3-1-1991

Dual Distribution in Franchising

Nancy Gallini

Nancy A. Lutz

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 973

DUAL DISTRIBUTION IN FRANCHISING

Nancy T. Gallini and Nancy A. Lutz

March 1991

DUAL DISTRIBUTION IN FRANCHISING

Nancy T. Gallini*

and

Nancy A. Lutz**

August, 1990
revised March, 1991

* University of Toronto

** Washington University and Yale University

The authors thank Scott Davis, Phil Dybvig, Thomas Gresik, Francine Lafontaine, Chris Lamoureux, Joe Tracy, Ralph Winter and Brian Wright for helpful discussions and comments. Gallini gratefully acknowledges the Social Science and Humanities Research Council of Canada for financial support.

1. Introduction

Franchising is a widely-used organization form for retail distribution, accounting for one-third of U.S. retail sales.[1] Although franchising is a thriving organizational form[2], the economic success of a franchise, especially of a new product, is not guaranteed. Purchasers of new franchises (franchisees) operate in uncertain markets, commit large sunk investments, and their profits depend on a complex relationship with the seller of the franchise (franchisor).[3]

Consequently, franchise profits are often far below those anticipated: a recent survey indicates that only 47% of franchisees were content with their profits.[4] Although disclosure laws provide some protection for franchisees, in a surprisingly large number of cases franchisees were misled by the franchisor at the time of purchasing the unit.[5] As noted in the franchise trade literature: "there is a kind of glow about franchising that, in the '70's, enabled some franchisors to sell seemingly too-good-to-be-true franchises to unsuspecting purchasers."[6]

In this paper, we focus on problems that arise and institutional remedies adopted when franchisors hold private information on the profitability of the franchise. Before distributing a new product widely, a franchisor often acquires private information on product demand

---

[1] See Caves and Murphy (1976).

[2] The number of franchised businesses have increased 26% since 1981 to nearly 329,000 operations. Sales are $576 billion, an increase of 60% since 1981.

[3] For example, see Hadfield (1990) for an thorough discussion of contractual issues in franchising.

[4] "Free Yourself From Servitude", by M. Brandenberg, Accountancy, October, 1986, pp. 82-86.

[5] Hadfield (1990) reports that in Consumer's Petroleum Co. v. Texaco, Inc. 804 F.2d 907 (6th Cir.1986), a gasoline franchisor assured his franchisee that his business was secure, even though the franchisor had been advised to terminate distribution in the franchisee's area, and eventually did. Similarly, Burger Chef promised high growth while the system was known to be in decline. The courts held that claims of future growth "could only be characterized as puffing...although [the franchisees] may in fact have relied on the representations of the defendants, they had no right to rely on them." In another case, two franchisees claimed the Los Angeles-based Devlyn Corporation misrepresented the potential sales and profits of the King Bear Auto Service Center which eventually filed for liquidation. The saga of the Pop-Ins maid service is full of empty promises of international expansion, inflated reports on the number of franchises sold, and false claims of prestigious franchise awards, and basic incompetence.

[6] "Franchising", by R. Hotch, Nation's Business, vol. 73, no. 9, Sept. 1985, pp. 37-38.

through marketing efforts. While prospective franchisees may have access to these marketing data, they are less able than the franchisor to identify whether product success to date has been location specific, attributable to the efforts of the owner, or due to high demand.[7] In this case, a franchisee will want some assurance that the franchise is profitable (or that the franchisor is competent). Under this assumption of asymmetric information, we show that the franchisor will use two instruments to convey information about the new product: dual distribution, the practice of selling products through "company-owned" outlets as well as through franchises, and a royalty in the franchise contract.[8]

This analysis provides one answer to the organizational puzzle posed in the economic literature on franchising: Why do firms simultaneously use different organizational forms?[9] Our explanation for dual distribution is related to Leland and Pyle's (1977) analysis of how an entrepreneur can signal high profitability of its product/project to potential shareholders by investing in the project himself. The basic intuition behind their results is also used here in that entrepreneurs (franchisors) are not willing to invest their own money in low value projects. In addition to predicting when a firm with private information will directly invest in

[7] A franchisee could find out information about the franchise by speaking to other franchisees. In the case of a failing franchise, however, the moral hazard problem is clear: Existing franchisees wanting to sell will have every incentive to encourage a prospective franchisee. In the Pop-Ins case, "every franchisee has a stake in [the franchisor] selling another franchisee. Besides franchisees were loath to admit they'd made a bad investment." ("Fear of Franchising," by C. Hartman, Inc., June 1987, pp. 104-122.)

[8] A natural question to ask is whether there are alternative, possibly less expensive ways, to convey information. For example, a franchisor might offer to refund a franchisee's investment if the franchise fails. Limited guarantees are used by a few franchisors, but we are not aware of any franchisor who offers full insurance to its franchisees. One explanation for this lack of full insurance is that a franchisor's risks are correlated: if the franchise concept is unprofitable, all franchisees will demand their money back and the franchisor will go bankrupt. For example, the franchisor of the Pop-Ins maid service offered a money-back guarantee on part of the franchisee's investment, but no franchisee actually received any compensation. Alternatively, there are clear moral hazard problems with using warranties (see Lutz (1989)). Hadfield's (1990) hypothesis that courts tend to protect the interests of the franchisor, and less so those of the franchisee, short of "blatant fraud" by a franchisor, may also explain why such guarantees are not enforceable.

[9] Of firms that sell products through franchises, nearly 80% also sell through "company-owned" outlets. In a study by Martin (1989), 431 out of 548 firms company owned some of their outlets. The mean percentage of outlets franchised was 82.75. The minimum percentage was 1.62.

the product (as in Leland and Pyle), our model explains the specific form of that investment (direct investment in company-ownership vs. share contracts with franchisees). That is, we analyze the tradeoff among multiple instruments for signalling.

Two alternative explanations for dual distribution have been advanced in the literature. Rubin (1978) assumes that contracts to franchisees are more incentive based; so under moral hazard, company outlets must be monitored more than franchised outlets. In his model, if monitoring costs are low, company outlets will be used; however, if monitoring costs increase rapidly with the distance from corporate headquarters, a firm may choose to own nearby outlets while franchising its remote outlets. A related explanation postulates that if the costs of monitoring a rapidly growing number of outlets is very high, rapidly expanding firms will want to franchise some of their new outlets to reduce monitoring costs. A second theory relies on capital market imperfections. According to this theory, firms franchise as a way to obtain capital for expansion. Any firm that could obtain capital through other means at lower expense would do so and own all outlets.[10]

Our explanation relies neither upon capital market constraints nor upon location-specific factors, and its predictions differ from those of the two previous explanations. For example, in contrast to the alternative theories, the signalling explanation for dual distribution predicts that over time the proportion of company-owned outlets falls as the franchisor's private information is revealed to prospective franchisees.

In section 2 we discuss the nature of dual distribution; in particular, we identify the critical economic distinction between company-owned and franchised stores on which our explanation turns. In section 3 we present a simple model of symmetric information. Section 4 outlines the assumptions of the model of symmetric information and section 5 derives the separating equilibrium franchise contract and configuration of company-owned and franchised outlets. Finally, preliminary empirical evidence for the predictions of the model is provided in section 6.

---

[10] This theory cannot explain the phenomenon that some franchisors provide financing to their franchisors. Out of 1,081 franchisors listed in Entrepreneur Magazines's 1990 franchise survey, 224 indicated that they provided financing to franchisees.

2. The Nature of Franchising

In this paper we consider an individual (firm) who wants to market a new product through several retail outlets. We seek to identify the profit-maximizing method of product distribution -- franchising, company-ownership, or both (dual distribution). Fundamental to this question is an assumption that ownership and franchising are economically different. Before describing our model of dual distribution we briefly discuss the critical difference between these two organizational forms.

