hunt also report that many franchisees did not receive the promised assistance, evidence of franchisor moral hazard. Only 47% of franchisees reported receiving promised formal training, while 71% received promised on-the-job training.

The contracts do little to obligate the central firm to make continuing investments in the success of the local unit. Only 28% of contracts surveyed obligated the franchisor to advertise. Even when royalties are collected from the franchisee to pay for advertising, the contract does not specify the nature, extent, or effectiveness of the franchisor's advertising decisions. While 45% of contracts offer continuing management support, 75% of these clauses allow the franchisor discretion in deciding whether or not such support is necessary. Mathewson and Winter (1985) report that out of the 9 Canadian franchise contracts they examined, none guaranteed advertising effectiveness. Based on this evidence, Hadfield concludes that 'There are essentially no clauses creating any obligation for the franchisor to develop the system, to continue in operation, to continue to advertise, or to maintain the trademark. (p. 944)' ¹⁴

Like franchise contracts, employment contracts leave significant scope for moral hazard on part of both the local manager and the central firm. The requirements imposed on franchisees by typical contract clauses are similar to the requirements placed on employee managers under employment law. ¹⁵ Despite these requirements, as long as the employer's monitoring is imperfect, employee managers will retain significant discretion over some decisions. Employers generally retain broad rights to choose their level of investment in the success of local units.

¹⁴ Franchise contracts also do not specify precise levels of another central firm action that contributes to local success. In a chain with several outlets and customers who patronize more than one outlet, profits at any one outlet depend on consumer's experiences at other outlets. Thus any one local manager is tempted to free ride on the quality produced by other managers. The central firm can reduce the extent of this free riding by frequent monitoring; therefore, monitoring is another action taken by the franchisor that affects the success of local units. Franchise contracts give the franchisor the right to monitor the local unit, but do not specify the frequency or extent of monitoring. In this paper, I do not consider the role of monitoring in preventing free-riding because I focus on the relationship between the central firm and a single local outlet. I discuss the issues created when there are several stores in the chain in more detail in the conclusion. See Rubin (1978) for the seminal discussion of the importance of monitoring in understanding the structure of retail chains.

¹⁵ See Masten (1988) for a discussion of these requirements.

Franchise contracts contain clauses specifying the ownership of the various assets used in production. Generally, the franchisor owns the trademark and the franchisee owns the local assets, including buildings and equipment. Ozanne and Hunt (1971) report that 77% of the contracts in their survey included clauses specifying that the franchisee has no ownership rights in the chain's trademark. They also found that in 60% of the cases they surveyed, the franchisee controls the land and buildings used at the local unit: this means that the franchisee either owns these assets or leases them directly from the non-franchisor owner. In 14 out of the 19 U.K. franchise chains surveyed by Dnes (1992), the franchisee is free to own his own site or to lease the land and buildings from someone other than the franchisor. In the remaining cases, the franchisee leases (or sub-leases) the land and buildings from the franchisor.¹⁶ Even when the franchisee does not own the local land and buildings, he does own the other physical assets used at the local unit.

Ownership of the assets gives the franchisee the right to make decisions about their use. The most important of these rights, for incentive purposes, is the franchisee's right to control the assets when the franchise relationship ends. The relationship ends in two possible ways: franchise contracts are of fixed duration, and can be terminated early due to material breach. The contracts generally contain specific clauses that limit the franchisee's control over the local assets when the relationship ends. However, these clauses do not eliminate the franchisee's control over the assets he owns. Ozanne and Hunt (1971) found that 44% of contracts they surveyed included clauses giving the franchisor an option to purchase the land, building, and/or equipment from the franchisee upon termination or expiration of the franchise contract. Of these clauses, 50% set the price of their assets at their depreciated value, 20% left the price up to negotiation or arbitration of some sort, and 30% set the price of the assets at market value. Clauses requiring the franchisor's approval before the unit can be sold to a new franchisee also limit, but do not eliminate, ownership rights; 83% of the contracts included such a clause, and 46% of the contracts gave the franchisor the right of first refusal in any sale. Kaufmann and Lafontaine (1994) report that McDonald's constrains the sales of franchises in several ways. McDonald's restricts the possible buyers to a pre-approved list of individuals, and allows only one buyer from the list to make an offer to the seller at any time.

¹⁶ In the United States, this policy seems to be followed mainly by McDonald's. In 1986, McDonald's owned the land and buildings used by 60% of its outlets, both company-owned and franchised. All but two of the remaining 40% of outlets used land and buildings leased by McDonald's and subleased to franchisees. However, even in the case of McDonald's, franchisees own the equipment used within the unit. See Love (1986) for more on McDonald's real estate policies throughout the chain's history. See Kaufmann and Lafontaine (1994) for estimates of the capital investments made by McDonald's franchisees in their equipment.

If the franchisee chooses not to sell his assets, his options are limited. 47% of the contracts surveyed by Ozanne and Hunt require that the franchisee cease local operation upon termination in some way: 33% of these clauses required that the former franchisee discontinue use of the franchisor's trademarks, and 47% of the clauses require alterations in the local building before it can be used for another purpose.¹⁷ The survey also found that 67% of contracts had noncompetition clauses of some kind, restricting the former franchisees ability to pursue the same line of business.

Despite these limitations on the franchisee's property rights, there is extensive evidence that these rights are not eliminated. There is an active market in franchises: 2,628 franchises requested franchisor approval of transfers in ownership during 1986, and all but 139 were approved. (USDOC (1986) p. 13). Ozanne and Hunt (1971) found that 64% of the contracts they surveyed gave a franchisee's heirs the right to inherit the franchise upon the franchisee's death. Kaufmann and Lafontaine (1994) document an active market in McDonald's franchises, and conclude that franchisees are able to sell their franchises at prices that allow them to capture at least part of the unit's future profit stream. All of this suggests that franchisees have far more extensive rights to control the local assets than those held by an employee manager. Employees do not have long term contracts (and most franchise contracts last at least 10 years).¹⁸ While their contracts may be renewed, employees have no right to sell their job to someone else or to pass it on to their heirs. Thus there is no mechanism for an employee to capture the effect his actions have on a unit's profitability after he leaves the active management of the local unit.

