THE KOREAN ECONOMIC REVIEW
Volume 20, Number 2, Winter 2004

THE COST OF DIVERSIFICATION: AN ORGANIZATIONAL PERSPECTIVE

NAMHOON KWON*

This paper approaches the issue of corporate diversification from the perspective of the theory of firm. The cost of diversification is identified by the loss of innovative information due to increased communication requirement within the firm. The level of diversification is shown to decrease with a better information processing technology and stabler business environment. These results are in accordance with historical and industrial patterns of diversification.

JEL Classification: L22, M20
Keywords: diversification, firm organization, lattice-theoretic method

I. INTRODUCTION

For most firms, one of the most important strategic decisions to make is what portfolio of businesses to maintain within their boundaries. These portfolios might include only the operations that are vertically related or the subsidiaries serving geographically separate markets. More frequently, they consist of diverse product lines with various degrees of relatedness. Intricacy of making such decisions is vividly evidenced by the fact that they often suffer bad consequences.[1] Despite the importance most firms put on their business portfolio strategies, however, economists still have only shallow understandings on the issue.

Diversification is defined as a business portfolio decision to expand a firm's

Received for publication: May 20, 2004. *Revision accepted:* July 20, 2004.

* Department of Economics, Konkuk University, 1 Hwayang-dog, Gwangjin-gu, Seoul 143-701, Tel.02-450-3622. E-mail address: namhoon@konkuk.ac.kr. This work was supported by the faculty research fund of Konkuk University in 2003. I thank two anonymous referees. All errors are mine.

[1] For example, Michael Porter (1987) studies the diversification records of 33 large U.S. companies over 1950-1986 period and finds that most of them divested many more acquisitions than they kept.

operational boundary into a new market, although the popular notion of the term often implies diversification into a new product line. According to the survey by Montgomery (1994), there are largely three perspectives for the causes of corporate diversification—market power, agency, and resource-based.[2] The market-power view argues that firms diversify to increase their commanding power over the markets and reap the benefit of reduced competition, while the agency view typically focuses on the separation of ownership and management and considers diversification as the result of the manager's pursuit of self-interest at the expense of the shareholders.[3] Finally, the resource view argues that firms diversify in response to excess capacity in productive factors, or *resources*.[4] This view is closely related to economies of scale and scope because utilizing resources through diversification would reduce the average production cost. Transaction cost theory of firm is also a key component of this view because it is assumed explicitly or implicitly that internal transfers are superior to market transactions in utilizing resources.

These views have their own drawbacks, however. The market-power view lacks empirical support; there is little evidence that diversified firms in general attain the sort of market power leading to increased profitability. On the other hand, prevalent versions of the agency view consider diversification as neither efficient nor profit maximizing. They assume that shareholders either lacks the power to stop managers or do not know that they are taken advantage of. Neither of these premises would fully explain such a widespread and persisting phenomenon as diversification. In contrast, the resource view is appealing because it considers diversification as an efficiency seeking behavior, but the view has the problem of not specifying the cost of diversification clearly. As Holmstrom and Tirole (1989) point out, while it is relatively easy to envision reasons for integration, it is substantially harder to articulate the cost. Yet, the existence of such cost is evident because a firm's optimal size would be limitless without the cost.[5]

---

[2] There is also an argument that corporate managers diversify on behalf of shareholders who do not want large swings in value. This rationale however would be sensible only if investors are, for some reason, unable to diversify their portfolios on their own. See Shleifer and Vishny (1986) for further discussion.

[3] There is an ample literature taking this view. Amihud and Lev (1981), Lewellen, Loderer, and Rosenfeld (1989) focus on risk-reduction motives by the managers. Mueller (1972) and Jensen (1986) emphasize free cash flow available for managers to build an 'empire.' On the other hand, managerial entrenchment theory by Shleifer and Vishny (1989) suggests that managers try to direct a firm's diversification strategy to increase demand for their skills. Avery, Chevalier, and Schaefer (1998) find that CEOs who undertake acquisitions are more likely to be appointed to *other firms*' boards of direoctors. See also Morck, Shleifer, and Vishny (1990).

[4] Penrose (1959) pioneered this approach. Prahalad and Bettis (1986) suggest that managers of diversified firms may spread their own managerial talent across nominally unrelated business areas. Teece (1982) and Caves (1982) emphasize the difficulties of trading intangible assets through market transactions.

In this paper, we propose an important source of diversification cost, namely, the loss of innovative information generated across the firm. Let us explain in more details. Once a firm diversifies for an efficiency purpose, at least some degree of reorganization should follow to exploit the *synergy*. Structurally, this in general results in extra layers of management to meet increased requirement for coordination, which involves transfers of key decision rights and increases in communication requirements.[6] Then costly and unreliable communication imposes costs for organizational expansion. More important problem, however, is that employees now have less incentives to produce and implement innovative business ideas. This is because information asymmetry exists about how much effort is exerted to produce such an idea, in which case a firm's incentive mechanism tends to depend upon how much the idea actually improves business results. If ideas are easily lost during internal communication and therefore less likely to be implemented, employees' efforts to generate those ideas cannot be rewarded properly.

Although this argument is intuitive, it alone is insufficient to generate interesting results unless some testable predictions could be made. Therefore, we set up a formal ageny-based model of diversification costs that allows comparative statics. In particular, we show that diversification is less costly when the business environment is unstable, or communication loss due to diversification is smaller. Also, if communication technology is effective, a firm with *smaller* hierarchy is more likely to diversify. On the other hand, if communication technology is ineffective, a firm with *larger* hierarchy is more likely to diversify.