The franchise literature commonly defines franchised and company outlets by the form of compensation offered to the "unit operator" (manager in the case of a company-owned store and franchisee in the case of a franchise). For example, Brickley and Dark (1987) suggest that "franchising can be viewed as near one extreme of the managerial compensation continuum; i.e., compensation is incentive-based, versus some mix of salary and incentive compensation for the managers of centrally owned units." Under this definition of organizations, the question of dual distribution becomes one of explaining the simultaneous use of different contracts.

In reality, franchising and company ownership do imply different contracts to the unit operator. We argue, however, that this contractual distinction is endogenous to a more fundamental ownership distinction between the two organizational forms. We define an organization by the assets owned by the unit operator, that is, the assets over which he has control.[11] For example, a franchisee owns the local unit, generally has control over production and sales, and receives the full capitalized value of the future income stream when the unit is sold.[12] The manager of a company-owned store does not hold these rights; his future income stream extends only until he leaves the firm, retires, or dies. Ownership of the outlet and control over production and sales decisions remain with the owner of the brand

---

[11] This is in the spirit of Grossman and Hart (1986).

[12] While franchisees own the building and the right to sell or lease it to others for other purposes, to sell it to a franchisee who will use it to sell the product requires the franchisor's approval, since she owns the brand name. However, in the U.S. the franchisor's ability to prevent resale has been limited by the courts.

name.

Even if the same set of payment rules is available to both forms of organizations, the incentive effects of a particular contract will differ for a franchisee and manager because of ownership rights described above. We examine this and other implications of this organizational framework in a simple model of symmetric information in the next section.

3. Optimal Product Distribution under Symmetric Information

Consider an individual (firm) who wants to market a new product through each of N outlets. The location of each outlet is fixed, and the owner of the outlet enjoys a local monopoly. The firm can either retain ownership and hire a manager for an outlet or transfer ownership to a franchisee, who then operates the unit for one period. Output decisions are made by the owner of the outlet but effort is chosen by the unit operator. The markets for managers and franchisees are competitive and all parties are risk neutral.[13] The demand at each outlet, indexed by t, is identical and observable to all.

We assume that while effort of the unit operator, e, is observable to all market participants, it is not verifiable to the courts. This assumption rules out forcing contracts on effort. Moreover, this assumption is key to the distinction between company-owned and franchised stores discussed in the previous section, since it implies that the future value of the outlet is not contractible.[14]

The ownership distinction between organizations described above has at least two implications in this model: First, contracts offered to franchisees will be relatively more incentive-based, consistent with reality and assumed in the literature. Second, ownership is less profitable than franchising when demand is observable; hence, there is no incentive for

[13] Adding franchisee risk aversion would change the form of the optimal contract to offer the franchisee under symmetric demand, for example, if there were an unobservable random element to demand. But franchising would still provide better incentives than company ownership, and the main result of our analysis, that dual distribution is used to signal franchise type, would not change.

[14] A manager could be paid shares in the firm; however, this would not adequately reflect the value of the local unit.

dual distribution.[15] To see the first implication, let $\pi_t(x)$ be first-period profits from an outlet in which x units are sold, g(e) disutility of effort, where $e \geq \underline{e}$, and $V_t(e)$ the future value of the outlet.[16] Consider first a company-owned store. The owner will maximize profits by setting x_t equal to:

$$(1) \qquad x_t^* = \text{argmax } \pi_t(x).$$

Moreover, since the manager leaves after one period and he has no ownership rights, he will put forth minimum effort $\underline{e}$. Therefore, the owner of the product will pay the manager a wage equal to the value of the manager's next-best employment opportunity, the wage earned in this opportunity minus the disutility of the effort required. We will normalize this net value of the next best opportunity to zero.[17] Hence, for a discount factor equal to 1 profits from a company-owned outlet are:

$$(2) \qquad W_t^{c*} = \pi_t^* - g(\underline{e}) + V_t(\underline{e}),$$

where $\pi_t^* = \pi_t(x_t^*)$.

Next, consider a franchised outlet. Since the franchisee controls output, the optimal contract makes the franchisee residual claimant over current and future profits. In this case, the franchisee will choose the profit-maximizing output level in (1). Moreover, since the franchisee owns the outlet, he cares about the future value of the firm. Hence, the franchisee

---

[15] This is important for our signaling model in the next section. If company-owned stores were more profitable, then a franchisor with private information would have no incentive to franchise.

[16] The unit operator's effort, e, can increase future profits by increasing the future demand for the product at the outlet, or by decreasing the outlet's future costs. Efforts to improve service, for example, would increase future demand by increasing the number of repeat customers. Efforts to train and retain current workers would decrease future labor costs.

[17] The fact that the contract is not incentive-based is attributed to two assumptions. The first assumption is a simplifying one that the unit operator's effort does not affect current profits. If this assumption were relaxed, the manager's contract would change, but without affecting the main results. (See Lutz (1989) for an example of the effects of agent moral hazard on signaling equilibria.) Second, since effort is not verifiable to the courts, the future value of the firm is not contractible.

will choose effort level:

(3) $e^* = \text{argmax } [- g(e) + V_t(e)].$[18]

The franchisor can capture the entire profit stream with a fixed fee equal to the unit's profits,

(4) $W_t^{f*} = \pi_t^* - g(e^*) + V_t(e^*).$

Since $W_t^{f*} > W_t^{c*}$, franchising dominates company-ownership. Hence, dual distribution does not arise in this model when the level of demand is common knowledge.[19]

4. Dual Distribution under Asymmetric Information

To provide an explanation for dual distribution, we relax the assumption of symmetric information on demand. The assumptions of the model are outlined here; in section 5, the equilibrium is derived.

4.1 Informational Assumptions

In contrast to the model in section 3, the owner of the new product is assumed to have private information on the demand for the product.[20] We assume that there are two possible levels of product demand, indexed by type t = L,H. As before, effort of the unit manager is

---

[18] For simplicity, we assume that $e_t^* = e^*$; that is, the optimal effort level is independent of demand. For example, it might be the case that $V_t(e)$ is an additively separable function of t and e.

[19] This would not be the case if the franchisor of type t could offer managers the option to buy the franchise at the end of the period at a price equal to $V_t(e^*) - g(e^*) + g(\underline{e})$ and pay a wage equal to $g(\underline{e})$ in the first period. Such a contract would yield full rents to the franchisor and be acceptable to the manager. According to our ownership definition of organizations, the company-owned store with a buy option is equivalent to a franchised outlet: in each case the unit operator has some property right to the future value. Hence, the more precise statement is that franchising dominates company-ownership without a buy option.

[20] For example, there might be two levels of product quality or franchisor skill which affect demand. In this case, quality would be directly observed by consumers (and franchisees) before they purchased the product, but they would not know how many units other individuals purchased.

not verifiable to third parties.

In addition to these two informational assumptions, both of which are fundamental to the model, we make two simplifying assumptions. First, all market participants know marginal production costs, c, and observe price and sales at the end of the first period; however, only price and sales are verifiable to a court. The inability to verify production costs restricts the unit operator's contract to the more commonly observed share of sales, rather than profits. Second, we assume that α, the proportion of company-owned stores, and s^f, the franchise contract, are readily available to potential franchisees. The take-it-or-leave-it contract s^f takes the form (r,F), where r is the proportion of total revenue paid to the franchisor and F is the upfront fixed fee.[21] Hence, α, r and F can be used as signals of the franchisor's type.[22] Contracts struck with managers of company-owned stores are not observable to franchisees.

4.2 Demand Assumptions

We adopt the following assumption on demands:

Assumption 1: Demands are of the form $p_t = a_t - f(x)$[23], with the following characteristics (i) $f'(x) > 0$; (ii) $-2f'(x)-f''(x)x \leq 0$; (iii) $-3f''(x)-f'''(x)x \leq 0$.