Finally, franchise contracts display remarkable uniformity in the division of unit profits between the franchisor and the franchisee during the terms of the franchise contract. Here we have relatively good evidence on current contract terms, since these terms must be reported to potential franchisees in the franchisor disclosure documents. Bhattacharyya and Lafontaine (1992) survey the current evidence on contract terms and conclude that the most common contract form calls for the franchisee to make a fixed upfront payment to the franchisor at the beginning of the contract, and to make royalty payments over the life of the contract. Royalties are usually calculated as a fixed percentage of the unit's sales. This is also consistent with the earlier contracts analyzed by Ozanne and Hunt (1971), and with the U.K. contracts considered by Dnes (1992).

¹⁷ The two clauses are similar, since distinctive building design is a form of promotional capital for many chain stores.

¹⁸ USDOC (1986) reports that slightly over 80% of all franchisors active in 1984 offered franchise contracts of at least 10 year's duration.

3. The model

In this section, I present a simple model of the relationship between the central firm and the manager of a local unit. The major features of this model are consistent with the franchise and employment relationships I have just discussed: both parties make contributions to the success of the local unit and double moral hazard is present. Contracts cannot credibly commit the parties to a division of future unit profits; this division is determined by the ownership of the assets used in production. Contracts can commit the two parties to a division of current profits, and I restrict the sharing rule to a simple linear rule like that used in most franchise contracts.

We will focus on the relationship between two parties, the developer of a trademark (individual T) and the manager of a local unit operating under this trademark (individual L). If the local unit is franchised, then T is the franchisor and L is the franchisee. If the unit is company-owned, then T is the employer and L is the employee manager.¹⁹

I assume that the local unit operates for two periods. At the beginning of the first period, a contract is signed that specifies the ownership of the assets used at the local unit and the division of first period returns. T and L then simultaneously choose actions; these actions will determine the unit's profits in each period. The unit's first period profit is then realized and divided between the two parties in accordance with the terms of their contract. At the beginning of the second period, T and L must reach some agreement about the division of the second-period profits. This agreement is determined by the ownership of the local assets. Once the agreement is reached, no further actions are necessary. The second period profit is realized and divided according to the terms of the agreement.

The profitability of a local unit depends on the actions of both parties. Assume that T chooses an action $t \in [0, \infty)$ at a cost or disutility $c(t)$ such that $c_t > 0$, $c_{tt} > 0$, and $c_{ttt} \geq 0$.²⁰ L chooses an action $l \in [0, \infty)$ at a cost or disutility $d(l)$ such that $d_l > 0$, $d_{ll} > 0$, and $d_{lll} \geq 0$. Since these actions are taken during the first period, the disutilities are incurred during the first period as well. We can think of action t as representing T's effort to enhance the firm's trademark by advertising, monitoring the quality of other outlets with the same trademark, and developing new lines of business. This suggests that T's action works to increase demand for the local unit's product or

¹⁹ A third alternative in this bilateral model would be to have L own the trademark developed by T's efforts. T would then be an advertising agency or business consultant. I do not explore this possibility here.

²⁰ All subscripts denote partial derivatives.

products; however, it is also possible that T's action reduces the cost of operating the local unit through the development of more efficient production methods or negotiating low prices with suppliers. Action l is L's effort to improve the profitability of the local unit. L might affect demand at its unit via advertising and high levels of service and product quality. L could also lower its costs through careful monitoring of unit employees.

Neither action is observable, so that L cannot directly observe t and T cannot directly observe l ; more importantly, I also assume that the actions t and l cannot be verified by third parties. Therefore the contract cannot specify levels of either action.

The profits of the local unit are realized over two periods. The unit's first period profits are $\Pi(t, l)$. Assume that $\Pi_t > 0$, $\Pi_l > 0$, $\Pi_{tt} \leq 0$, $\Pi_{ll} \leq 0$, $\Pi_{tl} \leq 0$ and $\Pi_{ll} \leq 0$; either action increases the unit's profitability, but at a decreasing rate. We will also assume that $\Pi_{tl} = 0$; an increase in the effort taken by one party does not decrease the marginal payoff to the other party's effort. This assumption simplifies the later analysis considerably, since it helps to insure that each party's action is a function only of the terms of the contract, and is independent of the other party's action.²¹

The unit's second period profits are $V(t, l)$.²² I assume that $V_t > 0$, $V_l > 0$, $V_{tt} \leq 0$, $V_{ll} \leq 0$, $V_{tl} \leq 0$ and $V_{ll} \leq 0$, and also assume that $V_{tl} = 0$.²³

This second period can be thought of as the time after the original franchise or employment contract between T and L expires. In this case, $V(t, l)$ can be viewed as the discounted present value of the stream of profits

²¹ The major results can still hold if $\Pi_{tl} > 0$, so that the two kinds of effort are complements. Additional assumptions are necessary, to guarantee that there is a unique Nash Equilibrium in effort choice, and that the equilibrium effort levels are continuous and piecewise differentiable functions of the royalty rate. A different tractable assumption would be to change the timing of the actions, and have one party choose an action before the other party.

²² I could relax the assumption made here that T and L take no second-period actions. In this case, the bargaining at the beginning of the second period would not just be over the division of a fixed profit, but over the incentive terms of the second-period contract. As long as first-period actions affect second-period profits, the major conclusions I reach about the determinants of the optimal ownership structure would not change.

