

Contract Mix in Franchising as an Efficient Monitoring Device under Asymmetric Information*

Thierry PENARD (corresponding author)
CREREG, University of Rennes I
Faculté des Sciences Economiques, 7 Place Hoche,
35065, Rennes cedex, France
Email: thierry.penard@univ-rennes1.fr

Emmanuel RAYNAUD
INRA SADAPT & Centre ATOM (U. of Paris I)
INRA SADAPT, 16 rue Claude Bernard,
75231, Paris cedex 05, France
Email: eraynaud@inapg.inra.fr

Stéphane SAUSSIÉ
University of Nancy II & Centre ATOM (U. of Paris I)
ATOM, Maison des Sciences Economiques, 106-112 Blvd de l'Hôpital,
75013, Paris, France
Email: saussier@univ-paris1.fr

This version: September 2003

* The authors would like to thank the French Ministry of Finance (Decas) and the Lorraine Region for their financial support. This paper is a revised version of a paper previously titled "Dual distribution and Royalty Rate in Franchised Chains: An Empirical Analysis using French Data". We would like to acknowledge James Brickley, Eric Brousseau, Vivian Dos Santos Silva, Josef Windsperger, Joyce Young for their useful comments.

ABSTRACT:

A strong regularity that emerges from empirical studies pertaining to franchising is the coexistence of franchised and company-owned units within the same chain (hereafter referred to as “dual distribution”). This paper supports the idea that dual distribution emerges as an efficient organization when both the abilities and the behaviors of the managers are costly to observe. We first show that partial monitoring - i.e. when only the franchisor monitors a subset of its outlets- is optimal for the franchisor. Then we show that the coexistence of franchised and company-owned units economizes on monitoring costs.

KEYWORDS: dual distribution, monitoring, asymmetric information.

JEL classification: D23., L14., L22.

1. INTRODUCTION

Franchising has received significant attention in the empirical and theoretical literature on contracting. Most of these papers have focused either on the determinants of contractual provisions, especially royalty rates (Mathewson and Winter, 1985, Lafontaine 1992, Lafontaine and Shaw 1999, Scott 1995, Brickley, 2002) or the tradeoff between franchising/company ownership (Brickley and Dark, 1987, Lafontaine, 1992, Brickley, 1999). Incentives and monitoring issues are crucial in this literature. For instance, Brickley and Dark (1987) emphasized the role of monitoring costs in the franchising/vertical integration tradeoff. Using geographical distance as a proxy for direct monitoring costs, they showed that outlets far from the franchisor's headquarters tend to be franchised. When monitoring the behavior of managers is costly, chains rely more on residual claimancy to reduce managerial incentives to shirk (see also, Norton, 1988, Lafontaine, 1992, Bercovitz, 2001, and Lafontaine and Slade, 2002, for a survey).

A puzzling empirical regularity, mentioned in most empirical studies, is the coexistence of franchised and company-owned units within chains (sometimes defined as "plural form", "contract mix" or "dual distribution"). Most of the literature deals with dual distribution as a transitory phenomena. For example, it is sometimes argued that franchising is more profitable than company ownership. However, at the beginning of their business, firms operate some units directly either to signal their type to potential franchisees (Gallini and Lutz 1992, Lafontaine 1993) or to credibly commit to protecting the value of their brand name (Scott 1995). One of the consequences is that the extent of company ownership should

decrease with chain maturity. On the other hand, other authors assert that the rationale for franchising should disappear with chain maturity (Oxenfeldt and Kelly 1969). Franchising only gives rise to temporary access to certain scarce resources, either capital (Caves and Murphy 1976), managerial talent (Norton 1988), or local information (Minkler 1990) that eases their expansion. As firms become established, they should reduce their reliance on franchising.

More recently, literature on franchising has suggested that dual distribution should be an efficient and persistent organizational form. Furthermore, empirical studies using panel data have shown the stability of this organizational mix (see for instance Lafontaine and Shaw, 2001, for US and Canadian chains). Several explanations have been put forward. Following Bradach and Eccles (1989), Bradach (1997) emphasized the complementarities between the two contractual arrangements in order to maintain quality and homogeneity of the business concept throughout the units while, at the same time, promoting innovation. He argued that “chain organizations are more than the sum of their parts: by having both company and franchise arrangements together, a chain can leverage some of the strengths and overcome some of the weakness associated with each arrangement” (1997, p. 279).

Lafontaine and Shaw (2001) explained the extent of company ownership in a chain as a way to exert direct managerial control in order to protect the brand name value against franchisee’s free riding. Lewin-Solomon (1999) justified the existence of dual distribution as a commitment device used by a franchisor to give franchisees incentives to innovate. In most chains, franchisees pay a royalty rate as a fraction of their revenues (and not their profit). With this royalty structure, chains will favor innovations that are good for revenues, sometimes at the expense of franchisees’ profits. By owning some units in the chain, the franchisor’s interest is more aligned with that of its franchisees. Bai and Tao (2000) constructed a multi-task model à la Holmström and Milgrom (1991) to show that dual

distribution is a way to induce effort towards both brand name development and sales. Finally, Sorenson and Sorensen (2001) explained franchise mix as the result of a trade-off between exploration (franchising) and exploitation (company-owned units) in organizational learning. In all these explanations, dual distribution appears to be an efficient arrangement.

This article provides another rationale concerning the efficiency of dual distribution. As in previous contributions, we emphasize monitoring costs as a major determinant of chain organization. The basic idea is that dual distribution enables the franchisor to economize on monitoring costs and to minimize informational rents allowed for an outlet's manager. In our model, the managers are heterogeneous in their abilities to run a store, and both managerial abilities and behaviors are imperfectly observed by the franchisor. A key issue is how to select "good" managers for running the stores and how to give them appropriate incentives to behave honestly and efficiently. The latter issue constitutes a big challenge since the franchisor faces two kinds of moral hazards depending on the statute of the managers. A salaried manager has an incentive to shirk and provide the minimum level of effort whereas a franchisee has an incentive to under-declare his actual revenues in order to minimize the royalties due. These opportunistic behaviors can be deterred by controlling or auditing the units. Literature on auditing is rather abundant and has dealt with many issues such as tax evasion, utility regulation or antitrust enforcement (Baron and Besanko 1984, Mookherjee and Png 1989, Khalil 1997). In this literature, the model of Besanko and Spulber (1989) can be fruitful for modeling of the organizational choices of a franchisor under asymmetric information. The authors characterized the optimal monitoring policy of an antitrust authority that imperfectly observes the production costs of the firms and their pricing behaviors. They proved that it can be optimal to tolerate minor collusion to some extent. The optimal policy consists in industry auditing when the observed market price is too high (above a threshold