The first part of Assumption 1 is the usual negatively-sloped demand assumption; the second part is a necessary and sufficient condition for franchisee's profits to be concave in output; and as we will show, the third guarantees that marginal revenue is concave in output, a sufficient condition for a unique separating equilibrium, under an appropriate refinement.

---

[21] Most franchise contracts are of this form. See Lafontaine (1990) for further discussion.

[22] This is consistent with the information franchisors are required to disclose under the Federal Trade Commission's Franchise Rule of October 1979 and more restrictive state laws in CA, HA, IL, IN, MD, NY, ND, OR, RI, SD, TX, VA, WA, and WI.

[23] Since most franchised products have substitutes, the assumption of an intercept on demand is reasonable. However, note that only the intercept changes with type. This restriction on demand is not necessary; for example, many of the main propositions can be shown to hold under constant elasticity demand.

4.3 Timing of Events

There are two periods in the model. At the beginning of period 1, the franchisor draws the product "type" -- high (H) or low (L) -- from a probability distribution, which is known to all players. The H-type franchisor sends a message on the product type through $\{\alpha, s^f\}$.

For both α and s_f to be effective signals, the franchisor must be able to commit to them. As is conventional in signaling models, we assume the terms of the contract s_f cannot be renegotiated. Since the franchisee does not verify the type until the end of the period, this assumption is not unreasonable. To see this, note that changing the terms of s_f after the contract is signed would require the franchisee's approval. However, any change in s_f made by the franchisor to reduce the signaling costs would be profitable to both types, if accepted. Hence, such changes would not be acceptable. In contrast, a change in α would not require the approval of the franchisee. We assume that the franchisor commits to αN managerial contracts before approaching the franchisee, and that decision is irreversible; for example, any change in such contracts would require large transaction costs.

If the unit operator accepts the contract, then the upfront fixed fee is paid and production takes place. The owner of the outlet chooses output, x, while the unit operator chooses effort, e. Price is chosen instantaneously by the owner of the outlet when the output is sold.[24]

The value of the outlet is common knowledge to all market participants in period two. Managers leave and franchisees sell their outlets. The sequence of events is summarized in Figure 1. We turn now to the separating equilibrium of this model.

5. <u>The Separating Equilibrium under Asymmetric Information</u>

Given the assumptions of the model outlined in section 4, we look for a perfect Bayesian Nash equilibrium satisfying the Cho-Kreps (1987) Intuitive Criterion. We

---

[24] Franchisors can and do suggest retail prices, but their direct power to control pricing or output levels is limited by antitrust laws.

demonstrate in the appendix that no pooling equilibrium satisfies this refinement. Thus we restrict our attention to separating equilibria. First, note that in any separating equilibrium, the low-demand type will offer the contract that would be optimal under symmetric information. That is, in equilibrium, the low-demand type uses 100% franchising, and offers a franchise contract with $r = 0$ and $F = W_L^{f*}$, where W_L^{f*} is defined in (4) with $t = L$. The separating contract for the high-demand type is derived in the remainder of this section.[25]

5.1 The Franchisor's Problem

A separating contract α and s^f for the H-type must satisfy two constraints. First, the H-type franchisor must choose α and a franchise contract that dissuades the L-type franchisor from mimicking her. Second, the franchise contract must be acceptable to the franchisee. To determine these conditions, we first derive the form of the unit operator's contract.

Contracts to managers

Recall that the number of company-owned stores and the corresponding managerial contracts are committed prior to the franchise contract. Since a manager's contract is not a signaling instrument, the franchisor of either type will offer a fixed wage contract generating the maximum possible profits, W_t^{c*}, given in (2).[26] It will be useful for the development of intuition to split this profit into two parts: a period 1 return of $\pi_t^* - g(\underline{e})$, and a period 2

[25] This separating equilibrium is also the only equilibrium surviving iterated elimination of weakly dominated strategies. See Milgrom and Roberts (1986) for an application of this method.

[26] An alternative contract would be a wage contract with the buy option discussed in footnote 18. However, under asymmetric information the H-type cannot extract its full information rents by use of such a contract. Suppose that the H-type firm offers to pay its managers a wage of $g(\underline{e})$ and gives them the option to buy the unit at a price of $V_H(e^*) - g(e^*) + g(\underline{e})$. The managers, knowing that they will exercise the option if the franchise turns out to be H-type and they choose effort e^*, will indeed choose e^* if they believe that the franchisor is H-type. However, the L-type firm would then be willing to offer the same option; misled franchisees would take effort e^* but not exercise the option, yielding total profits of $\pi_L^* + V_L(e^*) - g(\underline{e})$. Knowing this, the manager would not believe upon observing the option contract that the franchisor is certainly H-type, and would put forth less than e^* effort. Option contracts, like signalling, induce an information loss for the H-type franchisor. Whether the loss is greater under signalling or under an option contract is beyond the scope of this paper.

return of $V_t(\underline{e})$.

Contracts to franchisees

Next consider franchised outlets. The owner of the high-demand product would like simply to charge an upfront fee equal to W_H^{f*} given in (4). But this would encourage an L-type franchisor to offer the same contract. Consequently, profits from an H-type outlet must be distorted to signal the type.

Under a long-term contract, the H-type franchisor could signal its type with a sales royalty in period 1, period 2, or both periods.[27] Any two-period contract of the following form induces efficient effort: In the first period, the franchisee pays a franchise fee, F, and a sales royalty, r, on revenues; in the second period, only a fixed payment, F_2, is specified in the contract.[28] This induces optimal effort $e*$,[29] so that the future value of the H type unit is $V_H(e*)$. This future value can be extracted by the franchisor through either the first period or second period fixed fees. We will assume without loss of generality that this is done through the second period fee, so that $F_2 = V_H(e*)$. A franchisee who has been deceived by an L-type franchisor is obligated to pay the same F_2; since the future value of this franchisee's unit is only $V_L(e*)$, he will suffer a loss in the second period.[30]

[27] A series of one-period contracts would be inferior to a two-period contract: Since the franchisee would know that the benefits from his efforts would be extracted in the subsequent period, short-term contracts would not induce the efficient level of effort.

[28] Alternatively, a second-period sales royalty could be specified in the contract. But since demand is revealed to all by the beginning of the second period, this contract would be renegotiated. The franchisee would expect to receive only part of the reward for his effort after renegotiation, and so his choice of effort would be distorted.

[29] Since the franchisee expects a second-period return on effort equal to $V_H(e)$-$g(e)$-F_2, he will choose the efficient level of effort, $e*$.

[30] Another contractual alternative would be for the H-type franchisor to offer the franchisee an option to turn the unit over to the franchisor, thus avoiding the second period fee. A deceived franchisee would suffer no second period loss under a contract of this type, and the L-type franchisor would find it harder to mimic the H-type firm. However, the H-type firm would have to reduce F_2 below $V_H(e*)$; otherwise it would be profitable for its franchisees to take minimum effort in period 1 and turn the unit over in period 2. All the qualitative results of this model go through when the H-type franchisor signals with a contract of this type: proofs are available from the authors upon request.

