

Dual Organizational Structure of Franchise Contracts

K.J. Wainwright
Simon Fraser University

May 29, 2002

1 Introduction

In service based industries one of the fastest growing forms of market structure is that of franchise agreements. Certain aspects of franchise contracts tend to be idiosyncratic in nature thereby attracting a great deal of interest by academics and business analysts in recent years. Various explanations have been proposed for the widespread use of franchise contracts in certain industries. While a great deal of the franchise contract has been explained in the literature, there remains certain aspects of this form of arrangement that has yet to be addressed. This paper intends to address two of these issues as well as proposing an alternative modelling approach to franchise contracts.

The second section of this paper describes the basic structure of franchise contracts. The third section discusses the various explanations that have been proposed to explain franchising. The fourth section sets two aspects of the franchise contract that has not been addressed in the literature. The first of these is existence of both corporate owned outlets and franchised outlets within the same organization. Some authors have predicted that one form or the other would come to dominate the organization. Others have tried to explain under which conditions one form would be preferred by the parent company (or Franchisor). Yet many organizations exist as a mixture of both types of contracts and have chosen both forms of contract when expanding the number of outlets. The second unexplained observation is apparent rigidity in various organizations' franchise fee structure; both over time and between individual franchisees. The fifth section presents a model that proposes that both these observations can be explained when franchise agreements are modelled in a spatial context. The sixth section presents a digression into a dynamic model that uses optimal control theory to analyze explicitly the length of term of a franchise agreement, addressing attributes of franchise contracts designed to influence the term of commitment by the franchisee. A summary of empirical evidence is found in the Appendix.

2 Structure of the Franchise Contract

A basic result derived in modern property rights literature is that when any given set of rights is exchanged, the principals involved will select the institutional framework that minimizes the sum of production and transaction costs¹. The most commonly observed of these arrangements (or governance structures) are price mediated markets and centralized employment within firms². These are not the only forms of arrangement within which transactions are carried out, and the distinction between the two mentioned above is not as clear as it is suggested. An example of an alternative institutional framework is a franchise arrangement, and the purpose of this paper is to analyze the nature and purpose of franchise contracts.

In a franchise contract, a parent company contracts out the right to produce or market its product to an agent. Contractual stipulations involve rules governing the behavior of the agent including pricing, mode of production, and territorial or market restrictions. A frequently observed feature of a franchised industry is that certain aspects of the parent company's product have limited scale economies that require production at the local market level.

A principle characteristic of franchise contracts is the agent's right to use a national brand name in exchange for a share of the profits. The brand name is a signal to consumers in a local market that the agent supplies a product of a certain quality. The effectiveness of the brand name as a quality signal will decide its value to consumers. Given the nature of brand names and the characteristics of certain industries that rely on them, franchise contracts as a form of governance structure may be the most efficient for enhancing and protecting the value of the brand name.

¹Williamson, O. E., "Transaction Cost Economics: the Governance of Contractual Arrangements", *The Journal of Law and Economics*, 22, Oct. (1979) 223-261

²Cheung, S.N.S. "The Contractual Nature of the Firm," *The Journal of Law and Economics*, 26 April (1983) 1-21.

Franchise contracts have certain common characteristics³. The franchisor sells or leases the right to produce or sell some product to a franchisee. Written into the contract are various obligations and commitments required by both parties.

First, with the right to use the franchisor's brand name, the franchisor also agrees to supply various types of assistance. This includes orientation with the production process, managerial and accounting assistance, site selection and development, and any ongoing assistance or advice, as required. The franchisor also takes responsibility for national marketing and advertising also any research and development of the product. Second, the franchisee agrees to operate the business in the manner stipulated by the franchisor. This includes hours of operation, pricing scheme, inventory levels, and adherence to the operating manual – if one is supplied. Third, the franchisee agrees to pay royalties to the franchisor. This is usually in the form of a non-linear outlay schedule, comprised of a fixed fee plus a share of the revenues.

Fourth, there will be a monitoring and auditing clause in the contract. This may be spelled out explicitly, but will usually give the franchisor arbitrary and discretionary power. Fifth, the contract will have a termination clause. The termination clause will heavily favour the franchisor who can practically end at will. The franchisee, on the other hand, also can terminate, but at unfavourable terms, usually incurring a heavy penalty. Finally, the contract will contain miscellaneous clauses dealing with sale of the franchise, rights of heirs, territorial restrictions and any other conditions that may be specific to the particular product.

3 Explanations Of Franchising

3.1 *Franchising As a Method of Capital accumulation*

It was believed that franchising first arose as a form of capital accumulation and rapid expansion⁴. This line of reasoning can be discredited on two accounts. First, if an

³See, for example, Rubin, P. "The Theory of the Firm and the Structure of the Franchise Contract," *Journal of Law and Economics*, 21 (1978) 223-233; or Caves, R.E. and Murphy, W.F. "Franchising: Firms, Markets and Intangible Assets," *Southern Economic Journal*, 42 (1976)

⁴See, for example: Hunt, S.D. "The Trend Toward Company-owned Units in Franchise Chains," *Journal of Retailing*, vol. 49, 2 Summer (1973), "Firms often choose the route of franchised units because they simply do not have access to the capital required . . ."; Caves and Murphy, *Supra note 3*, "For financing outlets the capital supplied by franchisees has no ready substitute. . .".

individual is to buy a franchise, he bears all the risk (uncertainty of the residual claim) of that one outlet, whereas the franchisor has his risk spread across all outlets. To bear this higher risk, a risk averse franchisee will demand a higher risk premium (share of the profits). The franchisor could therefore put together a package of shares from all the outlets, and sell them to the individual store managers. The franchisor thus lowers the risk premium he must pay while maintaining full control of the outlets. Being the less costly arrangement, this form of organization will dominate.

Second, franchisees tend to have little or zero wealth. Therefore, the funds they invest in a franchise must be acquired. With imperfect capital markets, it is unlikely that an individual would be more successful at raising the needed capital than an already established firm. Therefore, capital accumulation is not an adequate explanation of franchising⁵.