From the existing literature, Rotemberg and Saloner (1994) is the most related to our paper. They argue that focusing on streamlined businesses could increase the incentives for innovation. However, their model is based on a special environment in which contract is very incomplete, and a diversified firm cannot commit to provide proper rewards for outcomes. Another problem is that they believe that diversified firms are not innovative but exploit existing technologies. This is not necessarily true, since many large diversified firms are very R&D oriented, while there are only mixed evidences regarding the innovativeness of small firms. For example, Cardinal and Opler (1995) show that large diversified

[5] This point can be best illustrated by the organization of *selective intervention* suggested by Williamson (1975, 1985). He argues

> If, for example, demand or cost inter-action effects are such that net gains can be had by moving decisions to the top, it will be done. Those decisions, however, that are most efficiently made at operating levels will remain there. Intervention at the top thus always occurs selectively, which is to say only upon a showing of expected net gains. The resulting combined firm can therefore do everything that the two autonomous firms could do previously and more [1985, p.133].

[6] Of course, it is not always necessary to add layers of managers to handle the coordination role. One possibility is to assign a joint CEO that makes all necessary coordinations. However, increased workload of the CEO implies reduced attention for each division, leading to the same kind of communication problem.

firms maintain similar innovation efficiency to less diversified firms when the efficiency is measured by new products per R&D spendings.

The paper is organized as follows; In section Ⅱ, we discuss in details how diversification incurs organizational costs. In section Ⅲ, we set up a formal model in an agency-type framework. We analyze the model and draw main results in section Ⅳ. In section Ⅴ, we discuss the results in the historical context and conclude the paper with the limitations of our analysis.

Ⅱ. ORGANIZATIONAL COSTS OF DIVERSIFICATION

As a motivation, let us consider a hypothetical example. Suppose there are two firms producing goods 1 and 2, respectively. For simplicity, let us treat them as if they each had only one worker. Production process of good 1 involves N_1 decisions (or actions) throughout, while good 2 needs N_2 decisions. Now, suppose benefits exist in carefully coordinating N of these decisions, where $N \leq N_1, N_2$. When two firms are separate entities, coordination would be achieved in the market or through direct bargaining. On the other hand, integration of two firms enables centralized coordination. A typical way to achieve this is by creating a corporate office responsible for N coordinating decisions. Each firm then becomes a division of the integrated firm and still makes $N_1 - N$ and $N_2 - N$ decisions on their own.

Now, suppose that the division producing good 1 comes up with an idea that would increase profit if proper actions are followed. If these actions involve only the decisions the division is entitled to, the cost required for implementation would remain the same. However, if the actions involve some of the centrally coordinated(N) decisions, at least one extra communication step is required to convey the idea to the corporate office. More generally, integration of two firms would increase communication steps in proportion to the current levels of hierarchy. This is because a firm's hierarchy reflects its current requirements of information processing and communication.

There are several reasons that additional communication requirements can be costly. Building communication channels itself requires some positive setup costs, but a more important cost comes from the imperfectness of the media; there exists a fundamental limit in accurately conveying ideas from one person to another. Communication within a firm is generally made through documentations and/or presentations that are prepared under tight deadlines. It is therefore quite likely that useful information is lost or distorted during the process.

We also note that at each communication step managers are located to screen the content before conveying it to the next superior manager. This practice exists to avoid information overload for the top manager, or it could be because lower level managers are entitled to some of relevant decision makings. In either case, the net result is that additional layers of managers make the organization more conservative in implementing new ideas.

Whether it is because of this increased conservatism or the deficiency of communication, additional organizational layers reduce the likelihood that an employee's idea is actually utilized. This poses a significant incentive problem for an employee who exerts efforts to produce valuable business ideas.

A couple of important caveats; first, increased conservatism is not necessarily bad because it also reduces the possibility of adopting wrong ideas. However, firms are generally assumed to be near risk neutral and therefore would not consider the risk factor very importantly in designing organizational structure. Second, preferences of a firm regarding expected gains from decisional changes may be different before and after integration, but its effect on incentives cannot be evaluated uniformly. That is, some actions that are expected to bring in positive net gains for the firm producing good 1 only may no longer be attractive when they are considered by the integrated firm. On the other hand, changes that were undesirable for an independent firm may result in positive net gains for the integrated firm. One could certainly develop a theory based on the characteristics of the business ideas, but this is not the approach we take in this paper. The primary reason is that we are more concerned with the unpredictable sources of ideas, which are difficult to take into account *ex ante*.[7]

Ⅲ. MODEL

Consider a firm that faces an opportunity to diversify into a new market. This could be done either by setting up a new division or by merging to an existing company. For our purpose, we assume there is no difference between the two. Let us define $\varDelta$ to take the value of one if the firm diversifies, and zero otherwise. The firm's organizational designer chooses $\varDelta$ in order to maximize the net benefit. That is, the organizational designer's problem is

$$\max_{\varDelta \in \{0,1\}} B(\varDelta) - C(\varDelta; \Theta) \qquad (1)$$

where B and C respectively represent the expected benefit and cost of diversification, and Θ is the set of parameters that affect the cost.

We do not attempt to fully describe the forms of B and C, because they would differ widely across firms and the situations they face. Instead, we focus on how the cost part of the problem is affected by changes in parameters. If some or all parameters change such that $C(1)-C(0)$ increases, we could be

[7] Sah and Stiglitz (1986) compare two project selecting mechanisms which they call polyarchy and hieararchy. Their treatment of communication cost and assumption of the serial screening for hierarchy have much in common with our comparison of centralized and decentralized coordination. However, their focus is whether a particular structure peforms better when the expected profit distribution of implementing ideas is given. On the other hand, we focus on the possibility that the distribution is affected by the changes in the structure.

certain that fewer firms would choose to diversify *ceteris paribus*. This marginal cost could be the forgone benefit from reduced effectiveness in communication, which we will describe in the following model.