Profits under the contract s^f

Given these contractual arrangements, we can define profits from a choice α and s^f = (r,F) for both types of franchisors. Let $x_H(r)$ be monopoly output of a franchisee who is offered s^f and $R_i(r)$ be total revenues generated from this output when demand is type i. Then, the H-type franchisor's two-period profits from a franchised outlet are given by the sum of the unit's period 1 return from the fixed fee and the royalty payment, and the period 2 return:

$$(5) \qquad W_H^f(r,F) = F + rR_H(r) + V_H(e^*).$$

A deceptive L-type franchisor who adopts this contract will earn first-period revenues of $F + rR_L(r)$. Under s^f, the franchisee, believing demand to be high, will choose $x_H(r)$, but when he sells the product on the market, price will be lower than expected. Under free disposal, any output produced greater than the output that maximizes total revenue of the low demand will not be offered on the market. Let r' be the sales royalty for which $x_H(r')$ maximizes total revenue for the low demand. Then for any $r \leq r'$, revenues generated by a deceived franchisee will equal $R_L(r')$. The two-period profits from a franchised outlet earned by a deceptive L-type franchisor are again the sum of the first period and second period returns from the unit:

$$(6) \qquad W_L^f(r,F) = \begin{cases} F + rR_L(r') + V_H(e^*) & \text{for } r \leq r' \\ F + rR_L(r) + V_H(e^*) & \text{for } r > r'. \end{cases}$$

Conditions for a separating equilibrium

We are now able to specify the conditions for a separating equilibrium. First, s^f must satisfy the separation constraint:

$$(7) \qquad \alpha\, W_L^{c}{*} + (1-\alpha)W_L^{f}(r,F) \leq W_L^{f}{*},$$

where $W_L^{c}{*}$, $W_L^{f}(r,F)$, and $W_L^{f}{*}$ are given by (2), (6), and (4), respectively. The separation constraint states that a deceptive franchisor must earn less under s^f and α proportion of company-owned stores than if she told the truth and simply adopted the organization under symmetric information.

Second, the contract must satisfy the rationality constraint:

$$(8) \qquad F \leq (1-r)R_H(r) - cx_H(r) - g(e^*).$$

That is, the fixed fee cannot exceed the share of revenues to the franchisee less production and effort costs. In a contract satisfying this constraint with equality, the franchisee will earn a total return of 0 over the two periods.

As is typical of signaling models, a plethora of signaling equilibria are implied by (7) and (8). By the Intuitive Criterion, all contracts given by the strict inequality of (7) are eliminated. Label the equality relationship in (7) as (7′). Then, the equilibrium contract and fraction of company-owned outlets solve the following problem:

$$(9) \qquad \max_{\alpha,r,F}\ \alpha\, W_H^{c}{*} + (1-\alpha)\, W_H^{f}(r)$$

subject to (7′) and (8).

The solution to (9), which will be unique under Assumption 1, gives the separating equilibrium organization for the H-type franchisor, as long as the H-type franchisor's payoff under this organization is greater than the maximum profits she could earn if she were believed to be L-type.[31] We solve this problem in two parts: In section 5.2 we find the contract (r,F) that solves (9) for a fixed α; in 5.3 the equilibrium α is derived.

5.2 The Equilibrium Franchise Contract

Propositions 1 and 2 characterize the equilibrium franchise contract, for a given α. Define $\Delta\pi = \pi_H(x_H^*) - \pi_L(x_L^*)$, the difference across types in the maximum first period unit profit. Let $\Delta V(e^*) = V_H(e^*) - V_L(e^*)$, the difference across types in the maximum future value of the unit, and let $\Delta V_t = [V_t(e^*)-g(e^*)] - [V_t(\underline{e})-g(\underline{e})]$ for t = L,H; this is the difference net of effort costs between the future value of a franchised and a company owned unit of type t.[32]

<u>Proposition 1</u>. For every α there exists a unique optimal franchise contract for the H-type franchisor. This contract has the following features: (i) Define $\hat{\alpha} = (\Delta\pi + \Delta V(e^*)/(\Delta\pi + \Delta V(e^*) + \Delta V_L)$; then for $\alpha \in [\hat{\alpha},1)$, r = 0 and $F = \pi_H^*$; for $\alpha < \hat{\alpha}$, r > 0. (ii) The separating constraint is (weakly) decreasing in r at the optimal r.

Proof in the appendix.

---

[31] A sufficient condition for this to be so is that W_H^{c*} be greater than the maximum return the H type franchisor could earn if she were believed to be low type, the maximum over r of $rR_H(x_L(r)) + (1-r)R_L(x_L(r)) - g(e^*) + V_L(e^*)$.

[32] This term may also be thought of as the difference between the total return from a franchised unit versus a company owned unit when the firm's type is observable by franchisees.

Proposition 1 implies that the terms of the franchise contract change with the proportion of company-owned stores operated in equilibrium by the high demand type of franchisor. For a sufficiently large proportion of the relatively inefficient company-owned stores, the franchisor does not need to manipulate the terms of the contract offered to franchisees to convince them of the product's demand.[33] The high demand franchisor can offer the efficient contract, charging a zero royalty on sales and using the fixed fee to extract all profits from the franchisee. If α is sufficiently low, a sales royalty will be used to create an inefficiency in the H-type state to prevent deception by the L-type franchisor. Fewer outlets are company-owned, but the franchisor is not earning the maximum possible returns from its franchised outlets. The optimal franchise contract for a given α is illustrated in Figure 2.[34]

Proposition 1 implies that a positive sales royalty will be used to signal, for all $\alpha < \hat{\alpha}$. Proposition 2 provides a comparative statics result on the optimal sales royalty, $r(\alpha)$, and franchise fee, $F(\alpha)$.

<u>Proposition 2</u>. For a given $\alpha < \hat{\alpha}$, the sales royalty required to signal increases with a_H and the corresponding franchise fee decreases with a_H.

[33] Note that for $\Delta V_L > 0$, $\hat{\alpha} < 1$. Hence, 100% company ownership is never optimal.

[34] An implication of Proposition 1 is that the firm would be able to signal through r and F alone if dual distribution were unobservable. A similar result for licensing contracts is found in Gallini and Wright (1990).

5.3 The Equilibrium Dual Distribution Configuration

One implication of the previous section was that the firm could signal through the use of r and F alone, distorting the first period return from the franchised units in order to signal. It is easy as well to see that α alone could also be used to signal. If the H-type firm did not distort its choice of these variables, $r = 0$ and $F = \pi_H*$ would be chosen. The H-type firm could signal its quality by setting $\alpha = \hat{\alpha}$: in this case, it is the future return from some of the units that is being distorted to signal type. However, analyzing how these signals can be combined is a more complicated problem. By choosing a positive level of α, the firm is choosing to distort the future value of some units from the first best level: by choosing a positive royalty r, the firm is choosing to distort the current return from its remaining units from the first best level.

In this section, we determine the equilibrium combination of owned and franchised stores, $\alpha*$, that an H-type franchisor will choose, given the optimal franchise contract, $(r(\alpha), F(\alpha))$, derived in section 5.2. The separating equilibrium maximizes (9) over α, after substituting in $(r(\alpha), F(\alpha))$.

The first-order condition for the optimal level of dual distribution $\alpha*$, found by maximizing (9) with respect to α given the profit-maximizing choices of r and F, is given by:

$$(10) \qquad W_H^{c}* \ - W_H^{f}(r(\alpha*),F(\alpha*)) + (1-\alpha*)\ dW_H^{f}(r(\alpha*),F(\alpha*))/d\alpha \ = \ 0.$$

An increase in the proportion of company-owned units has a direct and an indirect effect on the H-type franchisor's profits. The direct effect is the difference between return from a company owned unit and return from a franchised unit; this is given by the first two terms in (10). The indirect effect is due to the effect of the change in α on r and F; this is given by the third term in (10).

From (10), we can now identify when the separating equilibrium is characterized by dual distribution; that is, when $\alpha^* > 0$. This requires two lemmata. Let $\bar{r}$ be the tangency of the isoprofit curve and the separation constraint[35] and let $\bar{\alpha}$ be that value of α for which the rationality constraint intersects the separation constraint at $\bar{r}$. The values $\bar{r}$ and $\bar{\alpha}$ are illustrated in Figure 3; in Figure 3a, $\bar{\alpha} > 0$; in 3b, $\bar{\alpha} < 0$. In the former case, note that for $\alpha < \bar{\alpha}$, $d\bar{r}/d\alpha = 0$. The following lemma states the condition under which $\bar{\alpha} \geq 0$.

<u>Lemma 1</u>. There exists an $\underline{a}_H$, such that for $a_H \geq \underline{a}_H$, $\bar{\alpha} \geq 0$.
Proof in the appendix.