price) and in renouncing an industry audit when the market price is reasonable (under the threshold price). The antitrust agency simply puts into balance the cost of industry auditing and the welfare gain of this audit through the deterrence of collusive behaviors. When the observed price is enough low, then it does not worth auditing the industry. Finally it is more efficient to tolerate minor opportunistic behavior in order to concentrate the monitoring efforts on more detrimental behaviors. The rationale for contract mix in franchising is based on similar arguments. We characterize the optimal monitoring strategy of a franchisor when he cannot observe both the ability and behavior of a unit's managers. The franchisor can either monitor all the stores in the chain or only a subset of stores. Like Besanko and Spulber, we find that complete monitoring of all the units is never optimal and that it is in the interest of the chain's owner to renounce controlling units that declare high revenues. Above all we demonstrate that dual distribution is more profitable than a pure franchised system or a pure company-owned system even if all the units are perfectly identical in matter of revenues.

Strikingly, most of the paper dealing with dual distribution explains the coexistence of franchised and company owned units by a heterogeneity of outlets (Lafontaine, 1992, Bai and Tao, 2000) such as distance from headquarters (Brickley and Dark, 1987), percentage of repeated business (Brickley, 1999), location specific factors (Chakrabarty and al., 2002). Our model complements this literature by developing an explanation based on the homogeneous outlet (but heterogeneous managers). In this respect, two papers are more similar to ours. Gallini and Lutz (1992) analyzed dual distribution as a signaling strategy, and Bai and Tao (2000) focused on moral hazard in a multitasking perspective, both papers taking into consideration the homogeneous units.

The paper is organized as follows. Section 2 presents a set of stylized facts about the organizational design of franchise systems. Section 3 presents the theoretical model. Section 4 expounds the main results. Section 5 concludes.

2. STYLIZED FACT: THE PERSISTENCE OF DUAL DISTRIBUTION

Before determining the grounds for dual distribution efficiency, we start by giving some empirical insights. Several papers have emphasized the fact that many chains used both franchised and company-owned units (Bradach, 1997). Interestingly, this pattern of organization was found for different countries. For instance, 70% of chains in France have a dual distribution form (Pénard, Raynaud and Saussier, 2003) whereas Furquim de Azevedo and Dos Santos Silva (2001) have noted the same proportion for Brazilian chains. In their study of the restaurant industry in the US, both Bradach (1997) and Lewin-Solomon (1999) showed that dual distribution was the dominant form in governing chains. Except for two chains (Dunkin' Donut and Subway), dual distribution is a widespread phenomena. We can also see that the extent of company-ownership varies from chain to chain. This diversity is interesting in itself but beyond the scope of the present paper (see Lafontaine and Shaw, 2001, for more on this subject).

As pointed out in the introduction, several explanations have been given for dual distribution either as a transitory or stable phenomenon. The persistence of dual distribution has, for quite some time, been an empirical puzzle. By studying the evolution and determinants of dual distribution in North America, Lafontaine and Shaw (2001) have largely contributed to the enlightenment of this issue. Strikingly, they found that the proportion of company stores tends to decrease in the first years of franchise and then stabilizes. This stability is interpreted as evidence that chains target a given proportion of company and franchised stores, called the managerial control target. Moreover, this target is rather chain specific, even if it is influenced by the activity sector and the experience acquired before franchising. Lafontaine and Shaw also showed that the value of a brand name increases the proportion of company owned units in chains. They analyzed it as a way to protect the brand

name value from franchisees' opportunistic behavior. By limiting the number of franchised stores, a chain can better protect its brand name investments. Similar empirical patterns have been observed in other countries (see Furquim de Azevedo and dos Santos Silva, 2001, for Brazil or Pénard, Raynaud, Saussier, 2003, for France). For example, figure 1 represents the evolution of company ownership with respect to franchising experience in the French case. This figure is interesting because it complements previous now well-known observed empirical regularities concerning the case of the US and Canada.

[INSERT FIGURE 1]

The shape is rather analog to the one exhibited by Lafontaine and Shaw (2001). After an initial decrease, the extent of company ownership tends to stabilize.

In the next section, we provide a theoretical framework to explain why dual distribution is such a widespread and persistent organizational form in the franchised chain. Our argument is that dual distribution is more efficient than a pure form (either fully franchised or fully company-owned chains) in a context where monitoring is costly and the franchisor cannot observe the behavior and abilities of the managers.

3 THE MODEL

The framework

Consider a chain that wishes to open several units or stores in a geographical market. Beside the location choice of each unit, the chain has to decide the statute of its units (franchised or company-owned) and must recruit adequate outlet managers. Depending on the outlet statute, the managers will be either salaried or franchisees. We assume that the applicants for running an outlet are risk-neutral and a priori may accept a statute of franchisee

or salaried manager.

The key assumption in the model is that the applicants are heterogeneous in their abilities to run a store. But their abilities are imperfectly observed by the chain's owners.

The managers ability is represented by the parameter θ that positively influences the expected store revenues. That means that a highly-talented manager (with a high θ) can achieve more sales than a less talented one (with a low θ). The abilities of the applicants are distributed between \underline{q} and \bar{q} , according to a density function f where F is the cumulative function.

All the units belonging to a same chain are supposed to be identical and we define $V_F(q)$ as the expected revenue of a store run by a manager of type θ . Let us assume that¹ $V'_F(q) > 0$ and $V''_F(q) < 0$. However, for a store operated by a manager θ , the actual revenue can be less than $V_F(q)$, because sales depend not only on ability but also on the efforts provided by the manager. These efforts can take a value between 0 and 1. Let de be the cost of efforts. Thus, these efforts have a positive impact on the revenue: the actual revenue of a store run by a manager of type θ , exerting an effort e is defined by $eV_F(q)$. A manager with ability θ who exerts no effort will achieve no sales whereas the same manager exerting the highest level of effort ($e = 1$) will generate the maximum level of revenues $V_F(q)$. Moreover, the current operating cost of a store, net of the manager's earnings, is supposed to be constant (independent of the sales) and equal to C , whatever the statute of the stores.