²³ In models with one-sided moral hazard, such as standard principal-agent models, it is usual to assume that observable and verifiable profits are a function of unobservable action and an unobservable random variable. This random element makes it impossible to infer the unobservable effort level from the observable profit level. The random element is not necessary here; with double moral hazard, a third party charged with enforcing the contract cannot tell whether low profits are due to low effort by T or low effort by L. Thus the contract cannot specify levels of either action. Of course, because each individual knows his own action, he can infer the other's action from the observed profit.

after the contract expires, with $\Pi(t, l)$ as the discounted present value of profits during the contract, assuming that demand and costs are stationary over time.²⁴

Thus total surplus over the two periods is

$$\Pi(t, l) + V(t, l) - c(t) - d(l). \quad (1)$$

The efficient levels of effort, t^* and l^* , then maximize total surplus. Under our assumptions, total surplus is a concave function of t and l , and the efficient effort levels satisfy the first-order conditions

$$\Pi_t + V_t - c_t = 0 \quad (2)$$

and

$$\Pi_l + V_l - d_l = 0. \quad (3)$$

Of course, these efficient levels are not achievable, since effort levels cannot be specified in the contract. Instead, given a contract, T will choose his actions to maximize his private return; L will do likewise.

We will consider two kinds of contracts, franchise contracts and employment contracts. Each specifies a payment rule and the ownership of the local unit. Under a franchise contract, L (now the franchisee) pays T (the franchisor) a fixed fee of F and a fraction r of the unit's first period profit.²⁵

The contract terms specify the division of the first period profit. The division of the second period profit depends on the ownership of the local assets. In a franchise contract, L owns these assets. Because the division of second period returns cannot be specified in the franchise contract, T and L must bargain at the beginning of the second period over the division of second-period profits. Both parties know t and l at the beginning of the second period: each observes his own action, of course, and can infer the other's action from the first period profits. Since L owns the local store, it is free to operate the store without the trademark, or to sell the store to another franchisee. Let $Z(l)$ be the second period return to the franchisee

²⁴ As I have already noted, most franchise contracts extend for more than 10 years. While there is little information available about the average job tenure of an employed manager, anecdotal evidence suggests that employee managers rarely remain with a local unit for more than a few years. Thus the "first period" is longer for a franchisee than for the manager of a company-owned store. We can ignore this asymmetry in contract duration for simplicity. Including it in the model would strengthen the franchisee's incentives, and thus would not change the basic thrust of the argument.

²⁵ This form for the profit-sharing rule is consistent with observed franchise contracts, as I have already noted. What is more, this assumption can actually be made without loss of generality; see Bhattacharyya and Lafontaine (1992) for a proof in a different model that can be adopted for this situation as well.

from these options.²⁶ Since T owns the trademark, it is free to transfer the trademark to a new store; let $Y(t)$ be T's second period return if it does so.²⁷ We assume that $V(t, l) > Y(t) + Z(l)$; otherwise there is no gain to L and T agreeing in the second period to operate L's store with T's trademark. We also assume that $Y_t \geq 0$, $Z_l \geq 0$, $Y_{tt} \leq 0$, and $Z_{ll} \leq 0$.²⁸

Since both parties know $V(t, l)$, $Y(t)$, and $Z(l)$, we can reasonably assume that the Nash bargaining solution obtains.²⁹ Under Nash bargaining, T's second period return is $1/2(V(t, l) + Y(t) - Z(l))$, and L's second period return is $1/2(V(t, l) + Z(l) - Y(t))$.

We will assume that both T and L are risk neutral. Therefore, T's total return under the franchise contract (r , F) is

$$r\Pi(t, l) + \frac{V(t, l) + Y(t) - Z(l)}{2} + F - c(t), \quad (4)$$

and L's total return under the same contract is

$$(1 - r)\Pi(t, l) + \frac{V(t, l) + Z(l) - Y(t)}{2} - F - d(l). \quad (5)$$

We can then consider how the two actions l and t are chosen under a franchise contract. The franchisee L chooses l to maximize his return, which is given by (5). His optimal choice of l is described by the first-order condition³⁰

$$(1 - r)\Pi_l(t, l) + \frac{V_l(t, l)}{2} + \frac{Z_l(l)}{2} - d_l(l) = 0. \quad (6)$$

²⁶ As we have seen, franchise contracts can impose significant restrictions on the franchisee's right to operate the unit without the franchised trademark as well as his right to sell the local unit. However, these restrictions can be thought of as reducing $Z(l)$. Therefore, restrictions on the franchisee's ownership rights will only eliminate the effect of ownership on incentives when they allow the franchisor to capture the full value of the franchisee's efforts upon contract termination. As long as the franchisor has to strike a bargain with the franchisee to achieve a second period profit of $V(t, l)$, franchising and company-ownership will have different incentive effects.

²⁷ Note that the franchisor is using the trademark on only a single store at any one time. While most franchisors do franchise multiple units, considering several stores is beyond the scope of this model.

²⁸ The assumption that each party's reservation payoff is independent of the other party's action could be relaxed, and is made primarily for simplicity.

²⁹ Any other fixed division of second-period surplus can be assumed instead. For example, we could assume that T is able to extract all second-period surplus. While changing the division of second-period gains to trade will change some of the later results, it will not alter the basic thrust of the argument. Even if T extracts all second-period surplus, L will receive $Z(l)$ in the second period; hence L's incentives will depend on returns over both periods.

³⁰ Given our assumptions on Π , V , Z , and d , the franchisee's return is a concave function of l , and the second order condition for maximization is satisfied.

Let $l^F(r)$ be the choice of l that satisfies this condition. Differentiating (6) with respect to r , we have

$$l_r^F(r) = \frac{\Pi_l}{(1-r)\Pi_{ll} + \frac{1}{2}(V_{ll} + Z_{ll}) - d_{ll}} < 0. \quad (7)$$

The franchisee's effort is therefore decreasing in the royalty rate r .

