RIETI Discussion Paper Series 02-E-001

Incentives and Option Value in the Silicon-Valley Tournament Game

AOKI Masahiko
RIETI

TAKIZAWA Hirokazu
RIETI

Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

RIETI Discussion Paper Series 02-E-001

Incentives and Option Value in the Silicon-Valley Tournament Game[1]

Masahiko Aoki

Department of Economics, Stanford University
Stanford, California 94305
&
RIETI
1-3-1 Kasumigaseki, Chiyoda-ku, Tokyo 100-0013, Japan

and

Hirokazu Takizawa

Faculty of Economics, Toyo University
5-28-20 Hakusan, Bunkyo-ku, Tokyo 112-8606, Japan

Abstract

This paper analyzes the "Silicon Valley model" as a novel economic institution in the domain of technological product system innovation such as computers. We focus on the information structural relationship as well as governance relationships between venture capitalists and a cluster of entrepreneurial firms. The informational conditions under which the Silicon Valley model is efficient are identified, leading to understanding the significance of standardization of interfaces, modularization and information encapsulation. We then examine the governance/incentive aspect of the model by integrating the models by Aoki (2001) and Baldwin and Clark (2000) to give comparative statics results regarding the optimal number of entrepreneurial firms competing in the same component product. The analyses enable us to evaluate the applicability of the model beyond specific localities and industries.

JEL Classification Numbers: D21;L23;O32;P51;P52

Keywords: Modular design; Silicon Valley; Option value; Tournament Game

[1]This study has been prepared within the UNU/WIDER project "Property Rights Regimes, Microeconomic Incentives and Development." We are grateful to Laixiang Sun, the project director, and two anonymous referees for helpful comments and suggestions.

1 Introduction

In the aftermath of the so-called dot.com bubble and crash, the previous enthusiasm for the Silicon Valley phenomena seems to have faded away somewhat. The fact still remains, however, that Silicon Valley has been successful in bringing a lot of outstanding entrepreneurial firms into existence. What mechanism made Silicon Valley a major driving force for product system innovation, especially in the information and communications industry? Can it be transplanted into a wide variety of local and industrial domains beyond Silicon Valley? The purpose of this paper is to analyze the Silicon Valley phenomena as a novel economic institution in the domain of technological product system innovation.

The most conspicuous example of the Silicon Valley phenomena can be found in the computer industry. As is documented by Baldwin and Clark (2000), the computer industry was virtually a monopoly market dominated by IBM for a long time until the early 1970s. However, a group of entrepreneurial firms, mostly small and funded by venture capitalists, have been set up since the 1970s and have been very agile in R&D activities. The apparent feature common to these entrepreneurial firms is that they usually develop and produce modular parts of a product system, rather than competing with IBM by producing a stand-alone product system. Many new sub-industries have thus been formed within the domain of the traditional computer industry, and a variety of R&D activities traditionally conducted within IBM are now conducted independently outside. This process has drastically changed the landscape of the computer industry. A new product system is now evolutionarily formed by selecting and combining *ex post* new modular products developed by entrepreneurial firms. In this sense, we may say that a novel and unique economic institution has emerged in the domain of product system innovation. We will henceforth call this system of product system innovation the "Silicon Valley Model"(Aoki, 2001).

At first sight, it might appear that the property rights theory, as developed by Grossman and Hart (1986) and Hart and Moore (1990), can be applied to explain why R&D

activities previously conducted within an established integrated firm began to be conducted independently by small entrepreneurial firms. However, this approach cannot easily explain the unique manner of processing information that is prevalently observed in Silicon Valley. Indeed, as Saxenian (1994) points out, in Silicon Valley there are substantial degrees of information sharing across competing entrepreneurial firms on the one hand, and information hiding (encapsulation) on the other. Understanding these ostensibly contradictory phenomena is the key to understanding the Silicon Valley model.[2]

Baldwin and Clark (2000) is an attempt to understand the Silicon Valley model by focusing on how information is processed in the design of a product system. They submit that the "modular design" of a complex system like a computer is the key concept for understanding the emergence of a large modular cluster of firms and markets in the computer industry. While their explanation of the power of modularity is persuasive, they do not explicitly analyze the incentive aspects of the Silicon Valley model. We submit that it is not sufficient to analyze the Silicon Valley model only from the information systemic aspects or the governance aspects. We extend their model of "substitution operator" by explicitly considering the incentives of each entrepreneur.