Now we can write the payoffs of the managers and the chain by distinguishing the franchised and company-owned units.

¹ We could also conclude that $V_F(\cdot)$ should be positively influenced by the value and the reputation of the brand name.

The franchisee behavior

For the franchised units, we assume that all franchise contracts are identical and involve the payment of revenue-based royalties². The royalty rate is set by the franchisor and is equal to b . Thus a type- θ franchisee's profit is given by $(1-b)eV_F(q) - C - de$, whereas the franchisor receives a revenue of $b e V_F(q)$. Let us assume that for a franchisee the marginal disutility of efforts is always inferior to his marginal benefits:

$$\text{Assumption 1: } (1-b)V_F(q) > d \quad (1)$$

Under this assumption it is always in the franchisee's interest to exert the maximum level of effort ($e = 1$). Assumption 1 rules out any under-provision of effort by the franchisee. For franchisees, moral hazard consists only in *under-declaring* their sales in order to reduce the royalties due (a behavior analog to tax evasion). This form of opportunism could be hard to detect for the franchisor since low declared sales can come from a franchisee with a low ability who behaves honestly or from a franchisee with high ability who behaves opportunistically.

The salaried manager behavior

For the company-owned units, the owner needs to hire salaried managers. Let w be the wage of an employee manager; this wage is supposed to be fixed (no bonus), whatever the effort. If the manager's abilities are θ and his effort is e , then the chain's expected profit is $eV_F(q) - C - w$. For the manager, his utility is $w - de$. Contrary to the franchisee, a salaried manager has strong incentives to provide a minimum level of effort ($e = 0$). Thus, for salaried managers, moral hazard consists in *shirking*, (*i.e.* in choosing a low level of effort). When the owner observes low sales in one of his units, he will not know if it is due to a low effort (opportunism) or to a low ability of his employee.

² We suppose that the franchisor does not claim initial fees from his franchisees.

Auditing, monitoring and moral hazard

The chain faces two kinds of moral hazard from their managers, depending on their statutes: under-declaring from the franchisees and shirking from the salaried managers. To overcome these forms of moral hazard, the franchisor can monitor his managers, by auditing their accounts and/or by visiting them. We consider that the norm of effort for the salaried manager is $e = 1$: if the owner detects an effort below 1, he has the possibility of punishing the manager. These punishments can take different forms from fines to lay-offs³. Similarly if the franchisor detects under-declaration of revenues, he can inflict financial penalties or terminate the franchise contract⁴.

We assume that the control is "perfect", *i.e.* an audit enables the franchisor to discover the true abilities of the manager with certainty. But this control is costly to implement. We assume that the cost of monitoring an outlet is fixed and denoted by K if the store is company-owned and Q if the store is franchised. Consequently, the monitoring policy consists in a probability of control based on the available information. Here the information is the declared sales. Through this declaration, it is as if the manager announces his abilities or talent. Thus, when a type q franchisee declares a revenue R to the franchisor, it is equivalent to announce a type q_R with $q_R = V_F^{-1}(R)$ where $V_F^{-1}(\cdot)$ is the inverse revenue function. If the franchisee behaves honestly (declaration of his actual sales), then $q_R = q$, and if he behaves

³ A termination of the labor contract represents a severe punishment if the manager is paid above the market wage, at an efficiency wage level (Shapiro and Stiglitz (1994), or if he has invested in specific human capital, whether he has partly or entirely financed it. See Krueger (1991) for empirical evidences suggesting that efficiency wage is at work for employees in company-owned units in the fast food industry.

⁴ The termination or the non renewal of the franchise contract appears to be a severe punishment if the franchisee holds specific assets or if he earns rents (Kaufman and Lafontaine, 1994). See Klein (1995) and Lafontaine and Raynaud (2001) for insights on self-enforcement in franchising.

opportunistically (under-declaration of his sales, $R < V_F(q)$), then $q_R < q$. Moreover, the franchisor's revenues are given by $bV_F(q_R)$.

Similarly when a type q salaried manager generates a revenue R , it is equivalent to announce a type q_E , with $q_E = V_F^{-1}(R)$. Thus, we have $q_E = q$ if the salaried manager behaves honestly ($e=1$) and $q_E < q$ if he behaves opportunistically (shirking $e<1$). Moreover the franchisor's revenues are equal to $V_F(q_E) - w - C$.

The monitoring policy implemented by the franchisor should be conditional on the declared sales q_R and observed sales q_E . Let $g(q_R)$ represent the probability of controlling a franchisee who declares an amount of sales $V_F(q_R)$ and let $v(q_E)$ represent the probability of controlling a salaried manager who achieves an amount of sales $V_F(q_E)$.

Moreover, the monitoring policy must plan for punishment schemes if the manager is detected as behaving opportunistically. Let a represent the penalty inflicted on an opportunistic franchisee and b represent the penalty for an opportunistic salaried manager. In both cases, penalties are supposed to be independent of the damage. Moreover, we suppose that

$$\text{Assumption 2: } a \geq b[V_F(\bar{q}) - V_F(q)]$$

$$\text{Assumption 3: } b \geq \frac{d(V_F(\bar{q}) - V_F(q))}{V_F(q)}$$

A2 and A3 mean that the penalties are supposed to be sufficiently high to deter any opportunistic behavior. Given the severity of the punishments, a manager who is certain to be controlled ($g(q_R) = 1$ or $v(q_E) = 1$) is deterred to cheat.

What is the optimal monitoring policy for minimizing the moral hazards of the

managers? Does this policy depend on the organizational design of the chain? Does dual distribution enable a better monitoring of store managers than a pure franchised system or a pure company system? To answer these questions, we consider the following sequential game between the chain's owner and the managers.

- In a first stage, the franchisor sets the proportion of company-owned units and of franchised units within his chain. The owner can opt either for a pure system or a plural form (dual distribution).
- In a second stage, the franchisor announces his monitoring policy $(g(q_R), v(q_E))$ (e.g. the probability of control in each store). This monitoring policy is likely to be different for the franchised units and for the company units.
- In the third stage, the manager decides to be opportunistic or honest. Thus, given the monitoring policy, the franchisee declares a level of sales $V_F(q_R)$ and the salaried manager chooses a level of effort that leads to a level of sales $V_F(q_E)$.
- In the last stage, the chain implements its monitoring program, depending on the observed or declared abilities of the managers (q_E and q_R).

In the next section, we solve this strategic game and determine the optimal monitoring policy for all the possible organizational forms of our store network.