Now let us consider T's choice of effort as a franchisor. The franchisor's expected return over the two periods is given by (4). His level of effort t is chosen to maximize this return; therefore his optimal effort level, $t^F(r)$, is given by ³¹

$$r\Pi_t(t, l) + \frac{V_t(t, l)}{2} + \frac{Y_t(t, l)}{2} - c_t(t) = 0. \quad (8)$$

Differentiating this with respect to r , we see immediately that

$$t_r^F(r) = \frac{-\Pi_t}{r\Pi_{tt} + \frac{1}{2}(V_{tt} + Y_{tt}) - c_{tt}} > 0. \quad (9)$$

Thus we know that increases in the royalty increase the effort taken by the franchisor. Of course, if $V_t < Y_t$, then for royalty rates close to 1, the franchisor will take an inefficiently high level of effort. He does so because high effort levels have a large impact on his bargaining position at the beginning of the second period. Note that even when the franchisor's action has a large effect on his bargaining position, for small levels of r he will choose an inefficiently low level of effort.

We can now consider whether or not there is any franchise contract that will resolve the double moral hazard problem and result in the first-best effort levels, t^* and l^* .

Proposition 1. If $V_t \geq Y_t$, then there does not exist any franchise contract resulting in the first-best effort levels. If $V_t < Y_t$ and $V_l \geq Z_l$, then such a contract may exist; it includes $r < 0$. If $V_t < Y_t$ and $V_l < Z_l$, then there also may exist such a contract, with $0 < r < 1$.

Proof: Given $V_t \geq Y_t$, we know that $t^F(r) < t^*$ for any $r < 1$. If $V_l \geq Z_l$, we know that $l^F(r) < l^*$ for any $r \geq 0$; therefore there is no r such that both actions are at their first-best levels. If $V_l < Z_l$, then $l^F(r) = l^*$ for some $0 < r < 1$; $l^F(1) > l^*$, and so there is again no r such that both actions are at their first-best levels.

When $V_t < Y_t$, then there exists $\hat{r} < 1$ such that $t^F(\hat{r}) = t^*$. If $V_l \geq Z_l$, then there exists $\bar{r} < 0$ such that $l^F(\bar{r}) = l^*$. If $\hat{r} = \bar{r}$, the first-best outcome is

³¹ Assumptions on Π , V , Y and c guarantee that the second order condition is satisfied.

achievable with a franchise contract with $r = \hat{r} < 0$. A similar argument shows that when $V_l < Z_l$, the first best outcome may be achievable with a royalty rate $0 < r < 1$.

The implication is that the first-best level of profit is not achievable in a franchised unit when $V_l \geq Y_l$. This is the standard result in a model with double moral hazard; payment schemes that improve one party's incentives damage the other party's incentives. However, if $V_l < Y_l$, then the first-best level of profit may be achievable with a franchise contract. In this case, the marginal effect of the franchisor's effort on future profits is smaller when the franchisee's unit is operated under the trademark than when the trademark is used alone. The franchisor's effort has a strong effect on his bargaining position at the beginning of the second period, and therefore the franchisor may choose the efficient effort level t^* even when $r < 1$. Thus payment schemes that enhance the franchisee's incentives (those with $r < 1$) may not damage the franchisor's incentives.

We can now turn to employment contracts. I will model an employment contract under the assumptions and notation already developed for a franchise contract. With this notation, L pays T a fixed fee of $F \geq 0$ as well as a fraction r of the unit's first-period profits. Interpreting this as an employment contract, L (the employee manager of a company-owned unit) receives from T (the employer) a (negative) payment of $-F$ and a fraction $(1 - r)$ of the unit's first-period profit as a bonus. Again, returns to T and L in the second period are determined by the ownership of the local unit. Since T owns the local unit, he captures the entire second period profit $V(t, l)$. Either T operates the local unit without L's assistance, or else T pays L his reservation wage, which we can normalize to zero. Thus the company captures all second-period return to the efforts of both parties.

We can now examine how l and t are chosen under the employment contract (r, F) . In this case, given the wage $-F$ and the bonus share $(1 - r)$, the local manager L will choose l to maximize his expected return, which is

$$(1 - r)\Pi(t, l) - F - d(l). \quad (10)$$

Since the employee manager does not expect to receive any second-period return for his first-period effort, he ignores the effect of his choice of l on second-period profits. His optimal choice of l is described by the first-order condition³²

$$(1 - r)\Pi_l(t, l) - d_l(l) = 0. \quad (11)$$

Let $l^{CO}(r)$ be the choice of l that satisfies this condition. We can see immediately that for any $r \geq 0$, $l^{CO}(r) < l^*$. The employee manager takes an

³² Given our assumptions on Π and d , the second order condition for maximization is satisfied.

inefficiently low level of effort in running the store. This is true no matter how the employee manager is compensated, as long as his bonus is not higher than the store's first-period profit. Even when the employee manager receives the store's entire first-period profit as a bonus payment, he is still not adequately compensated for the effect of his effort on future profits. Differentiating (11) with respect to r , we see that

$$l_r^{CO}(r) = \frac{\Pi_l}{(1-r)\Pi_{ll} - d_{ll}} < 0. \quad (12)$$

Increases in the royalty rate (decreases in the employee's bonus rate) result in decreased effort by the employee manager.

Now let us consider T's choice of effort when he owns the local unit. The owner's expected return over the two periods is

$$r\Pi(t, l) + F - c(t) + V(t, l). \quad (13)$$

His level of effort t is chosen to maximize this return; therefore his optimal effort level, $t^{CO}(r)$, is given by ³³

$$r\Pi_t(t, l) + V_t(t, l) - c_t(t) = 0. \quad (14)$$

Note that $t^{CO}(r) < t^*$ for any $r < 1$; if the owner shares any fraction of the first-period return with the manager, the owner has an incentive to take an inefficiently low level of effort.

Differentiating (14) with respect to r tells us how the owner's incentives change with changes in the bonus rate:

$$t_r^{CO}(r) = \frac{-\Pi_t}{r\Pi_{tt} + V_{tt} - c_{tt}} > 0. \quad (15)$$

Increases in the royalty (decreases in the employee's bonus rate) increase the central firm's incentive to take effort when it owns the local unit.