Consider an agent who can exert a costly effort e. e can take two possible values, which we normalize as a zero effort level and a positive level of one: $e \in \{0, 1\}$. By exerting effort, the agent gets disutility of $\phi(e)$, which we normalize such that $\phi(1) = \phi$ and $\phi(0) = 0$. The agent receives a compensation w from the principal. Let us assume that the agent's utility is $U = u(w) - \phi(e)$ with $u(\cdot)$ increasing and concave.

With effort, the agent obtains an innovative business idea which, if properly implemented, increases the value of production. Proper implementation requires a matching set of business decisions that become known only to the agent. We assume that the agent does not have the authority of making relevant decisions. The agent instead suggest necessary recommendations to the proper authorities. This however often requires several steps of reporting and re-processing of information, during which the original idea could be distorted or lost. To model this in the simplest way, we assume that the agent's idea is implemented only with the probability of $p \in (0, 1)$. In other words, $1 - p$ reflects the degree of organizational inefficiencies within the firm.

We also assume that, although the principal can verify whether the agent make any recommendations, she is not capable of evaluating them. The agent can fabricate meaningless reports with little extra effort, therefore reporting by itself does not have any informational value to the principal.[8] In the following, we define η as the random variable that takes Y if the idea is implemented but otherwise takes N.

It has to be emphasized that innovative ideas we consider in this paper is different from typical R&D outputs. For example, firms often have personnels whose primary task is R&D, for whom an organizational designer can specifically reduce the steps of communication to utilize R&D outputs more effectively.[9] In contrast, we are concerned with ideas from ordinary workers, whose primary tasks are mainly routine operations. Although such ideas collectively could be very important for a firm's success, the organizational designer cannot account for the sources of the ideas *ex ante* because they are *unpredictable*. This is not just because a firm's resources are too limited to build a complex communication network. A lot of firms have tried to increase communication efficiency through a dual reporting system called *matrix* organization, in which an agent reports to both product and functional managers. So far, however, not many firms have been able to implement the structure in

[8] This is a simplifying assumption. Our model can be easily extended such that reporting activity itself conveys information and is tied to the agent's compensations.

[9] Many firms have a separate R&D department placed right under the top management in the organizational chart. Diversified firms often set up a corporate wide R&D department that handles big research projects, while leaving small projects at the division level.

full scale. Among the cited problems are heavy burdens for employees and unclear responsibilities because of overlaps in decision authorities.

Next, we assume that production level is stochastic and can only take two values $x \in \{\underline{x}, \overline{x}\}$, with $\overline{x} > \underline{x}$. Therefore, in our model, x and η are observable signals, each of which can take two values. The principal is able to condition the agent's performance on four possible different states of nature. Each of these states is defined in Table 1.

[Table 1] Probability distribution conditional to effort levels

State of nature (x, η)	Probability when $e=0$	Probability when $e=1$
$(\overline{x}, Y)$	$pq\underline{\pi}$	$pq\overline{\pi}$
$(\underline{x}, Y)$	$pq(1-\underline{\pi})$	$pq(1-\overline{\pi})$
$(\overline{x}, N)$	$(1-pq)\underline{\pi}$	$(1-pq)\underline{\pi}$
$(\underline{x}, N)$	$(1-pq)(1-\underline{\pi})$	$(1-pq)(1-\underline{\pi})$

We also introduce another important source of communication costs. Suppose there is a possibility of *breakdown* in which the value of an agent's idea is nullified by an external change of business environment. This could happen for many reasons. For example, a cost saving idea that fits to a certain production technology may become irrelevant with the arrival of a new technology. Likewise, value of an idea that is specific to a narrow product line or a small market area would easily be subject to the fickleness of consumer tastes. In our model, the agent's reports are thrown out without further inspection if a breakdown occurs. We assume that a breakdown occurs with a positive probability $q \in (0, 1)$.

An explanation is in order regarding how we relate our model to the variables we are interested in. First, as we discussed in the previous section, diversification reduces p. The development of information processing technology is also reflected by the shift of p. Suppose ψ is the probability of success for each step of communication. If we assume that an average number of communication steps is $K (K > 0)$, then $p = \psi^K$. Diversification to a new product line increases total steps to $K(1+\alpha)$ for some fixed $\alpha \in (0, 1)$. That is, α reflects communication loss introduced by a firm's increased scope. Since both K and α enter our model only through p, we focus on the effect of p until further analysis is required. On the other hand, q can be considered to measure the instability of market environment, although we need to identify the relationship between this variable and the degree of diversification.