4 OPTIMAL MONITORING POLICY UNDER ASYMETRIC INFORMATION

4.1 Pure systems and complete monitoring

Monitoring Policy applied to the franchisees

Consider first the case of a pure franchised chain. If the franchisor wants to deter a

type q to declare q_R , then he must announce a probability of control (*i.e.* an intensity of control), equal to $g(q_R)$ such that it is in the interest of the type q manager to reveal his true type:

$$(1-b)V_F(q) - C - d \geq V_F(q) - bV_F(q_R) - C - d - g(q_R)a \quad (1)$$

The left-hand term corresponds to the expected payoff of a franchisee declaring his actual sales and the right-hand term his payoff when he behaves opportunistically (declaring $V(q_R)$ instead of $V(q)$). With this latter strategy, he can capture a higher part of the total revenues but at the risk of being detected and punished by the franchisor ($g(q_R)a$). After rearrangement, the intensity of monitoring $g(q_R)$ should be at least equal to:

$$g(q_R) \geq \frac{b(V_F(q) - V_F(q_R))}{a} \quad (2)$$

If the franchisor wants to deter all franchisees to behave opportunistically, then he can achieve this result by announcing the monitoring policy defined for any q_R by :

$$g(q_R) = \frac{b(V_F(\bar{q}) - V_F(q_R))}{a} \quad (3)$$

Since the highest talented franchisees (type \bar{q}) are the most difficult to monitor, it is sufficient to deter them to cheat to prevent all the other types of franchisees q from cheating (with $q < \bar{q}$). This can be easily proven by writing the expected cheating payoff for a type q :

$$V_F(q) - bV_F(q_R) - C - d - \left(\frac{b(V_F(\bar{q}) - V_F(q_R))}{a} \right) a \quad (4)$$

After simplification, we find $V_F(q) - bV_F(\bar{q}) - C - d$, an earning well below what an honest franchisee could expect (*i.e.* $V_F(q) - bV_F(q) - C - d$ since V_F is an increasing function of q).

Monitoring Policy applied to the salaried managers

For salaried managers, the monitoring policy is different. If the owner wants to deter a manager q to cheat, he has to control him with a probability of $v(q_E)$, defined by:

$$w - d \geq w - de - v(q_E)b \quad (5)$$

$$\text{with } e = \frac{V_F(q)}{V_F(q_E)}$$

After rearrangement, the intensity of monitoring $v(q_E)$ should be at least equal to:

$$\frac{d(V_F(q) - V_F(q_E))}{bV_F(q)} \quad (6)$$

If the owner wants to deter all salaried managers to behave opportunistically, then he must implement a monitoring policy $v(q_E)$ characterized for any q_E by:

$$v(q_E) = \frac{d(V_F(\bar{q}) - V_F(q_E))}{bV_F(\bar{q})} \quad (7)$$

With such a monitoring intensity, the highest talented managers (type \bar{q}) are deterred to shirk and we can also see that the less talented types $q < \bar{q}$ have no incentives to cheat.

From (7) the expected payoff of an opportunistic salaried manager is equal to:

$$w - d \frac{V_F(q_E)}{V_F(q)} - \left(\frac{d(V_F(\bar{q}) - V_F(q_E))}{bV_F(\bar{q})} \right) b \quad (8)$$

This expression is equivalent to $w - d - d \left(\frac{V_F(q_E)}{V_F(q)} - \frac{V_F(q_E)}{V_F(\bar{q})} \right)$, a payment inferior to

$w - d$, (the earnings of the honest manager), since $V_F(q) < V_F(\bar{q})$.

Now we will analyze to what extent controlling all the stores in a chain is an optimal policy. We start with the case of a vertically integrated chain.

4.2 Pure company-owned chain and partial monitoring

Following Besanko and Spulber (1989), we consider a two-part monitoring program defined by:

$$\text{if } \begin{cases} q_E > \hat{q}, & \text{then} \\ q_E \leq \hat{q}, & \text{then} \end{cases} \quad \begin{cases} v(q_E) = 0 \\ v(q_E) = d \frac{(V_F(\hat{q}) - V_F(q_E))}{bV_F(\hat{q})} \end{cases} \quad (9)$$

A complete control would consist in setting $\hat{q} = \bar{q}$ and a partial control is characterized by $\hat{q} < \bar{q}$. In the latter case, the franchisor does not control the salaried managers who achieve sales superior to $V_F(\hat{q})$. But managers who generate low revenues are controlled with a probability that decreases with the level of sales. Thus the owner is concentrating his efforts and financial means on the less profitable units (those generating sales below $V_F(\hat{q})$)⁵.

Given such a monitoring policy, managers with talent $q \leq \hat{q}$ are deterred to shirk and will always choose $q_E = q$ (i.e. an effort $e=1$). Managers with talent $q > \hat{q}$ are aware that they are never controlled as long as they achieve an amount of sales superior or equal to $V_F(\hat{q})$.

Therefore it is in the interest of these managers to choose the exact level of effort $e = \frac{V_F(\hat{q})}{V_F(q)}$

that enables them to reach the amount of revenues $V_F(\hat{q})$ ⁶.

Finally, with partial monitoring, the owner offers a rent to the more talented managers, in turn allowing them to exert a lower effort than the norm.

⁵ Let recall that all the units are homogeneous and differ only in the abilities and efforts of their managers. Two units run by identical talent managers should generate identical revenues.

⁶ We can easily prove that a salaried manager can never expect a higher payoff by exerting an effort below

$$e = \frac{V_F(\hat{q})}{V_F(q)}. \text{ Thus we have } w - d \frac{V_F(q_E)}{V_F(q)} - \left(\frac{d(V_F(\hat{q}) - V_F(q_E))}{bV_F(\hat{q})} \right) b < w - d \frac{V_F(\hat{q})}{V_F(q)}.$$

Proposition 1. *In a pure company ownership system, it is never optimal to control all the units, when the managerial abilities and efforts are imperfectly observed. A partial monitoring is more optimal for the chain.*

Proof:

The optimal monitoring policy is determined by maximizing the chain's profit with respect to the managerial threshold \hat{q} :

$$G = \int_{\underline{q}}^{\hat{q}} (V_F(q) - w - C - v(q)K) f(q) dq + \int_{\hat{q}}^{\bar{q}} (V_F(\hat{q}) - w - C) f(q) dq$$