Next we identify when $\alpha^* > \bar{\alpha}$, that is, when the proportion of company-owned stores will increase until the rationality constraint binds in equilibrium. Lemma 2 establishes a sufficient condition for this to occur:

<u>Lemma 2</u>. If $\Delta V_H \leq \Delta V_L$, then $\alpha^* > \bar{\alpha}$.
Proof in the appendix.

---

[35] This royalty $\bar{r}$ does not depend on α, since a change in α has no effect on the slopes of the isoprofit curve and the separation constraint.

This condition, $\Delta V_H \leq \Delta V_L$, says that the difference between the return from a franchised unit and the return on a company owned unit is larger for the L-type than for the H-type, when type is observable. Lemmata 1 and 2 directly imply our first main result on dual distribution:

Proposition 3. If $\Delta V_H \leq \Delta V_L$, a signaling equilibrium exists in which $\alpha^* > \tilde{\alpha} \geq 0$ when a_H is sufficiently large.

Proposition 3 gives a sufficient condition for dual distribution to be observed in equilibrium. One way to understand this result is to begin with an $\alpha < \tilde{\alpha}$, assuming that $\tilde{\alpha} > 0$, and evaluate the left hand side of (10); this is done in the proof of Lemma 2. The direct effect -- the sum of the first two terms in (10) -- is indeterminate in sign, but is greater than ΔV_H.[36] The indirect effect of the change in α on the H-type's return then comes solely through the effect of α on $F(\alpha)$: at this low level of α, an increase in α has no effect on the royalty and $r(\alpha)$ remains constant at $\tilde{r}$. Because F changes with α in such a way as to keep the L-type franchisor indifferent between mimicking the H-type and revealing its low type by franchising all its units, the total indirect effect is $\Delta V_L/(1-\alpha)$. The positive indirect effect outweighs the possibly negative direct effect under the condition given in the proposition. Dual distribution will then be used in the separating equilibrium as long as $\tilde{\alpha} \geq 0$; by Lemma

[36] The most that can be extracted from the franchisees is W_H^{f*}. Since the franchisor earns W_H^{c*} on each company-owned unit, the minimum difference between what the H-type franchisor receives from a company-owned versus a franchised unit is $W_H^{f*} - W_H^{C*} = -\Delta V_H$.

1, this will be true if a_H is sufficiently large.

For small a_H, the rationality constraint may bind at $\bar{r}$ when no company-owned stores are employed (α=0), as shown in Figure 3b. In this case, it can be shown that there is a local maximum at $\alpha = 0$.

The condition given in Proposition 3 is not necessary for dual distribution. Suppose instead that $\Delta V_H > \Delta V_L$. If the difference between these values is relatively small, then it can be demonstrated that $\alpha^* \geq \hat{\alpha}$, and if $\underline{a}_H$ is sufficiently large, a signaling equilibrium with $\alpha^* > 0$ exists.

In the signaling equilibrium (whether or not dual distribution is used), α^* will be strictly less than $\hat{\alpha}$, the proportion of company-owned stores which signals type without distorting franchise profits. That is:

<u>Proposition 4</u>. A positive royalty is always used in the separating equilibrium.
Proof in the appendix.

Proposition 4 implies our second main result on dual distribution: When dual distribution is used to signal, r and F will also be used. That is, the three signaling instruments are complementary. The intuition for this result is easily seen by considering the profits earned by the H-type franchisor when it signals through α alone. This can be done by choosing $\hat{\alpha}$, with r = 0 and F = π_H*. The effect on profits of a decrease in α, holding r constant, will be positive, and this effect is larger for the L type franchisor than for the H-type. To continue to signal its type, the H-type must increase the royalty rate it charges its franchisees (and decrease the fixed fee) to compensate for the increase in the L-type's profits. Because increases in the royalty rate from r = 0 hurt the H-type less than they hurt the L-type, the increase in r that just keeps the L type indifferent about mimicking the H-type actually improves the latter's profits.[37]

Finally, in a dual distribution equilibrium, the franchisor never shares rents with the franchisee, as described by the following proposition.

Proposition 5. When $\Delta V_H \leq \Delta V_L$, rents will never be shared in the separating equilibrium; when $\Delta V_H > \Delta V_L$, rents will be shared in this equilibrium only when 100% franchising is used.

Proof in the appendix.

Proposition 5 implies that in an equilibrium in which only franchising is used there may be a second source of signaling cost in addition to the distortionary sales royalty: the franchisor may have to share some rents with the franchisee. This may be true even when the franchisee's opportunity cost is zero. The franchisor is forced to set F low, given r, since a high fixed fee would encourage the L-type to lie. However, when these costs of franchising are sufficiently high, the franchisor will increase the proportion of company-owned stores until no more rents are shared.[38]

---

[37] This is demonstrated by showing that the third term on the left hand side of (11) is negative. The proof of this result provides further details.

[38] This result is in contrast to that in Gallini and Wright (1990). In a licensing model of asymmetric information, an innovator who signals with the contract offer must share some of the rents with an exclusive licensee when the difference in types is large. In contrast to this model where dual

To illustrate the results in this section, the model is simulated for the case of linear demand in the next section.

5.4 An Example

In this section we simulate the model for linear demand, $p_t = a_t - x$. The results were robust qualitatively for all parameter values considered; we present an example in Table 1. The values of the L-demand intercept range from .8 to 1.5, while the values of the H-demand intercept range from 1.8 to 4.[39] Marginal costs = .5. We assume $\Delta V_H = \Delta V_L = .04$, and $V_H(e^*)-g(e^*)$ is set equal to full monopoly profits for the H-demand.

Several interesting results can be noted. While the proportion of company-owned stores increases monotonically in profitability, the sales royalty increases only initially in a_H. An increase in a_H has two opposing effects on r: From Proposition 2, r increases for increases in a_H given α. However, an increase in α moves the separation constraint up the rationality constraint, resulting in a decrease in r. For moderate values of a_H, the first effect dominates so both the sales royalty and company-ownership signals; however, as α approaches the perfectly revealing value $\hat{\alpha}$ for large a_H, the sales royalty needed to signal falls. Similarly, an increase in a_H has two opposing effects on F; for linear demand, the effect from a higher α dominates and F increases for increases in profitability.

The model is also simulated for different values of the L-type profitability. An increase in the H-type profitability is similar to a decrease in the L-type profitability in that

---

distribution is considered, they show that nonexclusive licenses may alleviate that problem.

[39] These values correspond to the cases in which a_H is sufficiently large such that $\alpha^* > 0$.

the signaling problem becomes more severe. The qualitative results are preserved: a reduction in the L-type profitability increases the proportion of company-owned stores; moreover, the sales royalty initially increases then decreases as the L-type becomes less profitable.

6. Conclusions

The model presented in this paper provides an explanation for the dual distribution of new products: Owners of new products signal private information on demand by distributing their products through both company-owned stores and franchised outlets.

Several extensions are possible. First, the predictions of the model can be tested. Among the testable implications of the model is the following. In our model information on demand eventually becomes common knowledge to all market participants. Since less-profitable company-owned stores are used as a signal, after the need for signaling disappears then so too will company-owned stores. Hence we have:

P1: Over time, the proportion of company-owned outlets falls.

Note that the capital constraint and monitoring explanations for dual distribution given in the literature and briefly discussed in the introduction are inconsistent with **P1**. In contrast to our model, these two explanations assume that company ownership is, ceteris paribus, the preferred organizational form.[40] Under the capital market imperfection theory, one would

[40] For example, company ownership may minimize free-riding on the brand name.

expect that since franchises are used as a means for acquiring capital, this theory predicts that α would increase as the capital constraint is relaxed over time. Moreover, the monitoring costs explanation would not predict any systematic change in the organizational arrangement, to the extent that moral hazard persists over time.[41]

Lafontaine (1989) shows that the proportion of franchised outlets increases and the royalty rate tends to fall as the number of years in business increases. Similarly, Martin (1988) finds that the proportion of company ownership by firms engaged in dual distribution falls over time. Although these studies provide preliminary support for **P1**, they are not precise tests of **P1** since they follow an average trend of all franchisees, rather than single franchisees.