The Silicon Valley phenomena contain multifaceted interactions between a cluster of entrepreneurial start-up firms on the one hand, and venture capitalists (as well as leading firms in respective niche markets) on the other. In order to properly capture the essential nature of this model, it is necessary to identify the unique roles played by those actors. The next section offers our modeling background by describing stylized facts about their relationships. We submit that it is not sufficient to look only at the property rights relationship between a venture capitalist and a single entrepreneurial firm, and that the venture capitalists usually have dual roles in their relationships with entrepreneurial firms: their role as a mediator of information and that of structuring governance. In Section 3, we develop a team-theoretic model to capture the information-processing activities of venture capitalists and entrepreneurial firms in the course of R&D activities. This enables us to compare different R&D organizations and to identify the conditions under which the Silicon Valley model can be superior to a traditional R&D organization in large integrated firms. Section 4 formulates the relationship between a venture capitalist and entrepreneurial firms as a "VC tournament game," and analyzes the governance role played by venture capitalists. We extend the model in Aoki (2001) by endogenizing the number of entrepreneurial firms competing in the same component product. Our model can also be regarded as a natural extension of the model of Baldwin and Clark (2000), in which the developmental effort level by each entrepreneurial firm is an exogenous variable. Using this integrated model, we show how the effectiveness of the powerful "substitution operator" in Baldwin and Clark (2000) is limited by incentive considerations. Section 5 concludes the paper by evaluating the applicability of the Silicon Valley model beyond specific localities and industries.

2 Stylized Facts as the Modeling Background

Venture capital funds do not usually finance an entrepreneurial firm at too early a stage in its development. Angel investors often fill the need for smaller amounts of start-up capital. Angel investors are individuals who invest their own wealth in start-ups that are not directly related to them through family or prior friendship. In Silicon Valley, a particular type of angel investor, the successful executive who has made his/her fortune in his/her own company, has recently become increasingly important, and there is a close relationship between angels and venture capitalists. In this paper, we do not explicitly differentiate among venture capital funds, venture capital companies and angel investors, meaning that we refer to all of them simply as venture capitalists.

Venture capitalists seek promising investment projects, while potential entrepreneurs with planned projects but insufficient funds seek venture capital financing. There are more than 200 venture capital companies in Silicon Valley, and experienced venture capitalists are said to receive over 1,000 applications per year. Suppose that a promising match is found. Unless the reputation of an entrepreneur is already known to venture capitalists and the proposed project is judged to be certainly sound and promising, the venture capitalist initially provides only seed money to see if the entrepreneur is capable of initiating the project, while possibly extending aid to help the start-up. When a venture capitalist decides to finance a start-up, elaborate financing and employment agreements are drawn up between the venture capitalist and the entrepreneur.

At the time of start-up, the venture capitalist commits only a fraction of the capital needed to complete the project, with the expectation that additional financing will be made stepwise, contingent upon the project proceeding smoothly, which may not be contractible. This is a process that Sahlman (1990) called "staged" capital commitment. Financing by venture capitalists normally takes the form of convertible preferred stocks or subordinate debt with convertible privileges (Kaplan and Strömberg, 2000; Gompers and Lerner, 1996). This means that they are paid prior to holders of common stock in the event of project failure. Also they retain an exit option exercisable by refusing additional financing at a critical moment when a start-up firm needs an infusion of new funds to survive. However, a typical share-holding agreement allows an entrepreneur to increase his/her ownership share (normally in common stock) at the expense of investors, if certain performance objectives are met. Fired entrepreneurs forfeit their claims on stock that has not been vested.

Venture capitalists are well represented on the boards of directors of start-up firms. In addition to attending board meetings, leading venture capitalists often visit entrepreneurs *cum* senior managers at the site of venture-funded firms. They provide a wide range of advice and consulting services to senior management; help to raise additional funds; review and assist with strategic planning; recruit financial and human resource managers; introduce potential customers and suppliers; and provide public relations and legal specialists. They also actively exercise conventional roles in the governance of the start-up firms, often firing the founder-managers when needed.