3.2 *Franchising to Ensure Agent Compliance*

A brand name is a mechanism by which certain measures (but not usually all) may be foregone⁶. The brand name provides an implicit guarantee of a certain level of quality, and as such removes the necessity of prospective consumers assessing the level of desirable attributes about the product. Since it is these attributes that are compared to relative prices in the consumption decision, and the need for measurement consumes resources, removing this need lowers the effective per unit price of the desired attributes. This characteristic of the brand name is analogous to the removal of an ad valorem tax, and can be represented by a rightward shift in the demand curve, the shift being the size of the measurement cost.

Having established a brand name, the benefits that are described above can be accrued at zero marginal cost, the establishment being effectively a sunk cost (although there may still be per period fixed costs in maintaining brand name status to some extent). It is natural, then, for a company to want to expand output, taking advantage of these large economies in the sale of brand name reputation. Such expansion will be subject to certain limitations. Technologically there may be diseconomies of scale, and in a spatial context the market will eventually become

⁵Rubin, P. *Supra note 3*.

⁶The need to establish a brand name is based on what Barzel calls "excess measurement", where the free attributes of a transaction are dissipated through excess measurement. {See Barzel, Y. "Measurement Costs and the Organization of Markets," *The Journal of Law and Economics*, 25, April (1982) 27 – 48.}

saturated. To overcome these constraints it becomes necessary to develop subsidiaries, each of optimal plant size, and each unlimited by the extent of its market.

If the brand name is successful in reducing excess measurement, then competitive forces will be mitigated, allowing the possibility of shirking by subsidiaries' employees. This may not be a problem if the output is clearly defined and straightforward to monitor. But, if the product has (intangible) attributes that are difficult to assess upon inspection, then monitoring may become prohibitively costly, the firm having great difficulty in fully monitoring the performance of subsidiaries. This rules out the possibility of one large, vertically integrated firm, making necessary the choice of an alternative governance structure of which the franchise is an example. This contractual arrangement allows the required expansion, whilst also providing behavior constraints as disincentives to shirk.

Shirking can be defined in two ways; quality or quantity shortfalls. There is an incentive to produce output other than the one preferred by the franchisor, since the franchisee's marginal cost curve is usually different from that of the franchisor's. The profit maximizing level of output is determined at the point marginal revenue equals marginal cost. Having incurred sunk costs establishing the brand name, the franchisor's marginal cost (of brand name production per unit of output) is zero. In this respect the franchisor is a sales maximizer. The franchisee produces those attributes of the product that experience limited economies of scale, and therefore his marginal cost is positive and often rises as output is increased.

Quality shirking is a form of the free-rider problem. When brand names allocated to many local outlets, this free-riding problem takes two forms. The first is vertical free-riding on the national brand name, and results from the franchisee having better knowledge of the state of the local market than the franchisor. This form of free-riding always exists in franchise arrangements. The second form, horizontal free-riding, arises when a percentage of customers from any one of the outlets are transient in nature. This portion of customers base their demand on the average quality of all outlets visited, and not just on the quality level supplied by the outlet they happen to be patronizing. This allows an individual to free-ride on the quality level of other franchises. In both cases the benefits to quality reduction (reduced production costs) accrue only to the free-rider while costs of quality reduction (devaluing the brand name) are shared by the franchisor and other franchisees.⁷ The devaluation could be compounded if

transient customers base their demand solely on a visit to a free-riding outlet, by which the brand name becomes irretrievably associated with a sub-standard product.

In response to these problems the franchisor includes stipulations in the franchise contract designed to render shirking uneconomical (ruling out perfect monitoring). The franchisor could consider charging a price for the franchise equal to the present value of an outlet that is operated according to its aims. But such a price is very difficult to assess, and such behavior will create a moral-hazard problem: there will exist the possibility that the franchisor will not maintain brand name status consistent with the calculation of the present value, making a wind-fall gain at the expense of the franchisee. The sunk expenditure in establishing the brand name may be insufficient to convince franchisees that their principal is interested in a long-term relationship.

Quantity shortfalls are made undesirable by the imposition of a non-linear outlay schedule. The non-linear outlay schedule is comprised of both a fixed and variable component, the latter being tied to gross revenues. Under this system the franchisee's per unit royalty decreases as output is expanded. The non-linear outlay schedule effectively 'flattens-out' the franchisee's marginal cost curve, causing output to rise.

Given their different cost curves, the franchisee may want a quantity/quality combination that differs from what is optimal for the franchisor. The value of the brand name as a signal will depend on the extent that the franchisor can effectively enforce his quantity/quality combination. Therefore the franchise contract will attempt to restrict the franchisee's ability to make such trade-offs. Contract provisions that set hours of operation, prices and outlet design serve to deter this form of franchisee behavior. The franchisor is particularly concerned with quality chiselling, because it devalues the brand name he has expended much investment in attaining. With the existence of free-rider problems the franchisor will be forced to engage in monitoring. If the costs of monitoring are positive, this prohibits perfect monitoring, and this is assumed to be the case (otherwise the franchisor would have chosen an alternative contractual arrangement). The franchisor will, therefore, require some incentive structure to ensure quality compliance, to supplement the necessarily inadequate level of monitoring.

The franchisor could require that the franchisee put up a

 handled through assigning territorial rights to individual franchisees. Vertical free-riding requires monitoring plus a reward or penalty system. The importance of this distinction in explaining franchise contracts is explored in greater detail in Mathewson, F. and Winter, R. "The Economics of Franchise Contracts," *The Journal of Law and Economics*, Oct. (1985) 503-526.