Let us assume that the reservation utility of the agent is zero. The optimal contract that induces a positive effort solves the following problem:

$$\max_{w_1,\ldots,w_4} pq\bar{\pi}(\bar{x}-w_1)+pq(1-\bar{\pi})(\underline{x}-w_2) + (1-pq)\underline{\pi}(\bar{x}-w_3)+(1-pq)(1-\underline{\pi})\ (\underline{x}-w_4) \tag{2}$$

$$s.t.\quad pq\bar{\pi}u(w_1)+pq(1-\bar{\pi})u(w_2)+(1-pq)\underline{\pi}u(w_3) + (1-pq)(1-\underline{\pi})u(w_4)-\phi \geq 0 \tag{IR}$$

$$pq\bar{\pi}u(w_1)+pq(1-\bar{\pi})u(w_2)+(1-pq)\underline{\pi}u(w_3) + (1-pq)(1-\underline{\pi})u(w_4)-\phi \geq pq\underline{\pi}u(w_1)+pq(1-\underline{\pi})u(w_2)+(1-pq)\underline{\pi}u(w_3) + (1-pq)(1-\underline{\pi})u(w_4) \tag{IC}$$

Two constraints of the optimization problem are individual rationality (IR) and incentive compatibility (IC) conditions, respectively. The optimized value of this problem constitutes the value function V. That is, if we denote w_i^* as the solution of the above problem, we define $V(p, q)$ as:

$$V(p,q) \equiv pq\bar{\pi}(\bar{x}-w_1^*)+pq(1-\bar{\pi})(\underline{x}-w_2^*) + (1-pq)\underline{\pi}(\bar{x}-w_3^*)+(1-pq)(1-\underline{\pi})\ (\underline{x}-w_4^*) \tag{3}$$

Based on this setup, we derive the property of V in the next section to describe the cost of diversification.

Ⅳ. ANALYSIS

The maximization problem of (2) is a slightly modified version of a simple moral hazard problem. Therefore, applying the standard techniques leads us to the following characterization:

Proposition 1 *In the optimal contract, both (IR) and (IC) are binding, and the transfer* w *satisfies* $w_1^* > w_3^* = w_4^* > w_2^*$.

Proof. Let μ and λ be the non-negative multipliers associated respectively with the (IR) and (IC) constraints. Applying standard Lagrangian techniques yields the following first order conditions:

$$\frac{1}{u'(w_1)} = \mu + \lambda\frac{\bar{\pi}-\underline{\pi}}{\bar{\pi}} \tag{4}$$

$$\frac{1}{u'(w_2)} = \mu - \lambda\frac{\bar{\pi}-\underline{\pi}}{1-\bar{\pi}} \tag{5}$$

$$\frac{1}{u'(w_3)} = \frac{1}{u'(w_4)} = \mu \tag{6}$$

It is immediate from (6) that $w_3^* = w_4^*$. Multiplying (4) by $\overline{\pi}$ and (5) by $1-\overline{\pi}$, and adding the two equations, we get

$$\mu = \frac{\overline{\pi}}{u'(w_1)} + \frac{1-\overline{\pi}}{u'(w_2)} \tag{7}$$

From (6), (7), and the fact that $u' < \infty$ for nonzero transfer, we infer that $\mu > 0$ and (IR) is binding. Also, using (4) and (7), we get

$$\lambda = \frac{\overline{\pi}(1-\overline{\pi})}{\overline{\pi} - \underline{\pi}} \left(\frac{1}{u'(w_1)} - \frac{1}{u'(w_2)} \right) \tag{8}$$

Rearranging (IC), we have $u(w_1) - u(w_2) \geq \frac{\phi}{\overline{\pi} - \underline{\pi}} > 0$. From this, we get $w_1^* > w_2^*$, and since $u'' < 0$, it follows that λ is strictly positive. Therefore (IR) and (IC) are both binding, and by inspecting (4), (5), and (6), we get $w_1^* > w_3^* = w_4^* > w_2^*$. Q. E. D.

Combining (6) and (7), we get an equation of three unknowns: w_1, w_2, w_3. Also taking two binding constraints into account, we get three equations with three unknowns. Therefore, we can obtain the solution of the optimization problem. In this paper, however, we are not so much interested in finding the explicit solution of the problem as checking the properties of V.

We may invoke standard envelope argument to do comparative statics of V with respect to changes in p *or* q. However, this approach would not tell us about how p *and* q interact through V. The approach also relies upon local neighborhood properties, which provide an awkward framework for an analysis of a regime shift such as diversification. To see this, notice that diversification in our model is represented by an increase of communication steps, which leads to a discrete decrease in the level of p from ψ^K to $\psi^{K(1+\alpha)}$. We are interested in how the decision to diversify is affected by the changes in K, α and q.

The standard approach based on differential calculus cannot address this question properly. However, recently developed methods based on lattice-theory has advantages since it greatly reduces the set of assumptions required for comparative statics analysis. Lattice-theoretic methods are first introduced by Topkis (1978) and further developed by Vives (1990), Milgrom and Roberts (1990), Milgrom and Shannon (1994), and many others. In the Appendix, we provide a brief presentation of the method.

To apply the lattice-theoretic analysis to our model, we need to check the sign of $\frac{\partial^2 V}{\partial p \partial q}$ first. It may be done by expressing V explicitly as a function of p and q. However, this approach is quite cumbersome to apply, as we need to derive closed form solutions for $w_1^*, \ldots, w_4^*$. We therefore adopt an indirect approach to the problem. First, applying envelope theorem, we know $\frac{\partial V}{\partial p} = \frac{\partial L}{\partial p}$

where L is the Lagrangian of the original problem. Further differentiation of V with respect to q yields

$$\begin{aligned}\frac{\partial^2 V(p,q)}{\partial p \partial q} &= \frac{\partial^2 L(w_1^*, \ldots w_4^*, \mu^*, \lambda^*; p, q)}{\partial p \partial q} + \sum_{i=1}^{4} \frac{\partial^2 L}{\partial p \partial w_i}\bigg|_{w_i^*} \cdot \frac{\partial w_i^*}{\partial q} \\ &+ \frac{\partial^2 L}{\partial p \partial \mu}\bigg|_{\mu^*} \cdot \frac{\partial \mu_i^*}{\partial q} + \frac{\partial^2 L}{\partial p \partial \lambda}\bigg|_{\lambda^*} \cdot \frac{\partial \lambda_i^*}{\partial q} \\ &= \frac{\partial^2 L(w_1^*, \ldots w_4^*, \mu^*, \lambda^*; p, q)}{\partial p \partial q}\end{aligned} \tag{9}$$

The second equality comes from the fact that the first order condition should be satisfied for all levels of p. For example, $\frac{\partial^2 L}{\partial p \partial w_i} = \frac{\partial}{\partial p}\left(\frac{\partial L}{\partial w_i}\right)$, but since $\frac{\partial L}{\partial w_i} = 0$ at w_i^* for any p, $\frac{\partial^2 L}{\partial p \partial w_i}\bigg|_{w_i^*} = 0$.