If the project is successful, the relational financing terminates either with an initial public offering (IPO) or with acquisitions by other firms. Capital gains are distributed between the venture funds and the entrepreneur according to their shares at that time.

Before the dot.com bubble, it usually took five to seven years for the start-up firms to go to the IPO market. During the dot.com boom, this period was shortened, especially for e-commerce businesses. This is because the technology involved was not strikingly innovative in those businesses, and only new business models had to be contrived. For example, basic analytical algorithms of Internet auction sites have long been known in experimental economics. By contrast, in the biotechnology industry where R&D uncertainty is still relatively high, the period needed for the recovery of venture-capital investment returns has not been shortened significantly. After the crash, the period has tended to get longer again in the information and communications industry.

Recently successful start-up firms show the tendency to become targets of acquisition by leading firms in the same market rather than going to IPO markets. From the viewpoint of start-up entrepreneurs, they are said to prefer buy-outs to IPO's, particularly when they have only a single innovative product line (Hellman, 1998). Those acquiring firms are often themselves grown-up entrepreneurial firms that have been successful in assuming leadership in setting standards in their niche markets. Their aim is to acquire successful start-up firms, either to kill off potential sources of challenges to the standards they set, or to further strengthen their market positions by shortening the period of in-house R&D through acquisition and development (A&D). They also seek to establish a monopolistic position in the market by bundling complementary technologies. By doing so, these leading firms have exerted great influence on venture capitalists and entrepreneurial firms in guiding their activities. This mechanism as a whole enables a new technological product system to be formed evolutionarily by combining flexible new modular products *ex post*.

For the above mechanism to work, it is necessary that the standardized interfaces are prescribed among different modular products and that information-processing activities are encapsulated and/or hidden within each entrepreneurial firm in the course of developing respective modular products. This is a unique mechanism of information sharing/hiding that Saxenian (1994) found to be the key to the innovative nature of Silicon Valley firms. Standardization of interfaces is as much a product of the architecture defined by dominant firms (especially Cisco Systems and Microsoft in the current era) and of industry standard-setting organizations, such as Semiconductor Equipment and Materials International (SEMI) and the Internet Engineering Task Force (IETF), as a product of coordination by venture capitalists. Similarly, firms like Sun are competing with products like Jini and Java to define the interface standards for emerging markets. Even the leading positions of established firms in respective niche markets may not be secure in highly uncertain and competitive technological and market environments. Rather standards may be conceived to be evolutionarily formed and modified through the interaction of firms, large and small. In this process, venture capitalists also play an important role in intermediating necessary information among these actors, especially entrepreneurial firms.

The above discussion indicates that the venture capitalists play a wide range of roles vis-à-vis entrepreneurial firms, which include *ex ante* monitoring, i.e., screening of proposed projects to cope with the possible adverse selection problem; *ad interim* monitoring; *ex post* monitoring, i.e., the verification of a project result and the controlling decision as to which exit strategy is to be exercised; and mediation of information regarding standardization of interfaces. These functions are of course not fulfilled exclusively by a single venture capitalist. *Ex ante* monitoring requires risk-taking entrepreneurial instinct and ability to draw road maps of technological development. *Interim* monitoring requires professional engineering competence in specialized fields and management skills. *Ex post* monitoring requires financial expertise. As a consequence, specialization among venture capitalists emerges to meet the different monitoring needs at the different development stages of an entrepreneurial firm. We abstract from such complications in the real world and assume that a single venture capitalist fulfills all these functions.

During the dot.com bubble, a large number of start-up entrepreneurial firms were set up under the above described mechanism, and many of them have suffered losses or disappeared. These events might lead one to doubt the viability of the Silicon Valley model. However it should be stressed that the model had been effective even before the dot.com bubble, and the crash just returned things back to the way they had been. The cause of the bubble can be attributed to the lack of rational expectation on the side of investors regarding the value to be realized (Baldwin and Clark, 2001). The mechanism as such still remains effective for creating value and therefore deserves to be examined.

3 The Information-Systemic Aspect of the Silicon Valley Model

3.1 Comparative R&D Organizations

The previous section suggests that one of the major roles of venture capitalists lies in the mediation of information, and the formation of a new product system by selecting and combining modular products *ex post*. Thus it would be natural to ask under what conditions such a unique arrangement of R&D activities can be superior to traditional R&D organizations in a large integrated firm.