We can immediately see that

Proposition 2: The first-best effort levels t^ and l^* are not achievable in a company-owned store.*

Proof: As we have established, $l^{CO}(r) < l^*$ for any $r \geq 0$, while $t^{CO}(r) < t^*$ for any $r < 1$; thus there is no employment contract that can achieve the first-best outcome. ■

Assuming that the first-best outcome cannot be achieved with a franchise contract, only a second-best outcome is possible. Whether or not franchising

³³ Our assumptions on Π , V , and c guarantee that the second order condition is satisfied.

is more profitable than company ownership will then depend on which ownership form comes closest to achieving the first-best outcome.

We can compare the incentive effects of the two ownership forms for a partial answer to this question. Compare the actions chosen by both the central firm and the local manager when the unit is franchised with a royalty rate of r to the actions chosen by both when the unit is company owned with a bonus rate of $1 - r$.

Proposition 3: For any employment contract (r, F) and the equivalent franchise contract (r, F) , $l^{CO}(r)$ is smaller than $l^F(r)$. Also, $t^{CO}(r)$ is smaller than $t^F(r)$ if and only if $V_i > Y_i$.

Proof: The first statement is obvious by comparison of (6) with (11). Similarly, the second statement follows directly by comparing (14) and (8). ■

This result explains how the effects of any profit sharing rule are determined by the underlying ownership structure. The advantage of franchising over company ownership is that it provides relatively better incentives for the manager of the local unit. These strong incentives may come at a cost: when $V_i > Y_i$, franchising provides weaker incentives for the central firm than are possible under company-ownership. If the marginal effect of t on second period profits is larger when the existing local unit is operated under the trademark than when a new unit is opened, then franchising weakens the franchisor's effort incentives. Of course, this may not be the case; if the marginal effect of t on second period effort is higher for a new second unit, then by increasing its effort in the first period the franchisor improves its bargaining position in the second period. The net result is then that franchising provides better incentives than company-ownership for both types of effort.

This result focuses on how the two ownership forms imply different incentives given some specified royalty rate r . Incentives under either ownership form are also affected by the profit sharing rule specified in the contract. We turn next to the optimal sharing rule under each type of ownership.

4. Optimal contract terms

We can now consider the terms of the optimal contract under each ownership structure. A franchise contract has two components, the franchise fee F and the royalty rate r . These terms are set by the franchisor at the beginning of the first period, and can be accepted or rejected by the franchisee. If the franchise contract is accepted, the actions of both the franchisor and the franchisee are both determined by the specified r , as we have already seen. If the contract is rejected, the franchisee receives some

reservation value, which we can without loss of generality normalize to zero.³⁴

The franchisor will then choose F and r to maximize his total return (4), subject to three constraints. The first two constraints are the incentive compatibility constraints:

$$t = t^F(r)$$

and

$$l = l^F(r).$$

The third constraint is the franchisee's individual rationality constraint:

$$(1-r)\Pi(t,l) + \frac{V(t,l) + Z(l) - Y(t)}{2} - F - d(l) \geq 0. \quad (16)$$

Clearly, the third constraint will bind at the optimum: the franchisor will not maximize its return by leaving any surplus with the franchisee.³⁵ Therefore, the franchisor's problem is to choose r to maximize

$$\Pi(t^F(r), l^F(r)) + V(t^F(r), l^F(r)) - c(t^F(r)) - d(l^F(r)). \quad (17)$$

The assumptions we have already made are enough to ensure that this maximand is a concave function of r . The optimal bonus percentage then satisfies

$$(\Pi_t + V_t - c_t)t_r^F(r) + (\Pi_l + V_l - d_l)l_r^F(r) = 0, \quad (18)$$

where for simplicity I am suppressing the dependence on $t^F(r)$ and $l^F(r)$. Substituting in for (6) and (8), and rewriting gives the optimal franchise royalty r^F :

$$r^F = \frac{(\Pi_t + \frac{1}{2}(V_t - Y_t))t_r^F + (\frac{1}{2}(V_l - Z_l))l_r^F}{\Pi_t t_r^F - \Pi_l l_r^F}. \quad (19)$$

Note that the optimal royalty may be negative, or greater than 1.³⁶

³⁴ This assumption is best motivated by assuming that there is a competitive market of potential franchisees; then F will leave each franchisee indifferent between accepting the franchise contract and pursuing an alternative activity. As we have already determined, F does not affect either the franchisor's incentives or the franchisee's incentives; it serves only to divide the total surplus between the two parties. Other rules for the division of the surplus are possible.

³⁵ I am of course assuming here that the franchisee does not face capital constraints.

³⁶ As I have already noted, a negative royalty rate means that the franchisee not only retains all of the unit's first-period profit, but also receives a bonus from the franchisor. This bonus can be funded through two sources: the franchisor's share of second-period profits, and the franchise fee. A royalty rate greater than 1 means that the franchisee pays more than 100% of the first-period return to the franchisor. This can be funded through the franchisee's share of second-period profits.

The optimal employment contract can be derived in a similar fashion. The central firm makes an offer of an employment contract to the employee manager, who can accept or reject. This employment contract has the same two components as a franchise contract: here F can be interpreted as a fixed payment of $-F$ from the employer to the employee, and r can be interpreted as a bonus percentage for the employee of $(1 - r)$. The manager will accept any contract yielding an expected return equal to his reservation wage. The employer maximizes his profits by offering the employment contract that maximizes his return, (13), subject to three constraints:

$$t = t^{CO}(r),$$

$$l = l^{CO}(r),$$

and the employee's participation constraint,

$$(1 - r)\Pi(t, l) - F - d(l) \geq 0. \quad (20)$$

We can again assume that the employer maximizes his profit by using F to extract all surplus from the employee, so that the third constraint is binding at the optimum. Thus the employer chooses r to maximize

$$\Pi(t^{CO}(r), l^{CO}(r)) + V(t^{CO}(r), l^{CO}(r)) - c(t^{CO}(r)) - d(l^{CO}(r)). \quad (21)$$

Our assumptions are sufficient to ensure that this function is concave in r . Therefore, the optimal level of r then satisfies

$$(\Pi_t + V_t - c_t)t_r^{CO}(r) + (\Pi_l + V_l - d_l)l_r^{CO}(r) = 0. \quad (22)$$

Substituting in for (11) and (14) and rewriting defines r^{CO} , the optimal royalty when the unit is company owned:

$$r^{CO} = \frac{\Pi_t t_r^{CO} + V_l l_r^{CO}}{\Pi_t t_r^{CO} - \Pi_l l_r^{CO}} \quad (23)$$

It then follows directly that $r^{CO} > 0$ if and only if $\Pi_t t_r^{CO} \geq -V_l l_r^{CO}$, and that $r^{CO} < 1$.