We can now state the next proposition:

Proposition 2 *The principal's value function* V *satisfies the following:*

(i) V *is strictly increasing in* p *and* q

(ii) $V_{pp} = V_{qq} = 0$

(iii) V *is supermodular in* (p, q).

Proof. Let us rearrange the terms of L around pq, then it can be easily seen that L can be represented as the form $pqH + J$, where H and J are independent of p or q. Using the fact that $w_3^* = w_4^*$, we get the following relationship:

$$\begin{aligned}&H(w_1, w_2, w_3^*, w_4^*, \mu^*, \lambda^*) \\ &\quad = \bar{\pi}(\bar{x} - w_1) + (1 - \bar{\pi})\ (\underline{x} - w_2) - \underline{\pi}\bar{x} - (1 - \underline{\pi})\underline{x} + w_3^* \\ &\quad + \mu^*\{\bar{\pi}u(w_1) + (1 - \bar{\pi})\ u(w_2) - u(w_3^*)\} \\ &\quad + \lambda^*(\bar{\pi} - \underline{\pi})\{u(w_1) - u(w_2)\}\end{aligned} \tag{10}$$

$$\begin{aligned}&J(w_1, w_2, w_3^*, w_4^*, \mu^*, \lambda^*) \\ &\quad = \underline{\pi}\bar{x} + (1 - \underline{\pi})\underline{x} - w_3^* + \mu^*\{u(w_3^*) - \phi\} - \lambda^*\phi\end{aligned} \tag{11}$$

We will show that $H(w_1^*, \ldots w_4^*, \mu^*, \lambda^*) > 0$. Let us assume otherwise. If we take $\widetilde{w}_1 = \widetilde{w}_2 = w_3^*$ and substitute them into H, we get

$$H(\widetilde{w}_1, \widetilde{w}_2, w_3^*, w_4^*, \mu^*, \lambda^*) = (\bar{\pi} - \underline{\pi})(\bar{x} - \underline{x}) > 0$$

If $H(w_1^*, \ldots w_4^*, \mu^*, \lambda^*) \leq 0$, it means taking $\widetilde{w}_1$ and $\widetilde{w}_2$ instead of w_1^* and w_2^* strictly increases H without making any changes in J. However, this contradicts to Proposition 1 that w_1^* and w_2^* maximize L. This proves (i) since

$\frac{\partial V(p,q)}{\partial p} = qH$, and $\frac{\partial V(p,q)}{\partial q} = pH$. It also shows that $\frac{\partial^2 V(p,q)}{\partial p \partial q} > 0$ because $\frac{\partial^2 L(w_1^*, \dots w_4^*, \mu^*, \lambda^*; p, q)}{\partial p \partial q} = H(w_1^*, \dots w_4^*, \mu^*, \lambda^*)$. On the other hand, substituting p for q into (9) yields $\frac{\partial^2 V(p,q)}{\partial p^2} = \frac{\partial^2 L}{\partial p^2} = 0$. Likewise, $\frac{\partial^2 V(p,q)}{\partial q^2} = \frac{\partial^2 L}{\partial q^2} = 0$. Q. E. D.

From our earlier discussion, diversification decreases the probability of successful communication by $\psi^K - \psi^{K(1+\alpha)}$, which generates positive costs by strictly reducing V. Therefore, from (1), we may consider the marginal cost $C(1) - C(0)$ as the magnitude of the decrease in V by diversification.

Based on this observation and Proposition 2, we finally state our main proposition:

Proposition 3 *A firm is less likely to diversify if the business environment is more stable (higher q), and communication loss due to diversification is larger (larger α). On the other hand, the relationship between the current level of hierarchy (K) and diversifying decision is mixed; if internal communication is generally effective (both ψ^K and $\psi^{K(1+\alpha)}$ are sufficiently large), diversification is more likely with smaller hierarchy. If internal communication is currently ineffective(ψ^K is small), diversification is more likely with larger hierarchy.*

Proof. Let us define $\tilde{V}(\Delta; K, \alpha, q) \equiv V(\psi^{K(1+\alpha\Delta)}, q)$. From our earlier discussion, we may reformulate the problem (1) as

$$\max_{\Delta \in \{0,1\}} B(\Delta) + \tilde{V}(\Delta; K, \alpha, q) \tag{12}$$

which should yield the same comparative statics. Since this maximization problem has one variable and three parameters, to apply Theorem A in the Appendix, we only need to check if $\tilde{V}$ has increasing differences in Δ and the parameters. For notational simplicity, let us treat Δ and K as if they were continuous. This is an innocuous assumption because the signs of the derivatives remain the same within the intervals that we examine. We then should check the signs of $\frac{\partial^2 \tilde{V}}{\partial K \partial \Delta}$, $\frac{\partial^2 \tilde{V}}{\partial \alpha \partial \Delta}$, and $\frac{\partial^2 \tilde{V}}{\partial q \partial \Delta}$.

