Abstract

The following paper is an example of the appropriate style, layout and format for a term paper or essay in an economics course. All papers should have a title page that contains the following:

1. Title of the Paper
2. Course Number and Instructor
3. Your name and student number
4. Date

Any graphs should be on separate pages that are not counted as part of the written page requirements. All graphs must be discussed and explained within the body of the text.

Newspaper and magazine articles may be cited but do not count as part of the required references.
A REVIEW OF FRANCHISE THEORY

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. This section introduces spatial or geographical considerations to the problem of franchising. When placed in a spatial context a testable hypothesis is proposed in which both of the issues identified can be explained.
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 costs. The most commonly observed of these arrangements (or governance structures) are price mediated markets and centralized employment within firms. 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.

---


Franchise contracts have certain common characteristics\textsuperscript{3}. 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.

EXPLANATIONS OF FRANCHISING

Franchising As a Method of Capital accumulation

It was believed that franchising first arose as a form of capital accumulation and rapid expansion\(^4\). This line of reasoning can be discredited on two accounts. First, if an 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 franchising\(^5\).

Franchising to Ensure Agent Compliance

A brand name is a mechanism by which certain measures (but not usually all) may be foregone\(^6\). The brand name provides an implicit guarantee of a certain level

\(^{4}\)See, 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 . . .".

\(^{5}\)Rubin, P. *Supra note 3*.

\(^{6}\)The 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.
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 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. The devaluation could be

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 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.
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.

An example of franchisee shirking on quality is illustrated in figure one. Suppose the local demand for the product is an increasing function of the franchisee’s service (or effort). Further suppose that the local demand is stochastic (volitile) such that there are both high demand (good days) and Low demand (bad days) states. This volitility is exogenous and independent of any efforts by the franchisee.

It is assumed that the local franchisee has better knowledge of local demand conditions and, in most cases, the franchisor relies on the franchisee to convey information
regarding local demand conditions. The franchise contract will specify the level of service the franchisee must provide. This is denoted as \( S^* \) in figure one. At the specified level of service, both parties expect that sales to fluctuate due to high and low demand states (points A and B in figure one).

However, to economize on local costs, the franchisee may choose to reduce the service below \( S^* \). In order to successfully reduce service without detection by the franchisor, the franchisee can choose a level of service in a high demand state that produces the sales associated with the low demand state (point N in figure one). The franchisee then "misdeclares" the state of demand to the franchisor.

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 windfall 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 forfeitable bond that would be lost with non-compliance. However, this creates a reverse moral hazard problem: if the bond is sufficiently large the franchisor may renego 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 contract. Faced with wealth constrained franchisees, the franchisor will require a reward structure to ensure quality compliance. The reward

---


9 It 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.*
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 franchisees. 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. 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

---

10 This 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).)

of the high probability of success. 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 Winter. 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 Murphy 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.

**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. The first is

---


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


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.
the breakdown between corporate owned and franchised outlets 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. 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

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).
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.

**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.

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.

**BIBLIOGRAPHY**


5. Gallini, N. T. and Lutz, N.A. "Dual Distribution in Franchising" JLEO 1992