⁷ Though the end result is the same from both forms of free-riding, the distinction is important to the nature of the contractual constraints used to remedy the problem. Horizontal free-riding can be

forfeitable bond that would be lost with non-compliance⁸. However, this creates a reverse moral hazard problem: if the bond is sufficiently large the franchisor may renege on his promise to maintain the brand name and abscond with the bond. Also, if the franchisee was sufficiently wealthy to afford an adequately sized bond, then he would invest in a more diversified, less risky asset than a franchise, with fewer constraints on his managerial sovereignty. This implies a wealth constraint on the franchisee; which is a necessary condition for a franchise contract⁹. Faced with wealth constrained franchisees, the franchisor will require a reward structure to ensure quality compliance. The reward will be such that the return to the franchisee from quality compliance exceeds the expected savings from quality reduction.

The actual level of monitoring thus will be decided by the relationship between its cost, and the levels of benefits and penalties described above. It also will depend on the attitudes to risk of the parties. For the purposes of this paper it is assumed parties are risk-neutral, other attitudes can be included as simple extensions. Therefore if greater than normal returns exist in an industry one would predict an inflow of franchisees¹⁰. Besides the sunk investment in the brand name the firm also incurs non-salvageable investment in individual franchisees. The firm must invest both time and resources in training the franchisee and developing the new outlet in a way that allows the franchisee to operate the business. This form of investment is necessary to attract potential franchisees who lack experience or knowledge in the particular industry.¹¹ Only those with such inadequate human capital

will offer themselves as franchisees, since for individuals possessing the necessary expertise, the benefits of using the brand name do not outweigh the costs (profit sharing and behavior constraints).

It is natural then, for franchise arrangements to have great appeal to individuals who lack sufficient wealth and human capital to establish an independent operation. Though franchisees must pay a large portion of their revenues to the franchisor, their expected value of the franchise exceeds that of a totally independent operation because of the high probability of success.¹² Once granted, the right to use the franchisor's brand name will usually serve as sufficient collateral to raise any money. In the event that the brand name is not sufficient collateral, the franchisor may supply necessary funding accordingly.

Franchise contracts are best explained as a solution to a monitoring problem when reputation is an important factor in the exchange of a good. The best model of a franchise contract is found in Mathewson and Winter¹³. Franchise contracts allow an agent to earn a quasi-rent stream from producing and/or selling a parent company's product in a local market. The purpose of the quasi-rent is to ensure compliance on the part of the agent to the terms of the franchise contract.

Klein and Murphy¹⁴ argue that quasi-rents are not sufficient to ensure appropriate agent behavior. They argue that it is necessary for the parent company to engage in active monitoring. It is the costs of monitoring relative to the quasi-rent stream that determines the degree of vertical integration within these industries. As monitoring costs fall one would expect to see corporate owned outlets rather than franchised outlets.

4 Spatial Issues of Franchise Contracts

There are two observed facts in industries that use franchising to produce and distribute their product that has not been adequately explained.¹⁵ The first is the breakdown between corporate owned and franchised outlets

⁸For further discussion on this form of constraint see: Klein, B. "Borderlines in Law and Economics: Transaction Cost Determinants of 'Unfair' Contractual Arrangements," *American Economic Review*, 70, 2 May (1980) 356-362.

⁹It is a lack of collateral that makes a franchise contract superior to any privately negotiated loan agreement a bank could offer the individual. A limited wealth condition is equivalent to a default option on loans to franchisees so that banks incapable of writing performance contracts superior to franchisors will rationally limit their loans to franchisees that ease the purchase of the local right to the brand name, knowing incentives in a franchise contract. The limited wealth constraint as a necessary condition for franchising is a well established result in the literature. See, for example Mathewson, F. and Winter, R. *Supra* note 7; or Rubin, P. *Supra* note 3.

¹⁰This turns out to be the case. Established franchise firms have queues of up to two or three years for the granting of a franchise licence. McDonald's accepts less than one percent of all applicants, and territorial rights are sold several years before actual construction of the outlet takes place, as prospective operators wait for natural population growth to reach a level that can supporting an outlet. (Kroc, R. **Grinding It Out: The Making of McDonald's**, Henry Regnery Co., Chicago, Illinois (1977).)

¹¹Seltz, D.D. *The Complete Handbook of Franchising* Addison-Wesley Publishing Company Inc. (1982)

¹²In 1973 only 2% of franchise outlets in the United States declared bankruptcy (see Vaughn, C.L. **Franchising** Lexington Books, Lexington Mass. (1974).

¹³Mathewson, F. and Winter R. *Supra* note 7.

¹⁴Klein, B. and Murphy, K. "Vertical Restraints as Contract Enforcement Mechanisms," *The Journal of Law and Economics*, Oct. (1988) 265-297

¹⁵Simon, Carol J., "Franchising vs. Ownership: a contracting explanation", *University of Chicago working paper* (1991). This paper presents the results of an extensive survey of franchise contracts across the midwest United States.

found within a given organization. It is frequently observed that an organization that engages in franchising will frequently buy back certain franchised outlets and operate them as corporate stores while at the same time issue franchises in new areas. Furthermore, there appears to be little correlation between the size of the quasi-rent that individual outlets are earning and the decision to buy them back.

The second unexplained observation is the fact that franchise fees remain relatively fixed, both across outlets and over time, while across outlets there is a wide variability in rents being earned. This fact appears to be inconsistent with the proposition that franchise fees allow the parent company to capture some of the economic rents being earned by the agent.¹⁶ Incentive compatibility constraints determine the extent that a parent company can capture the economic rents being earned by the individual outlets. If one assumed that individual franchisees have similar opportunity costs then one would expect that the quasi-rent required to ensure compliance would be the same across franchises. Therefore, if economic rents vary across outlets, the residual (minus the quasi-rent) would be captured by a *variable* franchise fee. One would expect the parent company to set each outlet's franchise fee based on local market conditions.

One characteristic common to franchise industries is that aspects production and distribution are carried out by many small, geographically displaced outlets. Therefore, when the parent company wishes to monitor its outlets, the monitor engage in considerable travel. In a large chain this will require the monitor to cover great distances in the execution of his duties. Therefore one would expect the remoteness of an outlet to have a bearing on the choice of contractual arrangement between the parent company and the local operator.