Suppose that a new technological product system is created by combining component products. For example, a laptop computer as a technological product system consists of such component elements as an LC monitor, MPU, image-processing LSI, a hard disk drive, OS, application software, audio and communication devices, etc. In general, there are complicated dependencies among the design tasks of those component products. Therefore, developing a complex product system requires continual coordination among design tasks of different component products so that they will fit with one another to form a coherent product system.[3] The volume of information exchanged and processed among those design task units can be so huge that any single agent would not be able to marshal the whole process in a centralized manner. Since each human being is boundedly rational in his/her information-processing activity, we usually form an organization to partially transcend human limitations, and attempt to solve the problem by installing a structured information-processing system.

In order to capture such structured information-processing activities inherent in the development of a complex product system, suppose that a generic R&D organization is composed of a development manager denoted by M, and two product design teams denoted by T_i ($i = a, b$). M is engaged in such tasks as development strategy, the allocation of R&D funds, and so forth, while the product design teams are engaged in the design of component products of an integral technological system. They coordinate their activities so as to maximize the value of the product system in uncertain environments. Their activities are assumed to be segmented as follows. There is a systemic segment E_s (systemic environment, hereafter), say the availability of total R&D funds and emergent industrial standards, that simultaneously affects the organizational returns to decision choices by M as well as T_i's. Next, there are segments of environments that affect the organizational returns to decision choices by T_i's, engineering environments, which can be further divided into three subsets: E_e (systemic engineering environment, hereafter) common to both teams, and E_a and E_b (idiosyncratic engineering environment, hereafter) idiosyncratic to respective projects of the teams.

Assuming that the activities of each member are aligned linearly, the above described situation can be formulated by using a team-theoretic model developed by Marschak and Radner (1972). Suppose that the value of the technological product system, which is also the payoff common to all the members, is expressed as[4]

$$\begin{aligned} V(x, y_a, y_b) &= \gamma_s x + (\gamma_s + \gamma_e + \gamma_a) y_a + (\gamma_s + \gamma_e + \gamma_b) y_b \\ &\quad - \frac{A}{2} x^2 + D x (y_a + y_b) - \frac{K}{2} (y_a + y_b)^2 - \frac{L}{2} (y_a - y_b)^2 \end{aligned} \tag{1}$$

where x is the choice variable by M, and y_i's are choice variables by T_i's. There are two kinds of parameters in this payoff function: stochastic parameters and constant parameters. Constant parameters are related to technological complementarity among the members' activities, while stochastic parameters perturb the returns to those activities. Specifically, γ_s, γ_e, γ_a, and γ_b are stochastic parameters expressing uncertainty arising in environment E_s, E_e, E_a, and E_b respectively. Observe that γ_s affects the returns to x as well as y_i's, and γ_e affects those to y_i's, while γ_i affects only y_i. The members can do better by adjusting their activity levels based on the obtained information regarding those stochastic variables. Constant parameters are K, L, A, and D. Note that because $\partial^2 V/\partial y_a \partial y_b = L - K$ measures the degree of technological and/or design attribute complementarity, the choice variables of T_i's are complementary when $K < L$, and are substitutes when $K > L$. It would be natural to assume that the choice variables of M and T_i's are complementary (namely $\partial^2 V/\partial x \partial y_i = D > 0$). The value function is assumed to be strictly concave. Under the above assumptions, the sufficient conditions for the value function to be strictly concave in (x, y_a, y_b) are $A > 0$, $K+L > 0$, $AK - D^2 > 0$.[5] Without loss of generality, let K and L be positive, because any value of $K + L$ and $K - L$ can be produced by appropriately selecting positive K and L.