There are a number of remaining questions to be answered. In particular, we want to derive conditions under which the best franchise contract is more profitable for the central firm than the best employment contract. We can explore this –as well as additional issues –in more detail by making additional assumptions about functional forms. The next section presents two examples that illustrate the effect of different assumptions.

5. Two examples

Two different examples will allow us to examine conditions under which franchising is more profitable than company-ownership. We will also be able to use the examples to compare the profit-sharing rules used in franchised and company-owned outlets. The first example, adapted from Gallini and Lutz (1992), shows the effect of double moral hazard on the relative profitability of the two organizational forms. The second example assumes that all profits are linear in t and l , while the costs of effort are quadratic in their arguments; this will allow us to find closed form solutions for r^F and r^{CO} .

For the first example, assume that the central firm's effort t is observable and verifiable. It can then be specified in the contract, along with r and F . The situation is then one of managerial moral hazard rather than double moral hazard. Given a franchise contract specifying r , F , and t , the franchisee will choose $l^F(r)$ as defined by (6). The franchisor then chooses r , F , and t to maximize

$$r\Pi(t, l^F(r)) + \frac{1}{2}(V(t, l^F(r)) + Y(t) - Z(l^F(r))) + F - c(t),$$

subject to the franchisee's participation constraint (16).

Since this constraint will bind at the optimum, the franchisor will choose t to maximize the unit's total profit over the two periods. Thus the franchisor will choose the first-best level of effort, t^* .

The franchisor also chooses r to maximize total profit over the two periods. It is straightforward to show that

$$r = \frac{Z_l - V_l}{2\Pi_l}.$$

This is the (possibly negative) royalty rate that induces the franchisee to take the first-best level of local effort, l^* . Thus the first-best outcome is achievable under a franchise contract. If a negative royalty is not feasible, then if $V_l > Z_l$ we have that the optimal royalty is $r^F = 0$; the optimal franchise contract leaves all first-period profits with the franchisee. In this case, of course, $l^F(r^F) < l^*$.

If the unit is instead company-owned, the central firm chooses r , F and t to maximize

$$r\Pi(t, l^{CO}(r)) + V(t, l^{CO}(r)) + F - c(t),$$

subject to the employee's participation constraint (20). This constraint binds at the optimum, implying that the central firm chooses r and t to maximize total profits over the two periods. Therefore, the central firm takes the first-best level of effort t^* and chooses

$$r = -\frac{V_l}{\Pi_l},$$

which is the royalty rate that induces an employee manager to take effort l^* . Note that this royalty rate is negative.

If negative royalties are possible, then the profits to the firm are the same whether the unit is company-owned or franchised. However, if negative royalties are not possible, then we have that the optimal employment contract, like the optimal franchise contract, has $r = 0$. It is easy to verify that $l^F(0) > l^{CO}(0)$, since $V_l + Z_l > 0$. Therefore, if negative royalties are not possible, franchising is more profitable than company-ownership.

This example highlights the importance of double moral hazard: in its absence, company-ownership cannot be more profitable than franchising. The next example shows that with double moral hazard, either ownership form may be optimal.

Assume that $\Pi(t, l)$, $V(t, l)$, $Y(t)$, and $Z(l)$ are all linear functions, with $\Pi(t, l) = at + bl$, $V(t, l) = et + fl$, $Y(t) = yt$, $Z(l) = zl$. Also assume that $c(t) = \frac{c}{2}t^2$ and $d(l) = \frac{d}{2}l^2$.

In order to examine the optimal ownership form, we must first find the optimal franchise contract and the optimal employment contract. We can find r^F , the optimal franchise royalty, by substituting the above functions into (19). Therefore, we have

$$r^F = \frac{a\left(a + \frac{e-y}{2}\right)d - b\frac{f-z}{2}c}{a^2d + b^2c}.$$

It follows immediately that

$$t^F(r^F) = \frac{a}{c} \left(\frac{a\left(a + \frac{e-y}{2}\right)d - b\frac{f-z}{2}c}{a^2d + b^2c} \right) + \frac{e+y}{2c},$$

and that

$$l^F(r^F) = \frac{b}{d} \left(\frac{b^2c - a\frac{e-y}{2}d + bc\frac{f-z}{2}}{a^2d + b^2c} \right) + \frac{f+z}{2d}.$$

Next consider company-ownership. From (23), we know that

$$r^{CO} = \frac{a^2d - bfc}{a^2d + b^2c}.$$

Thus we have that

$$t^{CO}(r^{CO}) = \frac{a}{c} \left(\frac{a^2d - bfc}{a^2d + b^2c} \right) + \frac{e}{c},$$

and

$$l^{CO}(r^{CO}) = \frac{b}{d} \left(\frac{bc(b+f)}{a^2d + b^2c} \right).$$

With these results, we can derive the conditions under which franchising is the preferred organizational form. Assume that $V_t \geq Y_t$, so that $e > y$. Then $l^F(r^F) > l^{CO}(r^{CO})$ if and only if $a(f+z) > b(e-y)$, and $t^F(r^F) > t^{CO}(r^{CO})$ if and only if $a(f+z) > b(e-y)$. We know that $l^F(r)$ and $l^{CO}(r)$ are both less than the optimal value l^* for all r ; we also know that $t^F(r)$ and $t^{CO}(r)$ are both less than t^* for all r . Therefore, total profits are higher under franchising than company ownership if (and only if) $a(f+z) > b(e-y)$.