The derivative with respect to \hat{q} is given by

$$\frac{\partial G}{\partial \hat{q}} = [V_F(\hat{q}) - w - C - v(\hat{q})K - V_F(\hat{q}) + w + C] f(\hat{q}) + V_F'(\hat{q}) \int_{\hat{q}}^{\bar{q}} f(q) dq + \frac{\partial v}{\partial \hat{q}} K \int_{\underline{q}}^{\hat{q}} f(q) dq$$

After rearrangement:

$$\frac{\partial G}{\partial \hat{q}} = - \int_{\underline{q}}^{\hat{q}} \frac{dV_F'(\hat{q}) V_F(q) K}{b(V_F(\hat{q}))^2} f(q) dq + V_F'(\hat{q}) \int_{\hat{q}}^{\bar{q}} f(q) dq$$

Complete monitoring consists in setting a threshold $\hat{q} = \bar{q}$ (all the company-owned units are controlled whatever the declared level of sales). But such a policy is not optimal as the marginal profit decreases with \hat{q} at $\hat{q} = \bar{q}$:

$$\left. \frac{\partial G}{\partial \hat{q}} \right|_{\hat{q}=\bar{q}} = - \int_{\underline{q}}^{\bar{q}} \frac{dV_F'(\bar{q}) V_F(q) K}{b(V_F(\bar{q}))^2} f(q) dq < 0$$

The chain could obtain higher profits by relaxing its control on the most efficient units.

Besides, we can prove that a laissez-faire policy ($\hat{q} = \underline{q}$) is not more optimal:

$$\left. \frac{\partial G}{\partial \hat{q}} \right|_{\hat{q}=\underline{q}} = V_F'(\underline{q}) > 0$$

This result can be explained by the fact that partial deterrence allows for a reduction in

monitoring costs. Even if the revenues are less than expected in the stores run by the more talented managers, this is more than counterbalanced by the savings related to monitoring costs. Notice that the “informational rent” captured by the high-ability managers is an increasing function of monitoring costs and a decreasing function of penalty severity. Thus the proportion of controlled units should be close to 100% when monitoring costs tend towards zero or the punishments become extremely severe. However even in these extreme conditions, it is always in the interest of the owner to give up auditing the most profitable units (controlling these units is useless).

4.3 Pure franchised chain and partial monitoring

Now let us consider the monitoring program for a pure franchised system as follows:

$$\left| \begin{array}{ll} q_R > \hat{q}, & \text{then } g(q_R) = 0 \\ \text{if} & \\ q_R \leq \hat{q}, & \text{then } g(q_R) = b \frac{\left(V_F(\hat{q}) - V_F(q_R) \right)}{a} \end{array} \right. \quad (10)$$

Complete monitoring consists in setting $\hat{q} = \bar{q}$. With $\hat{q} < \bar{q}$, the franchisor makes the choice of partial monitoring and offers a rent to some of his more talented franchisees. More precisely, all the franchisees with a type $q > \hat{q}$ are allowed to under-declare their revenues ($V(\hat{q})$ instead of $V(q)$) and thus pay lower royalties than normally required ($bV_F(\hat{q})$ instead of $bV_F(q)$). As long as they pay royalties superior than or equal to $bV_F(\hat{q})$, they are not controlled. So they have no reason to declare revenues below $V(\hat{q})$ ⁷. Conversely, all

⁷ For franchisees with type $q > \hat{q}$, the benefit from declaring revenues below $V(\hat{q})$ are given by:

franchisees with a type $q \leq \hat{q}$ are audited and have strong incentives to be honest (to declare their actual revenues).

Proposition 2. *In a pure franchised system, it is never optimal to audit all the stores, when franchisee abilities and store revenues are imperfectly observed. A partial monitoring is more profitable.*

Proof: The optimal intensity of monitoring \hat{q} is given by maximizing the franchisor's profit with respect to the managerial threshold monitoring \hat{q} :

$$\text{Max}_{\{\hat{q}\}} G = \int_q^{\hat{q}} (bV_F(q) - g(q)Q) f(q) dq + \int_{\hat{q}}^{\bar{q}} bV_F(\hat{q}) f(q) dq$$

The derivative of the profit with respect to \hat{q} is as follows:

$$\frac{\partial G}{\partial \hat{q}} = -g(\hat{q})Qf(\hat{q}) - b \frac{V'_F(\hat{q})}{a} Q \int_q^{\hat{q}} f(q) dq + b V'_F(\hat{q}) \int_{\hat{q}}^{\bar{q}} f(q) dq$$

Complete monitoring consists in setting $\hat{q} = \bar{q}$ (all the franchisees may be audited whatever the declared sales). But such a policy is not optimal, since the franchisor can increase his profit by relaxing his monitoring efforts on the most efficient units. Indeed, marginal profit decreases with \hat{q} at $\hat{q} = \bar{q}$:

$$\left. \frac{\partial G}{\partial \hat{q}} \right|_{\hat{q} = \bar{q}} = -b \frac{V'_F(\bar{q})}{a} Q < 0$$

$$V_F(q) - bV_F(q_R) - C - d - \left(\frac{b(V_F(\hat{q}) - V_F(q_R))}{a} \right) a = V_F(q) - bV_F(\hat{q}) - C - d$$

So they are perfectly indifferent between declaring $V(\hat{q})$ and declaring lower revenues.

We can also see that a "laissez-faire" policy $\hat{q} = q$ is sub-optimal

$$\frac{\partial G}{\partial \hat{q}} \Big|_{\hat{q} = q} = g(q)Qf(q) + bV'_F(q) > 0$$

The rationale for this result is the same as for proposition 1. The existence of monitoring costs incites the franchisor to renounce auditing the most profitable franchised units. Thus, a franchisee with a type $q > \hat{q}$ benefits from an "informational rent" equal to $b(V_F(q) - V_F(\hat{q}))$. The higher the franchisee's abilities, the greater the rent he may capture. By offering a rent to the most talented franchisees, the franchisor can deter them to undertake more harmful cheating at a lesser cost. We notice that partial monitoring leads the franchisor to audit only honest franchisees, but it is absolutely necessary to comply with this if the franchisor wants to prevent the high-talented franchisees from passing themselves off as low-talented franchisees (by declaring low revenues).

Up to now we have demonstrated that complete monitoring is never optimal in a pure system. Let us turn to the issue of dual distribution. Is a mix chain more or less efficient than a pure system in containing moral hazards from managers?