In what follows, we assume that M is engaged in observing E_s, and E_i's are observed only by T_i's $(i = a, b)$. Other specifications about observation and/or information sharing via communication will characterize each type of R&D organization. In the sense that any agent cannot observe all environmental variables, and thus has to base his/her decision only on partial information, this is a second-best situation. Also assume that all the observations of environmental variables are accompanied by some error due to bounded rationality. In this team-theoretic setting, at first the R&D organization decides how to share the various kinds of information among the members, although complete information sharing is impossible as stated above. Given such an information structure, it then adopts second-best decision rules that maximize the expected payoff. A decision rule maps pieces of available information to choice variables. We are interested in what type of R&D organization emerges as the one that most successfully coordinates agents' choice variables under a specific set of parameters. A type of R&D organization is defined to be *informationally more efficient* than another if the maximized expected payoff to it is greater than that to another, which means that one type of R&D organization is superior to the other type as a coordination system under a given set of parameters.

We further assume that all environmental shocks are normally distributed with a mean of zero. The errors that accompany observation of E_s, E_e, E_a, and E_b are denoted by ϵ_s, ϵ_e, ϵ_a, and ϵ_b respectively. They are also assumed to be independently and normally distributed with a zero mean. Thus

$$\gamma_s \sim N(0, \sigma^2_{\gamma_s}) \quad \gamma_e \sim N(0, \sigma^2_{\gamma_e}) \quad \gamma_i \sim N(0, \sigma^2_i)(i = a, b)$$

and

$$\epsilon_s \sim N(0, \sigma^2_{\epsilon_s}) \quad \epsilon_e \sim N(0, \sigma^2_{\epsilon_e}) \quad \epsilon_i \sim N(0, \sigma^2_{\epsilon_i})(i = a, b)$$

Other errors due to the communication process or that are distinctive of a specific type of organization will be defined when it is necessary.

(1) Hierarchical R&D Organization In this type of R&D organization, M is the research manager of an integrated firm and T_i's are its internal project teams. Inserted between them is an intermediate agent IM, say a system engineer. M is specialized in monitoring E_s. Let us denote M's observation by $\xi_s = \gamma_s + \epsilon_s$, which is communicated to IM. IM is engaged in monitoring E_e as well as communicating M's and his/her observation to T_i's. We denote IM's own observation by $\xi_e = \gamma_e + \epsilon_e$. Thus T_i's receive ξ_s and ξ_e with some communication errors, as well as observe $\xi_i = \gamma_i + \epsilon_i$. As a result, in this mode, M's choice variable x depends upon ξ_s, and T_i's choice variable y_i depends upon $\xi_s + \epsilon_{si}$, $\xi_e + \epsilon_{ei}$, and ξ_i, where ϵ_{si} and ϵ_{ei} denote the communication errors on the side of T_i. We assume that $\epsilon_{sa}, \epsilon_{sb} \sim N(0, \sigma^2_{se})$, $\epsilon_{ea}, \epsilon_{eb} \sim N(0, \sigma^2_{ee})$ and they are all independent.

This type of organization can be regarded as reflecting the essential aspects of the R&D organization of a traditional, large hierarchical firm, sometimes referred to as the "waterfall" model (Klein and Rosenberg, 1986; Aoki and Rosenberg, 1989).

(2) Interactive R&D Organization In this type of R&D organization, M is the research manager and T_i's are interacting development teams. There is information sharing among them all regarding E_s. The two teams also share information regarding E_e, but work individually on technical and engineering problems arising in their own segments of the engineering environment E_i. Thus each project team in this type of organization has wide-ranging information about environments, partially shared and partially individuated. M's choice variable depends upon $\xi_s = \gamma_s + \epsilon_s$, while T_i's depends upon $\xi_s = \gamma_s + \epsilon_s$

(common to M and T_i's), $\xi_e = \gamma_e + \epsilon_e$ (common to T_i's), and $\xi_i = \gamma_i + \epsilon_i$ (idiosyncratic to T_i).

This type of organization may be considered as corresponding to what Stephen Klein conceptualized as the "chain-linked model" of innovation (Klein and Rosenberg, 1986; Aoki and Rosenberg, 1989). Information assimilation is realized through the feedback of information from the lower level to the higher level, as well as through information sharing and joint development effort across design project teams on the same level.