The intuition for this result is straightforward. If the central firm's marginal effect on current profits (a) is large relative to its effect on future returns when the unit is jointly operated (e), then $a/(e-y)$ is large. Similarly, if the local manager's marginal effect on current profits is small relative to his effect on future profits when the unit is jointly operated, $b/(f+z)$ is small. Thus we have that franchising is the optimal organizational form when the effect of the central firm's action is relatively important for current profit and the effect of the manager's action is relatively important for future profits. Since franchising gives the manager stronger future incentives but weaker current incentives (thanks to the higher royalty rate) than company-ownership, the result is consistent with our intuition. We can also see that franchising is the preferred organizational form when z and y are both large. Thus we have that franchising is optimal when the effects of current efforts on profits to the central firm and local manager when they do not jointly produce in the second period are large. Again, this is consistent with our intuition: under franchising, the original contract is renegotiated once it expires. If current action has a large effect on each party's reservation profit in that negotiation, then franchising has strong incentive effects.

We can also compare in this example the sharing rules used in outlets where company ownership is optimal with the rules used in outlets where franchising is optimal. This comparison might be made across outlets of different chains, or within outlets of the same chain that differ in the optimal ownership structure because of differences in local conditions.

We have established that high levels of r strengthen L's incentives at the expense of weakening T's incentives. We also know that as long as $V_t > Y_t$, for any given r franchising weakens T's incentives while strengthening L's. This suggests that the ownership form and the royalty rate r can substitute for each other, and in particular that the optimal sharing rule provides, ceteris paribus, for a higher fraction of profit to be given to T under franchising than under company-ownership. It is easy to show that this is true as long as $e > y$. The optimal sharing rule leaves a larger fraction of first-period profit with the manager when the manager receives less of the

second-period profit, as long as receiving more of the second-period profit does not weaken the central firm's incentives.

However, we cannot conclude that the managers of company-owned stores have first period compensation schemes that are more incentive based than those of franchisees. The central firm will choose whether to company-own or franchise a given unit depending on which ownership scheme is more profitable when combined with its optimal sharing rule. We must compare r^F for those parameter values such that franchising is more profitable to r^{CO} for the parameter values such that company-ownership is more profitable. It is straightforward to show that it is possible for the royalty used at company-owned units to be higher than the royalty used at franchised units.³⁷

6. Conclusion

This paper has presented a definition of franchising that highlights the role of asset ownership. A franchisee owns at least some of the location-specific assets used at his unit; an employee manager does not own any of the assets used in production. When contracts between the manager and the central firm are incomplete, the structure of ownership affects the manager's incentives under any incentive contract. The optimal ownership form then depends on the importance of providing strong incentives for the local manager. When the relationship between the central firm and the local manager is characterized by one-sided managerial moral hazard, franchising is always more profitable than company-ownership. If the central firm also makes non-contractible contributions to the unit, then the relationship is characterized by double moral hazard. In this case, franchising may or may not be more profitable than company ownership. We expect franchising to be likely to be more profitable when the local manager's actions have a relatively small effect on a local unit's current profitability, and a relatively large effect on the unit's future profitability.

This analysis of the advantages of franchising does not rest on any assumptions of innate differences between franchisees and employee managers. While it highlights the importance of incentives, it recognizes that incentives are important for the employee managers of company-owned stores as well as franchisees. Thus it extends the literature on the incentive advantages of franchising to consider seriously the choice of franchising versus company-ownership.

³⁷ This is easy to demonstrate by considering two units that are identical except for a site-specific level of f .

The analysis presented here could be usefully extended in several ways. The most important extension will be to examine the choice of ownership for individual units within a multiple-unit chain. Since we observe franchise relationships almost exclusively in these chains, this extension is crucial for developing empirical predictions. Such a multiple-unit model should have several key features. First, at least some of the central firm's effort creates a public good: improvements in the trademark will increase profits at all stores in the chain. Second, any local manager's action may also affect the profits at other units: this will be the case when a local manager's action affects the value of the trademark. Thus incentives for all parties are interdependent.

Such a model would be especially useful in analysing the use of dual distribution: many chains franchise some outlets while owning others. Any bilateral model can only explain this phenomenon by assuming that different locations have different characteristics. A multiple unit model can determine whether locational differences are necessary for dual distribution, or whether dual distribution can arise as a solution to incentive problems even when all units are *ex ante* identical. Several models of this type already exist. Gallini and Lutz (1992) consider a model in which the central firm has private information about the exogenous value of the trademark. They assume that the franchisee's effort affects the unit's profitability, but that the franchisor cannot affect the value of the trademark. When the franchisees can observe the value of the trademark, franchising is more profitable than company ownership. When the trademark's value is unobservable, the franchisor may company-own some units in the chain in order to signal this value to potential franchisees. This paper does use an ownership-based definition of franchising, but its assumption that the value of the trademark is independent of the franchisor's actions and unobserved by franchisees does not describe many established chains. Michael (1993) considers the choice of franchising or company ownership in a two store chain. He focuses on the role of the franchisor as a central party who can structure incentives to assure coordination between the outlets. He assumes that company-owned stores have exogenously higher costs than franchised stores.

Another possible extension would look at the role of monitoring in maintaining managerial incentives. The central firm may monitor local managers, as well as use incentive contracts, to induce them to take high levels of effort. Monitoring is then one of the actions taken by the central firm. While monitoring reduces the role of profit-sharing in providing managerial incentives, the contract must then give the central firm an incentive to carry out high levels of monitoring. Rubin (1978) first highlighted the importance of monitoring in explaining the structure of retail chains, and especially the use of dual distribution. However, a formal model of the choice of monitoring intensity at different outlets remains for future research.

We might also extend this model to consider the role of capital constraints

in determining ownership form. I have assumed that neither the franchisor nor the franchisee faces any such constraints. Capital constraints on the part of the central firm have been used to justify franchising by several authors; the basic argument is that the central firm franchises because its franchisees face lower costs of capital. Capital constraints on the part of local managers have also been used to explain franchising. The argument is that employee managers will not be able to make large upfront payments; the only way for the central firm to give its manager strong incentives while extracting all rents with a fixed fee is by franchising. If capital constraints of either sort do indeed exist, they would affect the relative profitability of the two different forms of ownership. Incorporating exogenous capital constraints would not be difficult.