4.4 Dual distribution and efficiency

A measure of the company ownership

First we have to specify how the store managers are recruited in a mix chain. Are managerial abilities identically distributed among franchised units and company-owned units? The answer is no, because as mentioned previously, the manager candidates are supposed to apply for the most profitable statute whenever a chain offers the two statutes. Thus, the more

talented managers should choose the statute of franchisee and the less talented the statute of salaried manager, since the earnings of a franchisee increase with his talent, whereas the gains of an honest manager is independent of his talent.

Let us define the manager with talent \tilde{q} as the type indifferent between being franchisee or salaried. If such a type exists, then all the managers with talent $q > \tilde{q}$ should prefer to have a franchisee statute and all the managers with talent $q \leq \tilde{q}$ should be salaried.

The type \tilde{q} is defined by:

$$w - d \frac{V_F(\hat{q})}{V_F(\tilde{q})} = (1 - b)V_F(\tilde{q}) - C - d \quad (11)$$

The left hand side corresponds to the gain of a salaried manager of talent \tilde{q} , who provides an effort $e = \frac{V_F(\hat{q})}{V_F(\tilde{q})} < 1$, because he is one of the most talented employee managers and knows that he will never be controlled. The right hand side corresponds to the profit of a type \tilde{q} franchisee who behaves honestly because he is certain to be controlled, by declaring the worst revenues of all the franchised units.

\tilde{q} is the threshold type that influences the organizational form of the system. If \tilde{q} increases, the franchise applicants should be less numerous and the proportion of company units should rise. Conversely, if \tilde{q} decreases, the applicants for a franchise should be more numerous and the proportion of company units should decrease. Thus choosing the proportion of company-owned units is to some extent equivalent to choosing the \tilde{q} .

Complementarity between royalty rates and company ownership

Moreover, modifying the intensity of control in the company-owned stores \hat{q} or the level of royalty rates b , has an impact on the extent of company ownership. Indeed, equation

(11) shows that \tilde{q} , \hat{q} and b are linked. Therefore, the franchisor has three interdependent strategic variables through which he can improve the chain's efficiency: the monitoring intensity \hat{q} , the level of royalty rates b and the proportion of company-owned stores \tilde{q} . If we fix one of the three variables, then the two others appear to be positively linked. This leads to an interesting implication:

Implication. *Assuming that the monitoring intensity of salaried managers is invariable, then royalty rates (b) and the percent of company-owned units (\tilde{q}) tend to be positively linked and can be used as complementary variables by the franchisor.*

This result gives some theoretical foundations to the strong empirical findings of a positive relationship between royalty rates and the proportion of company-owned units (Lafontaine 1992, Lafontaine and Shaw 1999, Pénard, Raynaud, Saussier, 2003).

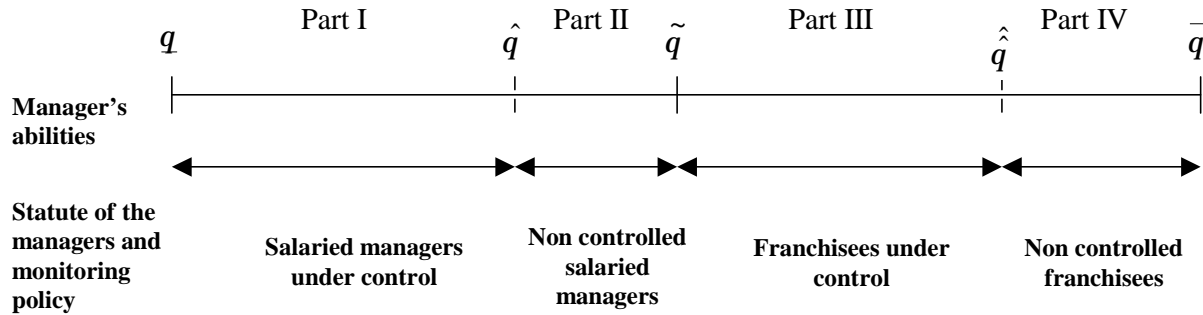
Now, we have to prove that dual distribution is the most efficient organization. For this purpose, we have to define the monitoring policy and the profits of the franchisor, given (\tilde{q}, \hat{q}, b) .

Monitoring policy under dual distribution

If \tilde{q} is the type indifferent between being salaried and franchisee, then the monitoring policy $(v(q), g(q))$ announced by the franchisor should be as follows:

$$\begin{aligned}
 & q \in [q, \hat{q}] \quad \text{then} \quad v(q) = d \frac{(V_F(\hat{q}) - V_F(q))}{bV_F(\hat{q})} \\
 \text{if } & q \in [\hat{q}, \tilde{q}] \quad \text{then} \quad v(q) = 0 \\
 & q \in [\tilde{q}, \hat{q}] \quad \text{then} \quad g(q) = 0 \\
 & q \in [\hat{q}, \bar{q}] \quad \text{then} \quad g(q) = b \frac{(V_F(\hat{q}) - V_F(q))}{a}
 \end{aligned} \tag{12}$$

The franchisor implements a four-part monitoring program. Part I refers to the monitoring of the less talented salaried managers and part II to the higher talented salaried managers. Part III is dedicated to the less talented franchisees and part IV to the high ability franchisees.



As the distribution of a manager's talents or abilities within the chain replicates the applicants' abilities (the cumulative function of the manager's abilities being F), then the profits of the mix chain are given by:

$$\begin{aligned}
G = & \int_{\underline{q}}^{\hat{q}} (V_F(q) - w - C - v(q)K) f(q) dq + \int_{\hat{q}}^{\tilde{q}} (V_F(\hat{q}) - w - C) f(q) dq \\
& + \int_{\tilde{q}}^{\hat{q}} (bV_F(q) - g(q)Q) f(q) dq + \int_{\hat{q}}^{\bar{q}} bV_F(\hat{q}) f(q) dq
\end{aligned} \tag{13}$$

$$\text{with } w - d \frac{V_F(\hat{q})}{V_F(\tilde{q})} = (1 - b)V_F(\tilde{q}) - C - d$$

The proportion of company ownership that maximizes the chain revenue is given by:

$$\frac{\partial G}{\partial \tilde{q}} = (V_F(\hat{q}) - w - C) f(\tilde{q}) - bV_F(\tilde{q}) f(\tilde{q}) + g(\tilde{q}) Q f(\tilde{q}) = 0 \tag{14}$$

This leads to the following propositions.