(3) V-mediated Information Encapsulation In this type of organization, information regarding E_s is shared among M and T_i's as in the interactive R&D organization. However, unlike the interactive R&D organization, there is no information sharing between T_i's regarding E_e. Thus development designs are completely encapsulated within each team and their new product design is based on individuated, differentiated knowledge. M's choice variable x depends upon $\xi_s = \gamma_s + \epsilon_s$, and T_i's upon $\xi_s = \gamma_s + \epsilon_s$ (common to M and T_i's), $\xi_{ei} = \gamma_e + \epsilon_{ei}$ (idiosyncratic to T_i's), and $\xi_i = \gamma_i + \epsilon_i$ (idiosyncratic to T_i's), where the same assumption as in the hierarchical R&D organization applies to ϵ_{ei}'s.

This model may be interpreted as an internal R&D organization, with each project team having high autonomy in information processing and product design. However, we regard this model as capturing some essential aspects of the relationship between venture capitalists and entrepreneurial firms, and that among entrepreneurial firms in Silicon Valley. According to this interpretation, M is a venture capitalist and T_i's are independent entrepreneurial firms. As we already noted, there is a substantial degree of information sharing among them about the emergent industrial systemic environment, and venture capitalists often take the role of intermediating such information by mediating contacts among entrepreneurs, engineers, university researchers etc.

3.2 Comparative Analysis of Information Efficiency

Since the objective function is quadratic and concave, the second-best decision rule for each agent is known to be linear in pieces of information available to and utilized by him/her (Marschak and Radner, 1972, Ch.5). In the course of calculating second-best decision rules, the coefficients appearing in them turn out to be proportional to the precision of information-processing activity. Here we adopt the following measure of precision of an observation according to the Bayesian theory of inference. Suppose that the prior variance of the observed environmental parameter is σ_j^2 and the variance of the observation error is $\sigma_{j\epsilon}^2$. Then the precision of observation is defined as $\Pi_j = \frac{\sigma_j^2}{\sigma_j^2+\sigma_{j\epsilon}^2}$.

For the purpose of comparison, suppose that the above three types of organizations face the same organizational environments. Namely random variables regarding E_s, E_e, E_i and all the constant parameters are the same across types of organizations in the following analyses. In addition, first suppose that the precision of processing information regarding those environments is equal across those types of organizations. Then tedious calculation shows the following.

Proposition 1 *Suppose that the three types of R&D organizations face the same stochastic parameters and constant parameters, and that, for each stochastic parameter, the precision of processing information is the same across those organizations. Then the V-mediated information encapsulation is informationally more efficient than hierarchical and interactive R&D organization if and only if $K > L$, namely when the choice variables of T_i's are not complementary. The interactive R&D organization is informationally more efficient than the hierarchical R&D organization.*[6]

Proof. See the Appendix.∎

The Intuition behind the above proposition is as follows. If the choice variables of design projects are complementary (namely the value function is supermodular in the decision variables), it is more profitable to coordinate them so that they move in the same direction (Milgrom and Roberts, 1995; Prat, 1996). Such a mechanism is internalized in the hierarchical and interactive R&D organizations since information is more assimilated in those types of organizations. In contrast, in the V-mediated information encapsulation, the observations of systemic engineering environments by entrepreneurial firms are mutually hidden, so that their decision choices are necessarily less correlated.

However, the above verbal description of information-processing activities in each type of organization suggests that the precision of processing information can be different across types of organizations. In the interactive R&D organizations, T_i's are collectively engaged in observation and communication of E_e, while in the V-mediated information encapsulation, E_e is observed separately together with E_i. Suppose therefore that the precision of processing information regarding E_i's is sacrificed relatively more often in the interactive R&D organization because attention is diverted to communications, although the precision regarding E_e may be improved because of pooling of data between the agents. The next proposition states that the V-mediated information encapsulation is informationally more efficient than the interactive and hierarchical R&D organizations when the environments surrounding T_1 and T_2 are statistically less correlated.[7]

Proposition 2 *Suppose that the three types of R&D organizations face the same constant parameters. Suppose that $\Pi_i^V > \Pi_i^I$ $(i = a, b)$, $\Pi_e^I > \Pi_e^V$ and $\Pi_s^V = \Pi_s^I$. If the systemic segment of the engineering environment is relatively unimportant (σ_{γ_e} is small) and the idiosyncratic engineering environment is relatively important (σ_i is large), then the V-mediated information encapsulation is informationally more efficient than the interactive and hierarchical R&D organizations.*