If the location of outlets and the distance between outlets is a function of market density, one would expect to see a clustering of outlets in more densely populated areas. This gives rise to an asymmetric distribution of stores which will have a significant effect on the costs of monitoring. If the monitor has to travel a significant distance to inspect a particular outlet, then frequent monitoring will be quite costly. However, if there is a second outlet in close proximity to the first outlet, then the marginal travel cost of monitoring the second store will be quite low.

This implies a non-convexity in the monitor's cost function that will effect the choice of contract between the parent company and the individual outlets. In the case of one outlet geographically displaced from the monitor it may be

more profitable to give the local agent a quasi-rent rather than frequent monitoring to ensure compliance. However, if a second store is established in close proximity to the first it may be more profitable for the parent company to switch to extensive monitoring and reclaim the quasi-rents.

While this point may seem straightforward with respect to the parent company's decision to franchise a new outlet, it implies something more. The decision to expand the number of outlets and the decision to change the form of the contract between the parent company and the local operator may be *two aspects of one decision*. This may explain why one form of contract has not come to dominate the other over time; something that has been predicted by analysts of these industries.¹⁷

With respect to the issue of fixed franchise fees, this too may be best explained in a spatial context. When a local market grows, so does the rents earned by the local franchisee. So why doesn't the parent company increase the franchise fee accordingly? One would expect that this would be a fairly straightforward clause to include at the outset of the franchise agreement.

It is assumed that the franchisee has better knowledge of local market conditions than the parent company. Therefore the franchisee would be in a better position to judge whether the local market could support expansion. In most franchise agreements the franchisee has the right of first refusal when a second outlet is being considered within his territory. The increase in economic rents accruing to the franchisee gives him the proper incentive to pursue expansion.

Furthermore, given diminishing returns to the ability of a single outlet to service a growing market, the parent company could better increase total royalty revenue from a given market by establishing a second outlet. The profitability of expansion will be further enhanced because of the non-convexity of the monitoring costs. The existence of the second store will lower the economic rents that were going to the first store before expansion. The lowering of rents will give the agent in the first store a greater incentive to shirk or alter his behavior in some way that is incompatible with the objective of the parent company. Therefore greater monitoring will be required. But, with the existence of the second store, the non-convexity of the monitoring costs may now make increased monitoring worthwhile relative to the pre-expansion period.

¹⁶See Tirole, J. *The Theory of Industrial Organization*, chapter 4 (1988).

¹⁷The list includes: Caves and Murphy *Supra note 3*; Hunt, S.D. "the Trend Toward Company-owned Units in Franchising in Franchise Chains", *The Journal of Retailing* vol. 49 (1973).

5 A Spatial Model of a Franchise

5.1 Initial conditions

A parent firm, or franchisor, sells his product in a set of geographically dispersed markets, or nodes. In each market there is an outlet where final production and sales are carried out by an agent, or franchisee. Demand conditions are assumed to vary across markets and each agent is assumed to have better information about local market conditions than the franchisor. The franchisor possesses a certain amount of monopoly power in each market.

At any given outlet the agent may be a franchisee or simply an employee of the franchisor. If the latter is the case, then the outlet is referred to as a corporate store. Denote the location, or address, of a local market by x_0 ($x_0 > 0$). The location of the franchisor will be normalized to be zero. Therefore x_0 represents the distance between the franchisor and the local market.

The franchisor produces a good at a constant cost of v per unit. The good is distributed by the agent to the local market. The agent also contributes additional input into the final good in the form of services or some other quality enhancing attributes. Let s denote the level of service provided by the agent and let $c(s)$ be the agent's cost of s where $c'(s) > 0$ and $c''(s) > 0$. Finally, each outlet incurs fixed cost of K .

Demand in the local market is a function of both price, p , and the level of services, s .¹⁸ Let q denote quantity demanded at location x_0 and the demand function is given by

$$q = D(s, p) \quad (1)$$

where

$$\partial D / \partial s > 0 \quad \text{and} \quad \partial D / \partial p < 0$$

The franchise contract specifies a payment schedule plus a level of s . The schedule for which the agent pays the franchisor royalties takes the form of a two-part tariff with a fixed and variable component:

$$f + \alpha(p - v)D(s, p) \quad (\text{where } 0 \leq \alpha \leq 1) \quad (2)$$

f is the lump-sum franchise fee and α is the share of sales revenue that accrue to the franchisor.

¹⁸Local demand is also a function of the strength of the national brand name. For our purposes, this is assumed exogenous and is therefore suppressed in the model.

5.2 The decision to shirk

Given the franchise contract, the agent may have an incentive to shirk on the level of services he is to supply. The decision to shirk will be a function of (i) the profits from shirking; (ii) the probability of detection by the franchisor; and, (iii) the penalty, or sanction from shirking. The probability of detection will, in turn, depend on the level of monitoring activity that the franchisor engages in and the degree by which the agent lowers the level of services below the contractually specified level.

Define ϕ as the frequency of monitoring carried out by the franchisor, which is normalized to be between 0 and 1. Furthermore, define Δs as the difference between the contractual level of s (denoted s^*) and the actual level of s supplied by the agent (i.e. $\Delta s = s^* - s$). Therefore the lower the actual level of services relative to the level specified in the contract, the greater will be Δs ($\Delta s = 0$ implies no shirking). The probability of the agent being detected shirking will be a function of both the frequency of monitoring and the degree of shirking by the agent¹⁹. Let δ denote the probability of detection, which can be expressed as follows:

$$\delta = \delta(\Delta s; \phi) \quad (3)$$

where

$$\partial \delta / \partial \phi > 0 \quad \text{and} \quad \partial \delta / \partial \Delta s > 0$$

In most franchise contracts the penalty for shirking is termination of the franchise agreement²⁰. Therefore the expected profit from supplying a low level of services can be expressed as

$$\pi_L = (1 - \delta(\Delta s; \phi))\pi(p, \Delta s) \quad (4)$$

where $\pi(p, \Delta s)$ is the agent's profits as he deviates from the contracted level of s .