First, since $\psi^{K(1+\alpha\Delta)}$ is a decreasing function of Δ, $\frac{\partial^2 \tilde{V}}{\partial q \partial \Delta} = \frac{\partial^2 V}{\partial q \partial p} \cdot \frac{\partial \psi^{K(1+\alpha\Delta)}}{\partial \Delta} < 0$. Next, applying Proposition 2, we know

$$\begin{aligned} \frac{\partial^2 \tilde{V}}{\partial \alpha \partial \Delta} &= \frac{\partial^2 V}{\partial p^2} \cdot \frac{\partial \psi^{K(1+\alpha\Delta)}}{\partial \Delta} \cdot \frac{\partial \psi^{K(1+\alpha\Delta)}}{\partial \alpha} + \frac{\partial V}{\partial p} \cdot \frac{\partial^2 \psi^{K(1+\alpha\Delta)}}{\partial \Delta \partial \alpha} \\ &= \frac{\partial V}{\partial p} \cdot \frac{\partial^2 \psi^{K(1+\alpha\Delta)}}{\partial \alpha \partial \Delta} \end{aligned} \tag{13}$$

Therefore, the sign of $\frac{\partial^2 \tilde{V}}{\partial \alpha \partial \Delta}$ is the same as $\frac{\partial^2 \psi^{K(1+\alpha\Delta)}}{\partial \Delta \partial \alpha}$. Likewise, $\frac{\partial^2 \tilde{V}}{\partial K \partial \Delta}$ has the same sign as $\frac{\partial^2 \psi^{K(1+\alpha\Delta)}}{\partial \Delta \partial K}$. Differentiation of $\psi^{K(1+\alpha\Delta)}$ with respect to Δ and α yields

$$\frac{\partial^2 \psi^{K(1+\alpha\Delta)}}{\partial \Delta \partial \alpha} = K \log \psi \{ \alpha K \Delta \log \psi \} \psi^{K(1+\alpha\Delta)} > 0 \tag{14}$$

On the other hand, differentiation with respect to Δ and K yields

$$\frac{\partial^2 \psi^{K(1+\alpha\Delta)}}{\partial \Delta \partial K} = \alpha \log \psi [1 + (1 + \alpha\Delta) k \log \psi] \psi^{K(1+\alpha\Delta)} \tag{15}$$

The sign of (15) critically depends on the terms inside the bracket. It turns out that $\frac{\partial^2 \psi^{K(1+\alpha\Delta)}}{\partial \Delta \partial K} \leq 0$ if and only if $\psi^{K(1+\alpha\Delta)} \geq \frac{1}{e}$, where e is the natural number. Therefore, $\psi^{K(1+\alpha)} \geq \frac{1}{e}$ implies V having increasing differences in $(\Delta, -K)$, and $\psi^K \leq \frac{1}{e}$ implies V having increasing differences in (Δ, K). On the other hand, the sign of $\frac{\partial^2 \psi^{K(1+\alpha\Delta)}}{\partial \Delta \partial K}$ is undetermined if $\psi^{K(1+\alpha)} < \frac{1}{e} < \psi^K$. The rest of the proof follows by applying Theorem A in the Appendix.

Q. E. D.

First two results of Proposition 3 are intuitive. A stabler business environment makes breakdown in communication less likely, resulting in higher expected value of communication. Under the circumstance, diversification becomes more costly to the firm because it reduces the probability of successful communication. Likewise, it is obvious that diversification is more costly if it incurs greater loss of communication.

On the other hand, the latter part of Proposition 3 shows that the level of hierarchy would have quite different implications. Let us remind that communication loss (measured by reduction in p) is $\psi^K(1-\psi^{\alpha K})$. Increased K reduces the first term, while increasing the second term. For a sufficiently small level of ψ, the first effect dominates, but the second effect dominates for a large ψ. Intuitively, increasing steps of communication geometrically reduces the probability of success. This is increasingly costly when communication is overall effective (large ψ), but the cost increase is mild if communication is already quite bad (small ψ).

V. DISCUSSION AND CONCLUSION

In this paper, we presented a formal model of diversification costs and derive some testable predictions; a firm is less likely to diversify when (i) the business environment is unstable, (ii) the level of hierarchy is small if communication is effective, (iii) the level of hierarchy is large if communication is ineffective. Whether these predictions are valid should be the subject of carefully designed empirical tests, which we leave as our future projects. Instead, we present a couple of historical events (albeit circumstantial evidence) that we think are relevant with our predictions.

Researchers have shown that R&D intensities and high growth rate of the target industry are two most important variables that explain the cross sectional pattern of corporate diversification (MacDonald, 1985). That is, diversification is more likely to occur between two R&D intensive industries and toward fast growing industries. In comparison, few studies are available regarding the characteristics of the industries that firms tend to divest from. However, in her study of corporate restructuring in the United States, Hall (1994) reports that

> [L]everaged buyouts and other private acquisitions of publicly traded manufacturing firms had taken place overwhelmingly in the sectors where R&D investment and innovation have not been important, at least to the industry as a whole. The industries and firms in question were those generating the steady cash flow necessary to service the added debt.

Although further investigation is necessary, this could be indirect evidence that firms de-diversify the divisions operating in the industry whose R&D and growth opportunities are exhausted. This is especially because LBOs and hostile takeovers were often motivated by the claims that the portfolio of the target firm is inefficiently structured.