A final extension would consider the structure of the incentive contracts. I assumed that incentive payments –the royalty payments –are calculated as a percentage of profit. Franchise royalties are usually calculated as a percentage of revenue; this raises the question of why this form is used, rather than basing the royalty on profit or cost. If only revenues are verifiable, then the answer is clear, since profits and costs are non-contractible. However, an alternative explanation might be that the franchisee's effort primarily determines the unit's costs, while the franchisor's effort primarily determines the unit's demand. In this case, basing the royalty on revenue may be a better resolution to the incentive problem.

References

- Bhattacharyya, S. and F. Lafontaine, 1992, Double-sided moral hazard and the nature of share contracts, mimeo.
- Brickley, J.A. and F.H. Dark, 1987, The choice of organizational form: the case of franchising, *Journal of Financial Economics* 18, 401420.
- Brickley, J.A., F.H. Dark and M.S. Weisbach, 1991, An agency perspective on franchising, *Financial Management*, 2735.
- Coase, R., 1937, The Nature of the Firm, *Economica*, reproduced in: G. Stigler and K. Boulding, eds., *Readings in Price Theory*, (Richard D Irwin, Homewood IL).
- Cooper, R. and T.W. Ross, 1988, An Intertemporal model of warranties, *Canadian Journal of Economics* 21, 7286.
- Dnes, A.W., 1992, *Franchising: a case-study approach*, (Avebury, Aldershot U.K.).
- Dybvig, P.H. and N.A. Lutz, 1993, Warranties, Durability, and Maintenance: Two-Sided Moral Hazard in a Continuous-Time Model, *Review of Economic Studies* 60, 575598.
- Grossman, S.J. and O.D. Hart, 1986, The costs and benefits of ownership: A theory of lateral and vertical integration, *Journal of Political Economy* 94, 691719.
- Hadfield, G.K., 1990, Problematic relations: franchising and the law of incomplete contracts, *Stanford Law Review* 42, 927992.
- Hallowell, R., 1993, Au Bon Pain: the partner manager program: teaching note, Harvard Business School Teaching Note 5-692-090.

- Hart, O.D., 1988, Incomplete contracts and the theory of the firm, *Journal of Law, Economics, and Organization* 4, 119139.
- Holmstrom, B., 1982, Moral hazard in teams, *Bell Journal of Economics* 13, 324340.
- Joskow, P., 1985, Vertical integration and long term contracts, *Journal of Law, Economics, and Organization* 1, 3380.
- Kaufmann, P.J. and F. Lafontaine, 1994, Costs of control: the source of economic rents for McDonald's franchisees, *Journal of Law and Economics*, forthcoming.
- Klein, B., and Saft, L.F., 1985, The law and economics of franchise tying contracts, *Journal of Law and Economics* 28, 354361.
- Krueger, A.B., 1991, Ownership, agency, and wages: an examination of franchising in the fast food industry, *Quarterly Journal of Economics* 106, 75102.
- Lafontaine, F., 1992a, Agency theory and franchising: some empirical results, *RAND Journal of Economics* 23, 263283.
- Lafontaine, F., 1992b, How and why do franchisors do what they do: a survey report, in: P. Kaufmann, ed., *Franchising: passport for growth and world of opportunity*, Sixth Annual Proceedings of the Society of Franchising, International Center for Franchise Studies, University of Nebraska.
- Lal, R., 1990, Improving channel coordination through franchising, *Marketing Science* 9, 299318.
- Love, J.F., 1986, *McDonald's: Behind the Arches*, (Bantam Books, Toronto).
- Masten, S.E., 1988, A legal basis for the firm, *Journal of Law, Economics, and Organization* 4, 181198.
- Mathewson, G.F. and R. Winter, 1985, The economics of franchise contracts, *Journal of Law and Economics* 28, 50326.
- Mathewson, G.F. and R. Winter, 1994, Territorial Restrictions in franchise contracts, *Economic Inquiry* 32, 181192.
- Minkler, A.P., 1992, Why firms franchise: a search cost theory, *Journal of Institutional and Theoretical Economics* 148, 24059.
- Minkler, A.P. and T. Park, 1992, Transaction costs determinants of vertical integration: an empirical investigation using franchising data, University of Connecticut mimeo.
- Norton, S.W., 1988, An empirical look at franchising as an organizational form, *Journal of Business* 61 197217.
- Norton, S.W., 1989, Franchising, labor productivity, and the new institutional economics, *Journal of Institutional and Theoretical Economics* 145, 5796.
- Ozanne, U.B. and S.D. Hunt, 1971, The economic effects of franchising, Report prepared for the Small Business Association, Senate Select Committee on Small Business, 92d Congress, 1st Session.
- Perry, M., 1989, Vertical integration: determinants and effects, in: R. Schmalensee and R. Willig, eds., *Handbook of industrial organization*, (North-Holland, Amsterdam).
- Rubin, P., 1978, The theory of the firm and the structure of the franchise contract, *Journal of Law and Economics* 21, 22333.
- Sasser, W.E. and S.H. Pettway, 1974, Case of Big Mac's pay plans, *Harvard Business Review* 3058.
- Schmidt, T., 1993, A theoretical analysis of franchise sharing and risk, University of New Hampshire mimeo.
- Sen, K., 1993, The use of initial fees and royalties in business format franchising, *Managerial and Decision Economics* 14, 175190.
- Sherman, A., 1991, *Franchising and licensing: two ways to build your business*, American Management Association
- Shimer, S.L., 1993, RKO Warner Video, Inc. incentive compensation plan, Harvard Business School Case Study 9-190-067.
- U.S. Department of Commerce, 1986, *Franchising in the economy 1984 –1986*, prepared by Andrew Kostecka.