Differentiating (4) with respect to Δs solves for Δs (and therefore s) that maximizes the agent's expected profits from shirking, or

$$(1 - \delta(\Delta s; \phi))\partial \pi / \partial \Delta s - \pi(p, \Delta s)\partial \delta / \partial \Delta s = 0 \quad (5)$$

¹⁹It is also possible that $s > s^*$, in which case the franchisee is supplying a level of service greater than the level specified in the contract. This may lead to *intra franchise* competition which lowers the franchisor's rents. Most franchise contracts will also attempt to minimize this form of behavior. For a more formal treatment, see Winter, R. A. "Vertical Control and Price versus Non-price Competition", *Working Paper* (1990).

²⁰See the Appendix for a summary of an empirical study of franchise contracts.

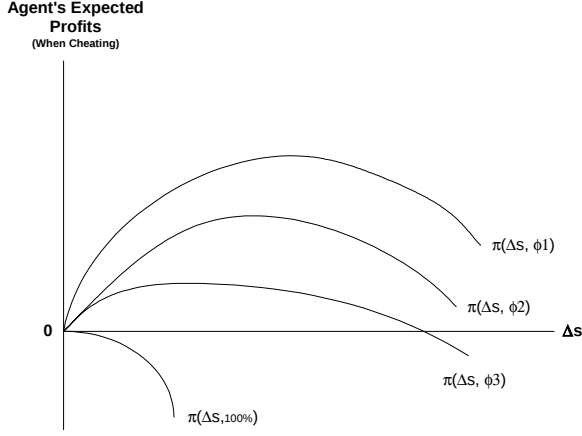


Figure 1: The agent's expected profit as a function of the level of shirking, given some known frequency of monitoring (ϕ). Changes in the frequency of monitoring will shift the agent's expected profit function.

Figure 1:

For any given level of ϕ the profit function of the agent (4) is at first increasing, then decreasing in Δs . Intuitively this results from the fact that as the level of service falls, the expected profits for the agent rises from the cost savings. However, as the level of services falls, the probability of detection rises, thus lowering expected profits. ϕ is a shift parameter in the expected profit function. Expected profits as a function of shirking on services (Δs) are illustrated in figure one.

Therefore, if $\pi_H = \pi(p, 0)$ is the profits of the agent when no shirking occurs, then the agent will choose to shirk if, at the s that maximizes (4),

$$\pi(p, 0) < (1 - \delta(\Delta s; \phi))\pi(p, \Delta s) \quad (6)$$

Equation 10 represents the incentive compatibility constraint faced by the franchisor.

If the franchisor decides to monitor the agent he will incur monitoring costs which are denoted as M . The costs of monitoring will be an increasing function of both the frequency of monitoring and the remoteness of the agent. Therefore the costs of verifying performance are

$$M = M(\phi, x_0) \quad (7)$$

where

$$\partial M / \partial \phi > 0 \quad \text{and} \quad \partial M / \partial x_0 > 0.$$

The Franchisor's objective is to maximize

$$\Pi(s, p) = f + \alpha(p - v)D(s, p) - M(\phi, x) \quad (8)$$

by choice of p, s, α, ϕ , and f subject to

$$\pi(p, s) = (1 - \alpha)(p - v)D(s, p) - c(s) - f - K \geq 0 \quad (9)$$

and

$$\pi(p, 0) \geq (1 - \delta(\Delta s; \phi))\pi(p, \Delta s) \quad (10)$$

Equation 9 is a non-negativity constraint on the agents profits²¹ and equation 10 is the incentive compatibility constraint. Let λ_1 and λ_2 denote the lagrange multipliers for equations (9) and (10) respectively. Applying Kuhn Tucker conditions and noting that (9) is non-binding in the presence of (10), we get the following results:

$$p(1 - 1/\epsilon_H) = v + k(1 - \delta)(p(1 - 1/\epsilon_L) - v) \quad (11)$$

$$\alpha(p - v)\partial D / \partial s = kc'(s) \quad (12)$$

$$\partial M / \partial \phi = \lambda_2((1 - \delta)\partial \pi_L / \partial \phi - \pi_L \partial \delta / \partial \phi) > 0 \quad (13)$$

$$\text{where } k = \frac{1}{\alpha/\lambda_2 + (1 - \alpha)} > 0$$

In equation (11), ϵ_H is the price elasticity of demand for a given s^* and ϵ_L is the price elasticity of demand when the agent chooses to shirk. Equation (11) implies that the price will be higher when the incentive to shirk is absent. Equation (12) determines s^* . If $\alpha < \lambda_2$ then the level of services will be set below the first best level. Equation (13) sets the level of monitoring and implicitly determines the rent stream accruing to the agent.

Since $\partial M / \partial x_0 > 0$, we can see from equation (13) that as the distance between the franchisor and the outlet increases, the level of monitoring will decrease and the rent stream to the agent will rise. This result is illustrated in figure two.

²¹ For simplicity, it is assumed that the agent's opportunity cost is zero.

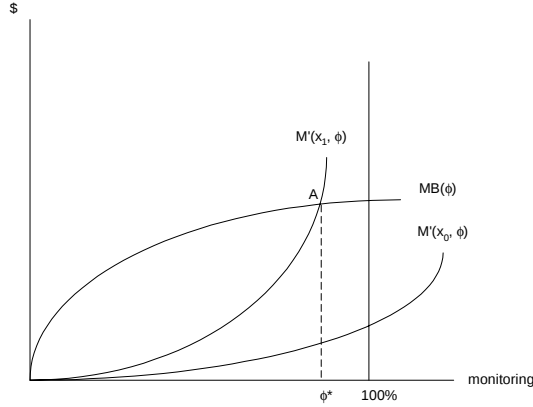


Figure 2: The marginal benefit of monitoring ($MB(\phi)$) is compared to the location specific marginal cost of monitoring. If the distance is small (x_0), it may pay to engage in 100% monitoring. If the distance is large enough (x_1), then there will be an interior solution of imperfect monitoring (point A).