Both high R&D and high growth rate reflect the instability of business environment described in this paper. The fact that R&D is important indirectly implies that technology is changing fast, increasing the likelihood that technology or product specific ideas are left unused. Also, it means that R&D has the priority of the resource allocation, implying that ideas for other aspects such as marketing are not likely to get the resources necessary for a trial. On the other hand, the fact that an industry is growing fast implies that the business environment is highly uncertain. Limited resources argument could be also applicable here as resources are likely to be used to increase the capacity of the production rather than for trials of innovative ideas benefits from which are highly uncertain.

On the other hand, there is empirical evidence that diversification of large firms increased until the 1970s but decreased afterward. Rumelt (1974) documents that 70 percent of major firms in the US drew more than 70 percent of sales from a single activity in 1949, but only 35 percent showed a similar focus in 1969. In contrast, using the data from Fortune 500 US firms, Davis, Dieckman, and Tinsley (1994) show that their measure of diversification decreased by 33 percent from 1980 to 1990. Likewise, Markides (1995) shows that 20 percent of Fortune 500 US firms were refocusing their businesses in 1980s, while only 8 percent were diversifying. As a comparison, only 1 percent of the top companies were refocusing in 1960s, while 25 percent were diversifying.[10]

Although we cannot make sure that this trend is actually related to our results, it seems to fit well with the predictions of Proposition 3. If we accept that large firms tend to have deeper hierarchies, the proposition implies that large firms are more likely to diversify when internal communication is ineffective. However, as communication technologies improve over the years, the relationship could be reversed so that large firms tend to de-diversify.

With the development of information and communication technologies, overall effectiveness of internal communication in fact has improved over the years. Diffusion of personal computers with powerful documentation and spreadsheet software greatly reduced the cost of report preparations. Reduced telecommunication costs, inventions such as FAX and e-mails, and even reduced travel costs all contributed to improved communication within firms, especially for those which operate in geographically vast areas.

We conclude the paper by pointing out some aspects that are not considered in this paper. First, in this paper we do not consider the methods of diversification, although their implications would be quite important in practice. For example, whethere a firm create a new division internally or merge into an already existing outside firm could make huge differences in terms of reorganization requirements. Further research is in need to sort out various modes of diversifications and their implications.

On the other hand, financial motivations are not considered in this paper, although they seem to be among the most important causes for diversification in practice. However, this does not mean that such concerns are unimportant. It is believed that imperfect financial market can create the incentives for diversification.

Another point to consider is the claim that de-diversification occurs because markets are changing faster and become more consumer-oriented, requiring firms that are more flexible in nature. In our paper, these factors are the results of the refocusing rather than the causes. In fact, causality in either direction is possible because individual firms need to adjust to the environmental changes whether these changes are purely exogenous or result from the other firms' actions. Still, we believe our approach is more appropriate because it is not easy to find intuitive reasons to believe that these factors are purely exogenous.

---

[10] Korean firms show similar historical patterns, although major refocusing trend occurred only after the currency crisis in late 1990s. Although the crisis is the single direct cause, it is not clear whether Korean firms were simply delaying inevitable reorganizations even without the crisis.

REFERENCES

Amihud, Yakov & Baruch Lev (1981), "Risk Reduction as a Managerial Motive for Conglomerate Mergers," *Bell Journal of Economics*, Autumn, 1981, 12, pp.605-17.

Avery, C., J.C. Chevalier, and S. Schaefer (1998), "Why Do Managers Undertake Acquisitions? An Analysis of Internal and External Rewards for Acquisitiveness," *Journal of Law, Economics, and Organization*, 14, pp.24-43

Cardinal, Laura B. & Tim C. Opler (1995), "Corporate Diversification and Innovative Efficiency: An Empirical Study," *Journal of Accounting & Economics*, 19, pp.365-81.

Caves, Richard E. (1982), *Multinational Enterprise and Economic Analysis*, Cambridge, Cambridge Univ. Press.

Davis, G.F., K.A. Diekman and C.H. Tinsley (1994), "The Decline and Fall of the Conglomerate Firm in the 1980s: The De-Institutionalization of an Organizational Form," *American Sociological Review*, 59, August, pp. 547-570.

Hall, Bronwyn (1994), "Corporate Restructuring and Investment Horizons in the United States, 1976-1987; Competitiveness and Capital Investment: The Restructuring of U.S. Industry, 1960-1990" *Business History Review*, March 24.

Holmstrom, Bengt & Jean Tirole (1989), "The Theory of the Firm" in Schmalensee, R. and R.D. Willig ed. *Handbook of Industrial Organization*, Vol. 1, Elsevier.

Jensen, Michael C. (1986), "Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers," *American Economic Review*, May, 76, pp.323-29.

Lewellen, Wilbur G., Claudio Loderer, & Ahron Rosenfeld (1985), "Merger Decisions and Executive Stock Ownership in Acquiring Firms," *Journal of Accounting and Economics*, April, 7, pp.209-31.

MacDonald, James M.(1985), "R&D and the Directions of Diversification," *Review of Economics and Statistics*, pp.583-590.

Markides, Constantinos (1995), *Diversification, Refocusing, and Economic Performance*, MIT Press, Cambridge.

Milgrom, Paul & John Roberts (1990), "Rationalizability, Learning and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58, pp.1255-1278.

Milgrom, Paul & Chris Shannon (1994), "Monotone Comparative Statics," *Econometrica*, Vol. 62, No. 1, January.

Montgomery, Cynthia, A. (1994), "Corporate Diversification," *Journal of Economic Perspective*, Vol. 8, No. 3, summer, pp.163-78.