Proposition 3. *Under asymmetric information, if a chain is completely company-owned, then it is optimal to franchise some units*

Proof:

As $g(\bar{q}) = 0$ and according to equation (14),

$$\left. \frac{\partial G}{\partial \tilde{q}} \right|_{\tilde{q} = \bar{q}} = (V_F(\hat{q}) - C - w) f(\bar{q}) - bV_F(\bar{q}) f(\bar{q})$$

As $w - d \frac{V_F(\hat{q})}{V_F(\bar{q})} = (1 - b)V_F(\bar{q}) - C - d$, then

$$\left. \frac{\partial G}{\partial \tilde{q}} \right|_{\tilde{q} = \bar{q}} = - (V_F(\bar{q}) - V_F(\hat{q})) \left(1 - \frac{d}{V_F(\bar{q})} \right)$$

From assumption 1, we know that $d < V_F(\bar{q})$.

$$\text{So, } \left. \frac{\partial G}{\partial \tilde{q}} \right|_{\tilde{q} = \bar{q}} < 0$$

A pure company-owned chain is never optimal.

The role of franchisees is to limit opportunistic behaviors of the salaried managers and

to reduce the rent that the chain is forced to offer them.

Proposition 4. *Under asymmetric information, if a chain is completely franchised, then it is optimal to convert some franchised units into company-owned units*

Proof:

According to equation (14),

$$\frac{\partial G}{\partial \tilde{q}} \Big|_{\tilde{q} = q} = (V_F(q) - C - w)f(q) - bV_F(q)f(q) + g(q)Qf(q)$$

As $w - d \frac{V_F(q)}{V_F(q)} = (1 - b)V_F(q) - C - d$, then we obtain

$$\frac{\partial G}{\partial \tilde{q}} \Big|_{\tilde{q} = q} = g(q)Qf(q) > 0$$

The role of company units is to limit the opportunism of franchisees and to incite them to declare their exact amount of sales. Thus by converting some franchises into company units, the franchisor can achieve some economies on the monitoring costs.

The two previous propositions (3 and 4) allow us to assert that franchise mix is an optimal organizational design when the franchisor cannot observe the abilities and behaviors of managers. This dual system is more efficient than a pure franchised system or a pure company-owned system.

Therefore, the franchisor has an ideal level of franchising mix (\tilde{q}) that enables him to maximize his profits. This company ownership target always belongs to the interval $]\underline{q}, \bar{q}[$ and should strongly depend on two key variables: royalty rates and the monitoring technology. As mentioned before, the ideal proportion of company owned units should increase with

royalty rates (complementarity). Besides, the extent of company ownership should increase with the monitoring costs of franchisees and decrease with the severity of punishments that the chain can inflict towards opportunistic franchisees. Indeed a chain will call for franchising more intensively if it has better means to monitor its franchised units. Conversely, the proportion of company-owned units should decrease with the monitoring costs of salaried managers and increase with the severity of punishments that the chain can inflict on opportunistic franchisees.

5. CONCLUSION

In this paper, we have given some insights on the predominance and the efficiency of dual distribution. The rationale for dual distribution relies mainly on a minimization of monitoring costs. This model is rather original when compared with previous literature. First, we assume that the franchisor cannot freely observe the actual level of a franchisee's sales and that franchisees have an incentive to under-declare the actual revenues in order to minimize the royalties due. By auditing franchise units, the franchisor can improve his information and deter the manager to provide a low level of effort although this process is costly. This type of opportunistic behavior is close to the one studied in Mathewson and Winter (1985) where a franchisee can under-declare his sales. However, in their model, the possibility of franchisee cheating arises because the franchisor cannot identify, without cost, the realized state of demand. Here franchisees are tempted to cheat because the franchisor must bear a cost to control the actual sales in the franchised outlets. In Bradach's (1997) study of chains in the restaurant industry, he states that, whereas headquarter management information systems exist for company units, this is not the case for the franchisee: in some chains *"the only financial information that the chain received was a revenue number each month from which the royalty*

was calculated' (1997, p. 288). Our model is also related to the work of Gal-Or (1995) that focuses on the incentive of the franchisor to monitor outlets. She found that it can be optimal for the chain to monitor only a subset of outlets. However, this result is founded on the outlets heterogeneity whereas partial monitoring in our case is obtained with homogeneous outlets. Our results would be strengthened if we relax this assumption and consider heterogeneous outlets.

REFERENCES

Azevedo P.F., and V.L. Dos Santos Silva., 2001, "Contractual Mix Analysis in The Brazilian Franchising", Unpublished Manuscript.

Bai, C.E, and Z. Tao, 2000, "Contract Mix in Franchising", *Journal of Economics and Management Strategy*, 9 (1), 85-113.

Baron, D.P., and D. Besanko, 1984, "Regulation, Asymmetric Information and Auditing", *RAND Journal of Economics*, Vol. 15, 447-470.

Bercovitz, J.E., 2001, "The Organizational Choice Decision in Business-Format Franchising: An Empirical Test", working paper, Fuqua School of Business, Duke University.

Besanko, D., and D. Spulber, 1989, "Antitrust Enforcement under Asymmetric Information", *Economic Journal*, 99, 408-425.

Bradach, J.L., 1997, "Using the Plural Form in the Management of Restaurant Chains", *Administrative Science Quarterly*, 42, 276-303.

Bradach, J.L., and R.G. Eccles, 1989, "Price, Authority and Trust: From Ideal Form to Plural Forms", *Annual Review of Sociology*, 15, 97-118.

Brickley, J., and F.H. Dark, 1987, "The Choice of Organizational Form: The Case of Franchising", *Journal of Financial Economics*, 18, 401-420.

Brickley, J.A., 1999, "Incentive Conflicts and Contractual Restraints: Evidence from Franchising", *Journal of Law & Economics*, XLII, 745-774.

Brickley, J.A., 2002, "Royalty Rates and Upfront Fees in Share Contracts: Evidence from Franchising", *Journal of Law, Economics & Organization*, 18(2), 511-535.

Caves, R.E., and W.F. Murphy, 1976, "Franchising: Firms, Markets and Intangible Assets", *Southern Economic Journal*, 42, 572-586.

Gal-Or, E., 1995, "Maintaining Quality Standard in Franchised Chains", *Management Science*, 41(11), 1774-1792.