To test **P1**, time series data on the proportion of company ownership by individual firms are required. Data on α during the period 1987-1989 for 295 companies that began operations during 1980-1987 are readily available.[42] For these data, we consider companies for which α decreased between 1987 and 1989 to be consistent with **P1**. Let X be the number of these companies.[43] Some companies did not have any own outlets during 1987-1989.[44] We argue that these data should either be eliminated on the ground that dual distribution was

---

[41] Other explanations for dual distribution may be consistent with **P1**; however, we focus on the two main explanations that have been advanced in the literature.

[42] These data are from the January 1990 issue of Entrepreneur Magazine.

[43] Companies with a constant $\alpha > 0$ were not included in X. This biases the results against the signaling theory: An α that declines to a positive constant may not refute the signaling theory since there may be other motivations for dual distribution which do not disappear when the franchisor's private information is revealed.

[44] Large companies with $\alpha \leq .01$ over 1987-1989 are assumed to have an $\alpha = 0$. We attribute the few company-owned stores to the possibility that with a relatively large number of franchised outlets, some are likely to be sold to the company in any year until a suitable franchisee is found.

never used, or included in X on the ground that company-owned stores were sold after information was revealed, as predicted by **P1**. We consider both cases: Case 1 eliminates companies with α=0 and hence biases results against **P1**; case 2, in which these data are included, biases results in favor of **P1**. The sample proportion ($\hat{\rho}$ =X/N), is given below for each year and for all observations.

Year Business Started	Case 1	Case 2
1980	.655	.722
1981	.686	.732
1982	.781	.833
1983	.750	.792
1984	.719	.750
1985	.889	.913
1986	.857	.893
1987	.800	.889
All years	.762	.810
N	235	295

The fact that $\hat{\rho}$ is significantly greater than .5 is support for **P1**. A careful test with an extended time series remains for future research.[45]

The dynamic aspect of this signaling model is admittedly simple in that everyone learns the type at the end of the first period. This assumption could be relaxed to allow for gradual learning of information over time as company stores are opened. Moreover in an

[45] Other predictions of the model could be tested. A second prediction which follows from Proposition 3 and discussion is:

P2: For a range of values above some minimum, the proportion of company-owned stores increases in profitability per outlet.

In a recent paper, Lafontaine (1990) finds empirically a negative relationship between company-ownership and firm growth. This does not necessarily refute **P2** since firm growth is endogenous and may not be positively correlated with profitability per outlet. Since N is fixed in our model, we do not offer any predictions on the relationship between α and firm growth.

Also, from Proposition 4 we have:

P3: A positive royalty is used on new products distributed through franchised outlets.

extended model, the number of outlets, fixed in our model, could be made endogenous. Then, if the optimal number of outlets differ according to type, N would be another signaling instrument. This change and the extension to a more dynamic model would allow for a more complete analysis of the evolution of organizational form and firm growth as private information is revealed to franchisees over time.

In empirical research Brickley and Dark (1986) and Norton (1988) find agency considerations, in particular moral hazard by the unit operator as analyzed by Rubin, to be important in the franchise-ownership decision.[46] Familiar principal-agent problems other than Rubin's and the one considered here (informed principal) may explain dual distribution. For example, as mentioned in the introduction, if franchising is more profitable in the absence of agency problems, company-ownership may arise when there is franchisor moral hazard (company-ownership is a commitment to improve the value of the brand name) or when franchisees have private information on local demand conditions (company-ownership is a means of acquiring information).[47] [48] Further theoretical and empirical development of these alternative theories remain for future research.

---

[46] Brickley and Dark focus primarily on location and other unit-specific variables in determining whether a given unit is franchised or company owned. Since our explanation of dual distribution does not rely on location-specific factors, conclusions regarding signaling cannot be drawn from their study. Norton makes comparisons within an industry between states to determine the extent that state-specific variables affect the percentage of franchised outlets in that state. Although some of his variables are relevant to signaling, conclusions about the percentage of outlets owned by franchisors cannot be drawn since his data set includes independently owned outlets (those without any relationship to a franchising firm), and he does not distinguish between company owned outlets and independently owned outlets.

[47] Mathewson and Winter (1985) identify both of these principal-agent problems in franchising, but do not analyze dual distribution as a way to alleviate them.

[48] Norton (1988) notes that when information on the skills of unit managers is privately held, then franchise contracts may be an effective screening device.

Appendix

Proof that no pooling equilibrium satisfies the Intuitive Criterion

Consider a pooling equilibrium in which the franchisor, regardless of type, offers r^p, F^p, and α^p. The franchisee can infer nothing about the franchisor's type from this offer, and so continues to believe that the franchise is of high type with probability ρ. If the franchisee accepts the franchise, he will produce $x_\rho(r^p)$ units, where

$$x_\rho(r) = \text{argmax} \ \ (1\text{-}r)\ (\rho R_H(x) + (1\text{-}\rho)R_L(x)) - cx.$$

Of course, the franchisee will choose to accept the franchise only if his expected returns from doing so are positive.

In this equilibrium, the high type franchisor will earn returns of

$$\alpha^p W_H^{c*} + (1\text{-}\alpha^p)\ (r^p R_H(x_\rho(r^p)) + F^p).$$

The returns to the low type franchisor depend on the number of units actually sold: define r such that $x_\rho(r) = x_H(r')$ as r''. Note that $r' > r''$. Then the returns to the low type franchisor are

$$\alpha^p W_L^{c*} + (1\text{-}\alpha^p)\ (r^p R_L(x_\rho(r^p)) + F^p) \quad \text{if } r^p \geq r'',$$
$$\alpha^p W_L^{c*} + (1\text{-}\alpha^p)\ (r^p R_L(x_\rho(r'')) + F^p) \quad \text{if } r^p < r''.$$

The pooling equilibrium will fail to satisfy the Intuitive Criterion if there exists a strategy (r,F,α) such that (i) the equilibrium payoff to the H-type franchisor is less than the payoff this type of franchisor would receive by deviating to (r,F,α) if the franchisee believes upon observing this deviation that the franchisor is indeed H-type, and (ii) the equilibrium payoff to the L-type franchisor is greater than the payoff under (r,F,α) no matter what the franchisee believes upon observing the deviation. In this model, the maximum payoff to the L-type under this or any strategy is obtained when the franchisee believes that the franchisor is H-type. Now note that the marginal revenue of a given level of output is always higher for the H-type than for the L-type. If then follows that there always exists some $F^d < F^p$ such that

$$\alpha^p W_H^{c*} + (1\text{-}\alpha^p)\ (r^p R_H(x_\rho(r^p)) + F^p) < \alpha^p W_H^{c*} + (1\text{-}\alpha^p)\ (r^p R_H(x_H(r^p)) + F^d) \text{ while}$$
$$\alpha^p W_L^{c*} + (1\text{-}\alpha^p)\ (r^p R_L(x_\rho(r^p)) + F^p) > \alpha^p W_L^{c*} + (1\text{-}\alpha^p)\ (r^p R_L(x_H(r^p)) + F^d) \text{ if } r^p \geq r',$$
$$\alpha^p W_L^{c*} + (1\text{-}\alpha^p)\ (r^p R_L(x_\rho(r^p)) + F^p) > \alpha^p W_L^{c*} + (1\text{-}\alpha^p)\ (r^p R_L(x_H(r')) + F^d) \text{ if } r' > r^p > r'',$$
$$\alpha^p W_L^{c*} + (1\text{-}\alpha^p)\ (r^p R_L(x_\rho(r'')) + F^p) > \alpha^p W_L^{c*} + (1\text{-}\alpha^p)\ (r^p R_L(x_H(r')) + F^d) \text{ if } r^p \leq r''.$$

Therefore, r^p, F^d, and α^p is a deviation that demonstrates that the intuitive criterion cannot be satisfied.