Morck, Randall, Andrei Schleifer, & Robert W. Vishny (1990), "Do Managerial Objectives Drive Bad Acquisitions?" *Journal of Finance*, March, 45, pp.31-48.

Mueller, Dennis C. (1972), "A Life Cycle Theory of the Firm," *Journal of Industrial Economics*, July, 20, pp.199-219.

Penrose, Edith T. (1959), *The Theory of the Growth of the Firm*, Oxford, Basil Blackwell.

Prahalad, C.K. and R.A. Bettis (1986), "The Dominant Logic: A New Linkage Between Diversity and Performance," *Strategic Management Journal*, 7, pp. 485-501.

Porter, Michael, E. (1987), "From Competitive Advantage to Corporate Strategy," *Harvard Business Review*, May-June.

Rotemberg, Julio J. & Garth Saloner (1994), "Benefits of Narrow Business Strategies," *American Economic Review*, December, pp.1330-49.

Rumelt, R. (1974), *Strategy, Structure, and Economic Performance*, Boston; Division of Research, Harvard Business School.

Sah, Raaj K. & Joseph E. Stiglitz (1986), "The Architecture of Economic Systems: Hierarchies and Polyarchies," *American Economic Review*, 76(4), pp.715-25.

Shleifer, A. & R.W. Vishny (1986), "Large Shareholders and Corporate Control," *Journal of Political Economy*, pp.461-468.

________________________ (1989), "Management Entrenchment: The Case of Manager-Specific Investments," *Journal of Financial Economics*, No. 25, pp.123-39.

Teece, David J. (1982), "Towards an Economic Theory of the Multiproduct Firm," *Journal of Economic Behavior and Organization*, March, 3, pp.39-63.

Topkis, Donald M. (1978), "Minimizing a Submodular Function on a Lattice," *Operations Research*, 26, 305-321

Vives, Xavier (1990), "Nash Equilibrium with Strategic Complementarities," *Journal of Mathematical Economics*, 19, 305-321.

Williamson, Oliver (1975), *Markets and Hierarchies: Analysis and Antitrust Implications*, New York, Free Press.

Williamson, Oliver (1985), *The Economic Institutions of Capitalism*, New York, Free Press.

Appendix: Lattice-theoretic methods of comparative statics

This paper applies lattice-theoretic methods of comparative statics. To make our paper self-contained, we provide some basic results that are used in this paper. Those who are interested should refer to the original papers.

Let us consider the general optimization problem:

$$V(\theta) = \max_{x \in X} f(x; \theta)$$

where f is a real valued function ($f: X \times \Theta \rightarrow R$), $X \subseteq R^N$ is a lattice, and Θ is a partially ordered set.[11] We want to examine how the set of optimal choices $X^*(\theta)$ responds to a change in θ.

For $x, y \in X$, let us denote their "join" to be $x \vee y$, which is defined as the number whose coordinates are $\max\{x_i, y_i\}, \forall i$. Likewise, we define "meet" ($x \wedge y$) to be the number whose coordinates are $\min\{x_i, y_i\}, \forall i$.

Definition A1 *A function f is supermodular in x if, for all $x, y \in X$,*

$$f(x \vee y; \theta) + f(x \wedge y; \theta) \geq f(x; \theta) + f(y; \theta)$$

If f is differentiable, this condition is equivalent to $\frac{\partial^2 f}{\partial x_i \partial x_j} \geq 0, \forall i \neq j$.

Definition A2 *A function f has increasing differences in (x, θ) if, for all $x' \geq x$ and $\theta' \geq \theta$,*

$$f(x', \theta') - f(x, \theta') \geq f(x', \theta) - f(x, \theta)$$

If f is differentiable, this condition is equivalent to $\frac{\partial^2 f}{\partial x_i \partial \theta_j} \geq 0, \forall i \neq j$.

It is easy to see that increasing differences is a weaker condition. That is, if f is supermodular in (x, θ), then f has increasing differences in (x, θ).

To incorporate the cases of multiple optimal choices, we use the concept of set ordering for comparative statics:

Definition A3 *Let $\theta, \theta' \in \Theta$. We say $X^*(\theta)$ is higher than $X^*(\theta')$ in set ordering ($X^*(\theta) \geq_S X^*(\theta')$) if, for $x \in X^*(\theta)$ and $x' \in X^*(\theta')$, $x \vee x' \in X^*(\theta)$ and $x \wedge x' \in X^*(\theta')$*

[11] A *partially* ordered set is a nonempty set X with a relation R defined on $X \times X$ where R satifies reflexivity (xRx), transitivity(xRy, yRz then xRz), and antisymmetry (xRy and yRx implies $x = y$). Consider a partially ordered set $(X, \geq)$. If $\forall x, y \in X$, $inf(x, y) \in X$ and $sup(x, y) \in X$ then X is a *lattice*.

Finally, we state the following main theorem by Topkis (1978):

Theorem A *(Topkis, 1978) Assume* $X \subseteq R^N$ *is a lattice,* $\Theta \subseteq R^N$ *is a partially ordered set. If* $f: X \times \Theta \rightarrow R$ *is supermodular in* x *and has increasing differences in* (x, θ)*, then*

$$X^*(\theta) \geq_S X^*(\theta'), \forall \theta \geq \theta'.$$

Supermodularity is not the necessary condition for this type of comparative statics, but it provides a simple condition to check as we only need pairwise cross derivatives. In particular, if f is differentiable, we only need to check the signs of $\frac{\partial^2 f}{\partial x_i \partial x_j}, \forall i \neq j$ and $\frac{\partial^2 f}{\partial x_i \partial \theta_k}, \forall i \neq k$.