Figure 2:

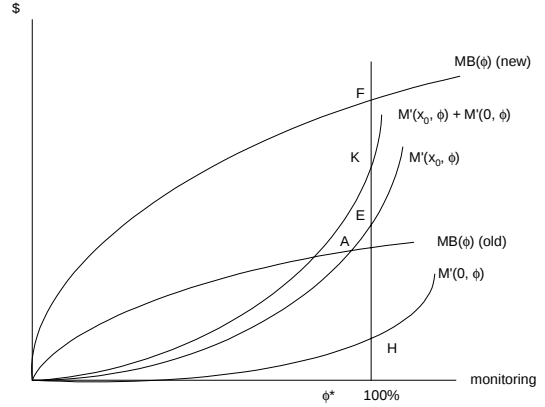


Figure 3: The marginal benefit of monitoring ($MB(\phi)$) shifts up due to growth in the market. If a second store is added to the local market, then its marginal cost of monitoring is lower than the first store since the cost of x is incurred monitoring the first outlet. Thus the marginal cost of monitoring both stores is less than twice the cost of monitoring the first.

Figure 3:

5.3 Expansion of the Market

Now at a certain point in the future the population allowed to grow. The increase in demand increases sales for the outlet. However, due to diminishing returns at the local level, the franchise is not able to fully supply the extra output at the given level of s . This will lead to an increase in the incentive to shirk. Therefore the franchisor will have to either increase the level of monitoring or allow the agent's rent stream to increase. The rent stream is implicitly increased whenever the fixed component of the franchise fee is held constant in the presence of growing demand.

The increase in demand may create an incentive for the franchisor to install a second outlet in close proximity to the first franchise. At this point the franchisor must decide whether it is more profitable to convert the franchises back to corporate owned and operated outlets or let them remain as franchise outlets. There are two factors that the franchisor must consider in making the decision to convert a franchise back to a corporate store. The first is the costs of monitoring two outlets which exist in close proximity to each other. The second is how the two outlets will interact while operating under a franchise arrangement.

5.4 The monitoring problem with Two Outlets

As before, the franchisor must travel to the outlet in order to engage in monitoring. Therefore the cost of monitoring the first outlet is $M(\phi, x_0)$. Now suppose a second outlet is located in the same market. Since the transportation costs must be incurred to monitor a single outlet at location x_0 , they become sunk costs, thus the cost of monitoring the second store is $M(\phi, 0)$ (where $M(\phi, 0) < M(\phi, x_0)$). The marginal cost of monitoring functions for each of the stores is illustrated in figure three.

In figure three $M'(\phi, x_0)$ intersects the original marginal benefit of monitoring schedule, MB_1 , at point A. The increase in demand shifts the marginal benefit schedule up to MB_2 . The marginal cost of 100% monitoring of the first store is given by point E. If a second store is also located at distance x_0 , then the cost of monitoring the second store is given by point H. The marginal benefit of 100% monitoring is given by point F. If the distance F to E is greater than the height to point H, then it will be worthwhile for the franchisor to convert the two stores to corporate outlets and engage in full monitoring. Regardless of the choice of contractual arrangement that the franchisor finally settles on, it is clear from figure 4 that

the marginal cost of monitoring a second store is less than that of a single store in the same geographical area.

In addition to the non-convexity in monitoring costs described above, there exists a further potential reduction in monitoring when there is a second outlet. The franchisor can use information from one outlet to infer local demand conditions of the second outlet. For example if one store reports high sales in the same market that the other outlet reports low sales, the franchisor may be able to form a better prior about the likelihood that the second store is shirking on quality rather than suffering from a random drop in demand.

6 Dynamic Extension

This section explores the intertemporal aspects of the franchise contract. As mentioned earlier, franchisees choose to enter into such agreements due to limited wealth constraints and capital market imperfections. However, if rents accrue to the agents then, over time, this condition will change. This change will have implications for the life of the contact.

For a given state of demand and a quality floor set below the first best level, at any given level of output the franchisee's profit function (π) at any given time is a direct function of the level of quality (s) at that particular time:

$$\pi(t) = \pi(s(t); p, v, \alpha, v, f, K) = \pi(s(t)) \quad (14)$$

where π is assumed to be a monotonically increasing function of $s(t)$ ($\pi' > 0$ and $\pi'' < 0$)²².

The marginal cost of $s(t)$ is assumed to be constant over time. p is assumed fixed for the purposes of this section's analysis. From his profits the franchisee must decide how much to reinvest into the production of future quality, and how much to retain as personal Income. The level of quality he decides to produce over time will be that which maximizes:

$$V = \int_0^T U(\pi(s(t)))e^{-\delta t} dt \quad (15)$$

Subject to:

$$Wo = \int_0^T \pi(s(t))e^{-rt} dt \quad (16)$$

where V is the franchisee's intertemporal utility function; Wo is the maximum present value of the income

²²percent Let F , the fixed fee be set equal to zero. This does not change the results since $dF/dt = 0$. It does, however, simplify the derivation.

stream over the life of the agreement. δ is the franchisee's internal rate of return. r is the market rate of interest.

The hamiltonian for the problem is:

$$H = U(\pi(s(t)))e^{-\delta t} - \mu\pi(s(t))e^{-rt} \quad (17)$$

from the maximum principle

$$(dU/d\pi)(d\pi/ds(t))e^{-\delta t} - \mu d\pi/ds(t)e^{-rt} = 0 \quad (18)$$

or

$$dU/d\pi = \mu e^{(\delta-r)t} \quad (19)$$

Equation 18 is the condition, in terms of $s(t)$, that maximizes the hamiltonian. μ is the shadow price of present utility in terms of forgone future utility. In equation 19 present marginal utility is set equal to the marginal user cost of future utility. Equation 19 implicitly defines the optimal path of quality production ($s(t)$) over time.