Gallini, N.T., and N. Lutz, 1992, "Dual Distribution and Royalty Fees in Franchising", *Journal of Law, Economics & Organization*, VIII, 471-501.

Holmström, B., and P. Milgrom, 1991, "Multitask Principal Agent Analysis: Incentive Contracts, Asset Ownership and Job Design", *Journal of Law, Economics & Organization*, 7, 24-52.

Khalil, F., 1997, "Auditing without Commitment", *RAND Journal of Economics*, Vol.28, N°4, 629-640.

Kaufman, P., F. Lafontaine, 1994, "Costs of Control: The Source of Economic Rents for McDonald's Franchisees", *Journal of Law & Economics*, 37(2), 417-453.

Klein, B., 1995, "The Economics of Franchise Contracts", *Journal of Corporate Finance*, 2, 9-37.

Krueger, A.B., 1991, "Ownership, Agency and Wages: An Examination of Franchising in the Fast Food Industry", *Quarterly Journal of Economics*, 106(1), 75-101.

Lafontaine, F., 1992, "Agency Theory and Franchising: Some Empirical Results", *RAND Journal of Economics*, 23, 263-283.

Lafontaine, F., 1993, "Contractual Arrangement as Signaling Devices: Evidence from Franchising", *Journal of Law, Economics & Organization*, 9, 256-289.

Lafontaine, F., K.L. Shaw, 1999, "The Dynamics of Franchise: Evidence From Panel Data", *Journal of Political Economy*, 107, 1041-1080.

Lafontaine, F., and K.L. Shaw, 2001, "Targeting Managerial Control: Evidence from Franchising", working paper, University of Michigan Business School.

<http://laf.bus.umich.edu/pub/wp.lafontaine.2001.07.14.pdf>

Lafontaine, F., and E. Raynaud, 2002, "Residual Claims and Ongoing Rent as Incentive Mechanisms in Franchise Contracts: Complements or Substitutes?", in E. Brousseau and J-M. Glachant, eds., *The Economics of Contracts: Theories and Applications*, Cambridge University Press

<http://www.nber.org/papers/w8868> .

Lafontaine, F., and M. Slade, 2001, "Incentive Contracting and the Franchise Decision", forthcoming in K. Chatterjee and W. Samuelson (eds.), *Advances in Business Application of Game Theory*, Kluwer Academic Press.

Lewin-Solomon, S.B., 1999, "Innovation and Authority in Franchise Systems: An Exploration of the Plural Form", working paper, Dept. of Applied Economics, Cambridge University.

<http://www.econ.cam.ac.uk/dae/people/lewin/empirical85.pdf>

Mathewson Winter, 1985, "The Economics of Franchise Contracts", *Journal of Law & Economics*, 28, 503-526.

Minkler, A., 1990, "An Empirical Analysis of a Firm Decision to Franchise", *Economic Letters*, 34, 77-82.

Mookherjee, D., and I.P.L. Png, 1989, "Optimal Auditing, Insurance, and Redistribution", *Quarterly Journal of Economic*, Vol.104, 339-415.

Norton, S., 1988, "An Empirical Look at Franchising as an Organizational Form", *Journal of Business*, 61, 197-217.

Oxenfeldt, A.R., and A.O. Kelly, 1969, "Will Successful Franchise Systems Ultimately

Become Wholly-Owned Chains?" *Journal of Retailing*, 44, 69-83.

Pénard T., E. Raynaud and S. Saussier, 2003, "Dual Distribution and royalty rates in Franchised Chains: An Empirical Analysis Using French Data", *Journal of Marketing Channels*, 10, 5-32.

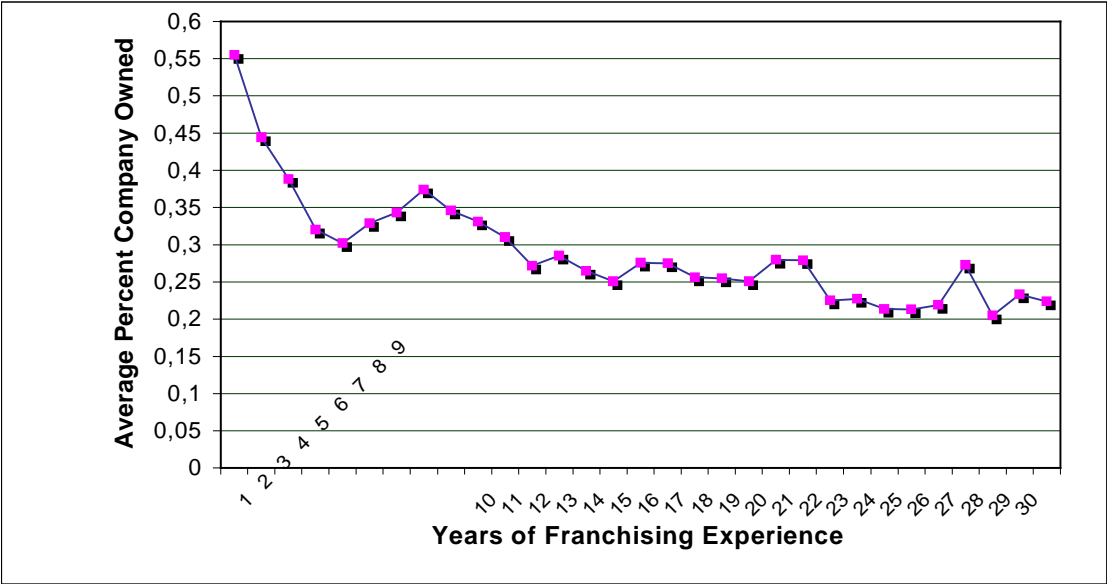
Rubin, P., 1978, "The Theory of the Firm and the Structure of Franchise Contract", *Journal of Law & Economics*, 21, 223-233.

Scott, F.A., 1995, "Franchising vs. Company Ownership as a Decision Variable of the Firm", *Review of Industrial Organization*, 10, 69-81.

Shapiro, C., J.E. Stiglitz, 1984, "Equilibrium Unemployment as a Worker Discipline Device", *American Economic Review*, 74, 433-444.

Sorenson O., and J.B. Sorensen, 2001, "Finding the Right Mix: Organizational Learning, Plural Forms, and Franchise Performance", *Strategic Management Journal*, 22 (6-7), 713-724.

FIGURE 1. The proportion of Company Units as a Function of Franchising Experience



Source: Pénard-Raynaud-Saussier 2003