__Proof of Proposition 1__

Rewrite the separation constraint as:

(A1) (i) $F = \pi_L^* + \alpha\Delta V_L/(1-\alpha) - rR_L(r') - g(e^*) - \Delta V(e^*)$ for $r \le r'$

(ii) $F = \pi_L^* + \alpha\Delta V_L/(1-\alpha) - rR_L(r) - g(e^*) - \Delta V(e^*)$ for $r > r'$,

and the boundary of the rationality constraint as:

(A2) $F = (1-r)R_H(r) - cx_H(r) - g(e^*)$.

The slope of (A2) is:

(A3) $dF/dr = - R_H(r)$.

Let $mr_i(x_H)$ be the marginal revenue of the i-type demand when output is set equal to $x_H(r)$ and $mr_i' = \partial mr_i/\partial x$. Then, the slope of (A1) is:

(A4) (i) $dF/dr = - R_L(r')$ for $r \le r'$

(A4) (ii) $dF/dr = - R_L(r) - r\, mr_L(x_H)\, dx_H/dr$ for $r > r'$.

Since $dx_H/dr = mr_H(x_H)/(1-r)\, mr_H'$ which is clearly less than 0, the slope of (A1) is greater than (A2) for every r. Inspection of (A3) and (A4ii) reveals this to be the case for all $r > r'$, since $mr_L(x_H)$ is positive in this range. Also, (A4i) is larger than (A3); $R_L(r') < R_H(r') < R_H(r)$ since $mr_H(x_H)$ is positive for all $0 < r \le r'$.

The fact that the separation constraint is flatter than the rationality constraint implies that (A1) lies everywhere above (A2) for $\alpha > \hat{\alpha}$, where $\hat{\alpha}$ is defined by:

$$\pi_L^* + \hat{\alpha}\Delta V_L/(1 - \hat{\alpha}) - \Delta V(e^*) = \pi_H^*$$

or $$\hat{\alpha} = (\Delta\pi + \Delta V(e^*))/(\Delta V_L + \Delta\pi + \Delta V(e^*))$$

where $\Delta\pi = \pi_H^* - \pi_L^*$ and $\Delta V_L = [V_L(e^*)-g(e^*)] - [V_L(\underline{e})-g(\underline{e})]$.

Hence, for $\alpha > \hat{\alpha}$, the separation constraint does not bind. When $\alpha \ge \hat{\alpha}$, profits of the H-type franchisor is maximized subject to the rationality constraint at $r = 0$ and $F = \pi_H^* - g(e^*)$. To see this, note that the H-type's problem is to maximize:

(A5) $\alpha(\pi_H^*-g(\underline{e})+V_H(\underline{e})) + (1-\alpha)(F+rR_H(r)-cx_H(r)-g(e^*)+V_H(e^*))$

subject to the rationality constraint. Note that the slope of the high demand isoprofit curve is flatter than the slope of the rationality constraint in (A3) for all r, where the slope of the isoprofit curve is:

(A6) $dF/dr = - R_H(r) - r\, mr_H(x_H)dx_H/dr$.

This implies that profits are maximized at $r=0$ and $F=\pi_H^* - g(e^*)$.

Next, we show that for $\alpha < \hat{\alpha}$, $r > 0$. To see this, suppose $r = 0$. Inspection of (A1) and (A2) reveals that only the separation constraint binds at $r=0$. Furthermore, the slope of the iso-profit curve at $r=0$ from (A6) is $- R_H(0)$ and the slope of the separation constraint from (A4) is $- R_L(0)$. Since the isoprofit curve is steeper than the separation constraint, the optimal royalty is strictly positive.

Finally, we show that the separation constraint is weakly decreasing at the optimal r. By Assumption 1, maximization of the H-type's profits with respect to the separation constraint alone yields a unique maximum. Call the tangency between (A1) and (A5) $\bar{r}$. Then, the optimal r is either given by $\bar{r}$ (when only the separation constraint binds), or by the intersection of the constraints in (A1) and (A2) (when both constraints bind). If $\bar{r}$ lies on the decreasing portion of the separation constraint, then so does the constrained optimum. Under Assumption 1, from (A4) and (A6) $\bar{r}$ solves:

(A7) (i) $R_H(r) + r\, mr_H(x_H)\, dx_H(r)/dr = R_L(r')$ if $r \leq r'$.

(ii) $(a_H - a_L)\,(x_H(r) + r\, dx_H(r)/dr) = 0$ if $r > r'$.

If $\bar{r}$ is given by (A7i), then the slope of the separation constraint, given by (A4i), is clearly negative. Now suppose $\bar{r}$ is given by (A7ii). Then, evaluating (A4ii) at $\bar{r}$ from (A7ii) and noting that the H-demand is more price elastic than the L-demand for any given x, we see that the slope of the separation constraint at $\bar{r}$ is strictly negative.

Proof of Proposition 2

First note that an increase in a_H shifts the intersection of the rationality and separation constraints to the right. So if the rationality constraint binds, the result in the proposition goes through. Second, if the rationality constraint does not bind, then the optimum will be defined by the tangency of the isoprofit curve and separation constraint. (See proof of Proposition 1.) As we show below, this tangency moves to the right under Assumption 1. These two features imply that an increase in a_H results in a higher r; moreover, given result (ii) in Proposition 1, F decreases in a_H.

To show this, first note that if the rationality constraint is not binding for some a_H, then the optimal r is $\bar{r}$ given by (A7). By Assumption 1, since the right-hand side of each part of (A7) is independent of a_H, if an increase in a_H for a given r results in an increase in the left-hand side, then $\tilde{r}$ will increase for an increase in a_H. It suffices to show that each of the following increase in a_H:

(A8) (i) $R_H(r) + r\, mr_H(x_H)\, dx_H(r)/dr$

and (ii) $x_H(r) + r\, dx_H(r)/dr$.

Then for demands in Assumption 1, the following must be true:

(A9) (i) $mr_H(r)\, \partial x_H(r)/\partial a_H + x_H(r) + r\, mr_H'\, (\partial x_H/\partial a_H)\, (dx_H(r)/dr) + r\, mr_H(x_H)\, (\partial^2 x_H/\partial r \partial a_H) > 0$

(ii) $\partial x_H/\partial a_H + r\, \partial^2 x_H/\partial r \partial a_H > 0$.

The franchisee's first-order condition reveals that $\partial x_H/\partial a_H = -\, 1/mr_H'$. Substituting this into (A9i) and rearranging terms reveals that (A9ii) implies (A9i). Then, it is straightforward, but tedious, to show that under Assumption 1, (A9ii) holds.

Proof of Lemma 1

To prove Lemma 1 we show (i) For a_H sufficiently large, $\tilde{r} \le r'$. (ii) When $\tilde{r} \le r'$ and $\alpha = 0$, the rationality constraint does not bind at $\tilde{r}$. To show (i), note that if and only if $\tilde{r} \le r'$, the slope of the H-type isoprofit curve is greater than the slope of the separation constraint at r'. For the demands in Assumption 1, (A6) and (A4) imply that we can rewrite this condition on the relative slopes as

(A10) $\Delta a\, (x_H(r') + r'(\partial x_H(r')/\partial r)) < 0.$

To show that (A10) holds for a_H sufficiently large, first note that by Assumption 1 $r' = 1 - c/\Delta a$. Thus when $\Delta a = c$, $r' = 0$ and the LHS (left-hand side) of (A10) is strictly greater than 0; when $a_H = \infty$ then $r' = 1$ and the LHS is strictly negative. The latter is true since $dx_H(r')/dr = mr_H(x_H)/(1-r')mr_H'$; so as $r' \to 1$, $dx_H(r')/dr \to -\infty$. Hence, if we can show that the LHS decreases monotonically in a_H then there exists an $\underline{a}_H$ such that for $a_H > \underline{a}_H$, $\tilde{r} \le r'$.