Given that utility, $U(\pi(s(t)))$, is a monotonically increasing function that exhibits diminishing returns, ($U' > 0$; $U'' < 0$) then, U will be inversely related to s and we can express $s(t)$ as:

$$s(t) = \Phi e^{(r-\delta)t} \quad (\Phi = \Phi(Wo, \delta) > 0) \quad (20)$$

If the franchisee's internal rate of return is less than the market rate ($r > \delta$) then the franchisee will be continuously increasing the quality of his outlet. If $r = \delta$, he will maintain the quality at some constant level. If $r < \delta$, then he will allow the quality to diminish. When franchisee first enters into a franchise contract, his opportunity costs are assumed to be zero. This is the result of the limited wealth and human capital he possesses. Limited or zero opportunity costs will be reflected in a low internal rate of return. The necessary condition for entering into a franchise agreement will be that $\delta \leq r$; consistent with increasing quality over time.

However, as the franchisee accumulates wealth and human capital, his opportunity costs will rise. Alternative investment opportunities (previously unattainable) offer the franchisee a higher internal rate of return. He will therefore discount the returns from future quality production at a much greater rate than at the outset of the contract.

When his discount rate becomes sufficiently high the franchisee will allow the quality level to deteriorate over time. Income retained by the franchisee for other (non-franchise) use is more highly valued than the additional increment in future revenues from further investment in future quality production. The optimal time paths of quality production are depicted in figure four.

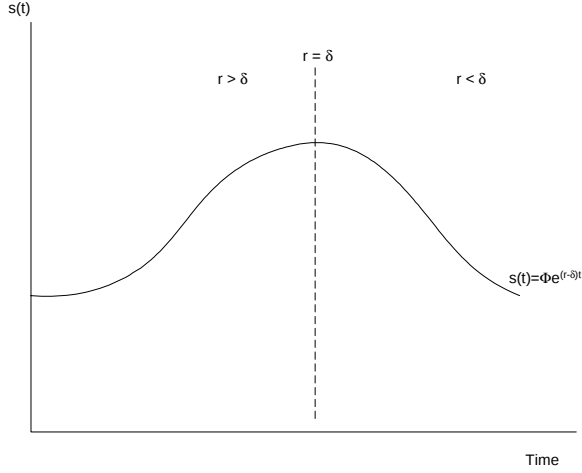


Figure 4: $s(t)$ is the optimal time path for quality (service) production by a utility maximizing agent with an increasing personal discount rate.

Figure 4:

The franchisor sets the quality floor, $\bar{s}$, from the solution to equation 8. If $s(t)$ is known explicitly then the length of the agreement can be determined by setting $s(t) = \bar{s}$ and solving for T . When the quality level deteriorates to the point it falls below the quality floor ($\bar{s}$) the contract will be terminated (see figure five). The quality floor is a function of the level of monitoring. Any increase in the marginal benefits of monitoring, or any decrease in the marginal cost of monitoring will raise the quality floor and reduce the life of the contract.

Because of the franchisee specific investment made by the franchisor, he will have a minimum contract duration that he will want the franchisee to honour. The franchisor will make adjustments in the contract up to the point that the marginal benefit of an additional period of commitment by the franchisee just equals the marginal cost of adjusting the agreement.

The franchisor cannot assess each potential franchisee, ex ante, as to his personal rate of time preference. In order to select franchisees whose discount rates are sufficiently low at the outset, the franchisor can use an initial licencing fee that will eliminate individuals with high discount rate.

Assume that an individual with $\delta \leq \delta^*$ will commit for the duration at or above the quality floor. Individuals

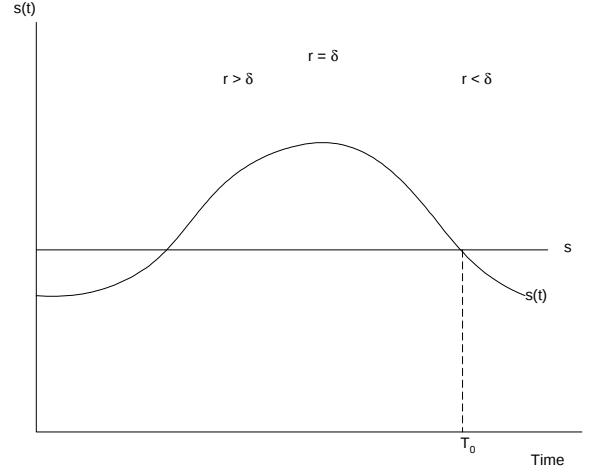


Figure 5: $\bar{s}$ is the quality floor set by the franchisor. When $s(t) < \bar{s}$ the franchise contract is terminated.

Figure 5:

with $\delta > \delta^*$ will violate the quality standard before the optimal T . The licence fee will be set according to:

$$L = \int_0^T (W_0 e^{-\delta^* t} - W_0 e^{-r t}) dt \quad (21)$$

where L is the licence fee paid up front (W_0 is defined in equation 19). Individuals with $\delta > \delta^*$ will not enter into the franchise contract because, given their higher discount rate, the differential between their valuation of the franchise and their next best alternative (discounted by r) has been removed by the imposition of the licence fee (See figure six).

The dynamic model gives other comparative "dynamic" predictions. If the quality floor increases then, in the absence of any increased reward to the franchisee, the duration of franchisee commitment will fall ($dT/d\bar{s} < 0$). Any action by the franchisor that reduces the franchisees opportunity costs will decrease δ and cause the $s(t)$ curve to shift upward. Increasing sunk investments required by the franchisee or increasing the cost of premature contract termination will reduce the franchisee's opportunity costs. Moreover, any increment in the initial net present value of the franchise (Φ at $t = 0$) will shift the $s(t)$ curve upward, extending the contract ($dT/d\Phi > 0$).