Furthermore, since $\partial x_H(r')/\partial r = mr_H/((1-r')mr_H'$, the LHS of (A10) becomes:

(A11) $\Delta a\, (x_H(r') + (\Delta a - c)\, \Delta a/(c\, mr_H'))$.

Taking the derivative of (A11) with respect to a_H, and noting that $\partial x_H(r')/\partial a_H = 0$ and $\partial mr_H'/\partial a_H = 0$, we have

(A12) $x_H(r') + (3\Delta a - 2c)\Delta a/c\, mr_H'$.

Since $x_H(r')$ is constant in a_H and the second term is negative, (A12) implies that $\tilde{r} \le r'$ for a_H sufficiently large.

To show (ii) above, it suffices to show that the rationality constraint does not bind at

r' when $\alpha = 0$. This will be true when the rationality constraint is above the separation constraint at $r = r'$ and $\alpha = 0$. That is,

(A13) $\quad (1-r')R_H(r') - cx_H(r') > \pi_L^* - r' R_L(r') - \Delta V(e^*)$.

Since $r' = 1 - c/\Delta a$, this becomes

(A14) $\quad R_H(r') - cx_H(r') - (\Delta a - c)\, x_H(r') > \pi_L^* - \Delta V(e^*)$.

(A14) will hold if $R_L(r') > \pi_L^*$, which is always true. Hence the rationality constraint does not bind when $\bar{r} \leq r'$ and $\alpha = 0$.

Proof of Lemma 2

First note that for $\alpha < \tilde{\alpha}$, $dr(\alpha)/d\alpha = 0$, because $r(\alpha)$ remains constant at $\bar{r}$. Substituting in for W_H^{C*} and $W_H^f(\bar{r},F)$ in (10) from (2) and (5) yields

(A15) $\quad \pi_H^* - V_H(\underline{e}) - g(\underline{e}) - \bar{r}R_H(x_H(\bar{r})) - F - V_H(e^*) + (1-\alpha)dF(\alpha)/d\alpha$.

By differentiation of (A1), $dF(\alpha)/d\alpha = \Delta V_L/(1-\alpha)^2$. Also, F must be small enough to satisfy the rationality constraint (A2):

$$F \leq (1-\bar{r})R_H(\bar{r}) - cx_H(\bar{r}) - g(e^*).$$

But then it must be the case that (A15) is greater than

(A16) $\quad \pi_H^* - \pi_H(x_H(\bar{r})) - \Delta V_H + \Delta V_L/(1-\alpha)^2$.

This implies that (A15) will be strictly positive if $\Delta V_H \leq \Delta V_L$. A similar proof goes through for the case of $\bar{r} > r'$.

Proof of Proposition 4

To see this, first rewrite (10), substituting for W_H^{c*} and $W_H^f(r,F)$ from (2) and (5) and for F from (A1); for notational simplicity, suppress the dependence of r and F on the optimal α. Then the first-order condition for α^* is

(A17) $\quad [(\pi_H^* - rR_H(r)) - (\pi_L^* - rR_L(r))] + [V_H(\underline{e}) - V_L(\underline{e})]$
$$+(1-\alpha)[(d\{r(R_H(r)-R_L(r))\}/dr)\ (dr/d\alpha)] = 0.$$

Evaluate this at $\hat{\alpha}$. At $\hat{\alpha}$, the optimal royalty equals 0. Then, the first two terms in (11) are given by:

(A18) $\quad \Delta\pi + V_H(\underline{e}) - V_L(\underline{e})$.

For the third term in (A17), first note that for α such that $\hat{\alpha} \geq \alpha \geq \tilde{\alpha}$, the optimal r is given by the intersection of the separation and rationality constraints. That is, r solves:

(A19) $\pi_L^* + \alpha\Delta V_L/(1-\alpha) - \Delta V^* - rR_L(r) = (1-r)R_H(r) - cx_H(r)$.

Then, $dr/d\alpha$ is found by total differentiation of (A19). At $\alpha = \hat{\alpha}$, this gives:

(A20) $dr/d\alpha = - \Delta V_L/\{(1-\hat{\alpha})^2) (R_H(0)-R_L(0)\}$.

Then, multiplying (A20) by $(1-\hat{\alpha})$ and $d\{r(R_H(r)-R_L(r)\}/dr = R_H(0)-R_L(0)$ (when evaluated at $r=0$) gives the third term in (A17):

(A21) $- \Delta V_L/(1-\hat{\alpha})$.

The expression for $\hat{\alpha} = (\Delta\pi + \Delta V(e^*))/(\Delta\pi + \Delta V(e^*) + \Delta V_L)$ is substituted into (A21) and the revised expression is added to (A18) to get $- \Delta V_H$. Hence, (A17) is negative when evaluated at $\hat{\alpha}$.

Proof of Proposition 5

The first part of the proposition follows directly from Lemma 2. To show the second part, note that for rents to be shared in equilibrium, it must be the case that the left-hand derivative of the profit function in (A17) with respect to α, evaluated at $\bar{\alpha}$, must be negative. That is, in equilibrium the following must hold:

(A22) $\pi_H^* - \pi_L^* + V_H(\underline{e}) - V_L(\underline{e}) - \bar{r}\,(R_H(\bar{r}) - R_L(\bar{r})) < 0$.

But since (A22) is independent of α for all $\alpha \leq \bar{\alpha}$, (A22) implies that $\alpha^*=0$.

References

Brickley, James A., and Frederick H. Dark, "The Choice of Organizational Form: The Case of Franchising," Journal of Financial Economics, vol. 18, 1987, pp 401-420.

Caves, Richard E., and William F. Murphy III, "Franchising: Firms, Markets, and Intangible Assets," in Southern Economic Journal, vol. 42, April 1976, pp 572-586.

Cho, In-Koo, and David M. Kreps, "Signaling Games and Stable Equilibria," Quarterly Journal of Economics, vol 102, 1987, pp 179-221.

Entrepreneur, January 1990.

Gallini, Nancy T., and Brian Wright, "Technology Transfer under Asymmetric Information", Rand Journal of Economics, vol. 21, no. 1, Spring, 1990, pp. 147-160.

Grossman, Sanford J. and Oliver D. Hart, "The Costs and Benefits of Ownership: A Theory of Lateral and Vertical Integration," Journal of Political Economy, Vol. 94, August 1986, pp 691-719.

Lafontaine, Francine, "Contract Theory and Franchising: Some Empirical Results," Carnegie Mellon Graduate School of Industrial Administration Working Paper, December 1988.

Lafontaine, Francine, "An Empirical Look at Franchise Contracts as Signaling Devices," Carnegie Mellon Graduate School of Industrial Administration Working Paper, February 1990.

Leland, Hayne E. and David H. Pyle, "Informational Asymmetries, Financial Structure, and Financial Intermediation," Journal of Finance, vol 32, May 1977.

Lutz, Nancy A. "Warranties as Signals under Consumer Moral Hazard", Rand Journal of Economics, vol. 20, no. 2, Summer 1989, pp. 239-255.

Martin, Robert E., "Franchising and Risk Management," American Economic Review, vol.78, December 1988 pp 954-968.

Mathewson, G. Frank, and Ralph A. Winter, "The Economics of Franchise Contracts," Journal of Law and Economics, vol 28, October 1985, pp. 03-26.

Milgrom, Paul and John Roberts, "Price and Advertising Signals of Product Quality", Journal of Political Economy, vol. 94, 1986, pp. 796-821.

Norton, Seth W., "An Empirical Look at Franchising as an Organizational Form," Journal of Business, vol. 61, 1988, pp. 197-217.

Rubin, H., "The Theory of the Firm and the Structure of the Franchise Contract," Journal of Law and Economics, vol. 21, April 1978, pp. 223-233.