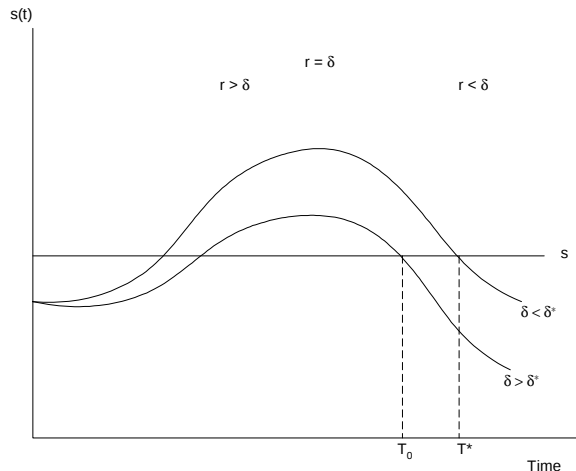


Figure 6: Time paths for two different personal discount rates. T^* is the franchisor's optimal contract length. Agents with $\delta \leq \delta^*$ will produce high quality or service to at least T^* .

Figure 6:

The effects of a change in the royalty, f , are ambiguous at this level of generality, to derive any further refutable propositions requires a more detailed specification of the utility and demand functions.

7 Conclusion

This paper has presented a simple model of a franchise contract. While capturing the essential elements of an incentive compatibility contract, the model is able to address some of the geographic issues inherent in franchise contracts. Specifically, the model focuses on the issue of the costs of monitoring to explain the contractual choices observed in franchise industry.

Several results are derived from the model. First, that in the presence of incentives to shirk and positive monitoring costs, increases in rents due to market growth may accrue to the agent rather than the franchisor. Second, the decision to expand are not independent of each other. When geographical considerations are taken into account, non-convexities in monitoring costs may arise that effect *both* the decision to expand and the decision to convert a franchise outlet to a corporate owned-store. Third, the results from the model questions the effectiveness of franchise fees at extracting economic rents being earned in the local market.

Finally, the franchisor will design the contract such that

it anticipates the changing opportunity costs of the franchisee. The franchise contract serves to govern the ongoing relationship between principal and agent, anticipating systematic changes between the two that occur during the life of the agreement. The franchisor will attempt to lower the franchisee's opportunity costs through a combination of contractual constraints and monetary incentives. Moreover, the franchisor will set the initial franchise fee in a manner that will result in self-selection of those potential franchisees with stronger commitment to the franchise.

This paper does not try to address all the complexities associated with the analysis of spatial market competition but instead illustrates the strength of this type of analytical framework.

The model in this paper is limited to the set of franchise contracts where some input on the part of the franchisee is a major component of the final product. The model does not apply to all forms of franchising observed in the economy, in particular franchise arrangements that are classified as manufacturer-retailer relationships. Such industries that experience large economies of scale in centralized production of the final product may find franchising simply an efficient method of delegating the responsibility of distribution.

8 VIII Bibliography

1. Barzel, Y. "Measurement Costs and the Organization of Markets," *The Journal of Law and Economics* , 25, April (1982) 27-48.
2. Caves, R.E. and Murphy, W.F. "Franchising: Firms, Markets and Intangible Assets," *Southern Economic Journal*, 42 572-586.
3. Cheung, S.N.S. "The Contractual Nature of the Firm," *The Journal of Law and Economics* , 26 April (1983) 1-21.
4. Demsetz, H. "The Exchange and Enforcement of Property Rights," *Journal of Law and Economics*, 7 Oct. (1964) 11-26.
5. Gallini, N. T. and Lutz, N.A. "Dual Distribution in Franchising" *JLEO* 1992
6. Goldberg, V.P. "Toward an Expanded Economic Theory of Contract," *Journal of Economic Issues*, 10, 1 March (1976) 45-61.
7. Helmers, H.O., Davisson, C.N., and Taggart, H.F. Two Studies in Automobile Franchising , The University of Michigan Ann Arbour, Michigan.
8. Hunt, S.D. "The Trend Toward Company-owned Units in Franchise Chains," *Journal of Retailing* , vol. 49, 2 Summer (1973) 3-13.
9. Klein, B. "Borderlines in Law and Economics: Transaction Cost Determinants of 'Unfair' Contractual Arrangements," *American Economic Review*, 70, 2 May (1980) 356-362.
10. Klein, B. and Leffler, K. "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 89, 4 (1981) 615- 641.
11. Klein, B. and Saft, L.F. "The Law and Economics of Franchise Tying Contracts," *Journal of Law and Economics* , 28 May (1985) 345-361
12. Klein, B. and Murphy, K. "Vertical Restraints as Contract Enforcement Mechanisms," *Journal of Law and Economics*, Oct. (1988) 265-297
13. Kroc, R. *Grinding It Out: The Making of McDonald's* Henry Regnery Co., Chicago, Illinois (1977).
14. Mathewson, F. and Winter, R. "The Economics of Franchise Contracts," *The Journal of Law and Economics*, Oct. (1985) 503-526.
15. Minkler, A.P. "Why Firms Franchise: A Search Cost Theory" *Journal of Theoretical and Institutional Economics* (1991)
16. Ozanne, U.B. and Hunt, S.D. *The Economic Effects of Franchising* (Washington, D.C.; U.S. Government printing office, 1971). Rubin, P. "The Theory of the Firm and the Structure of the Franchise Contract," *Journal of Law and Economics* , 21 (1978) 223-233.
17. Seltz, D.D. *The Complete Handbook of Franchising* Addison- Wesley Publishing Company Inc. (1982).
18. Simon, Carol J. "Franchising versus Ownership: a contracting explanation," *University of Chicago working paper* (1991)
19. Udell, G. "The Anatomy of the Franchise Contract," *The Cornell Hotel and Restaurant Quarterly* 13, no. 3 August (1972) 13-21.
20. Vaughn, C.L. *Franchising* Lexington Books, Lexington Mass. (1974).
21. Williamson. O.E. "Transaction Cost Economics: the Governance of Contractual Relations," *The Journal of Law and Economics* , 22, Oct. (1979) 223-261.
22. Winter, R.A. "Vertical Control and Price versus Non-price Competition" *working paper* (1990).