Proof. See the Appendix.∎

The above two propositions are instrumental to understanding the nature of the unique arrangement of R&D activities in the Silicon Valley model. Baldwin and Clark (2000) note that closely related to the unique informational arrangement of the Silicon Valley model is the concept of "modularization" of a product system. This concept involves at least three aspects: (1) partitioning a product system into relatively independent modules; (2) reduction in complementarity among the modules through standardization of interfaces among them; and (3) the unique mixture of information sharing and information encapsulation. We submit that the third aspect, which is the V-mediated information encapsulation in the present context, has a close relationship with the first and/or the second aspects.[8] Furthermore the second aspect enables the *ex post* formation of a new product system by combining new modular products. Let us elaborate these points in more detail.

First, modularization partitions a complex product system into several modules. A module is a unit of a system within which elements are strongly interrelated to one another, but across which they are relatively independent. In order to obtain this property of the system, the way in which partitioning is carried out cannot be arbitrary at all. Albeit in a different context, Cremer (1980) showed that the optimal way of partitioning an organization is the one that minimizes the statistical correlations among the units. In the present context, the whole design problem should be divided into two tasks in such a way that statistical correlation between the two are minimized. This means that the systemic engineering environment for each unit would become relatively unimportant as compared with idiosyncratic engineering environment. As Proposition 2 shows, V-mediated information encapsulation is viable in such environments. In this sense, "good" modularization, namely "good" architecture of a product system, is complementary to the unique informational arrangement observed in Silicon Valley.

Second, all the modules created through the process of partitioning as mentioned above have to be compatible with one another and work together in a smooth manner. In order to assure such compatibility, the interfaces among modules have to be explicitly and clearly determined. In other words, the interfaces have to be standardized. Under well-defined interfaces, R&D activities in respective modules can be conducted in parallel. This means reduction in technological complementarity between two tasks in our model. It may be the case in general that the choice variables of T_i's exhibit some complementarity; namely

$K < L$. However, by the standardization of interfaces, K and L become sufficiently close. Thus standardization of interfaces also makes V-mediated information encapsulation a viable organizational arrangement.

If K is sufficiently close to L, the value function is nearly separable, meaning that the improvement of the whole system results from that of each modular product, rather than from the coordinated and simultaneous improvements of several modular products. This sets the technological basis for a product system to be formed evolutionarily by combining new modular products. In order to talk about the *ex post* evolutionary formation of a product system, however, we have to see the situation where multiple entrepreneurial firms are present in each module and the standardized interfaces are made publicly open to them. Such a situation will be analyzed in the next section.

The above observation also helps us understand why most success stories in Silicon Valley are concentrated in the information and communications industries. The technological development in those industries has been spurred by setting standards for various interfaces arising in the information and communications systems. The modular design of the IBM System/360 is a notable example. Another example can be found in Internet/Web services. The Internet can be seen as a collection of protocols concerning the "platform layer," such as TCP/IP and HTML, that are independent of the "physical layer." This structure also enables various application software to be developed independently. Once good modular architecture is set, innovations usually take place in individual modules, and architecture and interfaces will change less frequently. In such an environment, complementarity between activities in different modular parts will be reduced, and the degree of uncertainty in the systemic segment of the engineering environment will be low. Thus V-mediated information encapsulation, which we think captures the essence of Silicon Valley model, would become effective.

4 The Incentive Aspect of the Silicon Valley Model

In Section 2, we argue that the other major role played by venture capitalists is that of governance. This section explores "governance by tournament" in Silicon Valley. The stylized VC tournament game shows that this aspect is deeply interconnected with the information-systemic aspect of the Silicon Valley model analyzed before.

4.1 Description of the VC Tournament Game

Assume that time consists of an infinite sequence of stage games. Each stage game is played over four dates between venture capitalists and entrepreneurial firms. The venture capitalists live permanently, competing with one another to nurture valuable firms, while entrepreneurial firms start up at the beginning of date 2 and exit at the end of date 4, either by going public, being acquired by other firms, or being terminated. When terminated, an entrepreneur can come back to the next stage game as a new candidate for a start-up firm. In the present paper, we do not explicitly explore the repeated nature of the game, but concentrate on the analysis of the single stage game between one venture capitalist and multiple start-up firms.[9]

Before going into the details of the stage game, let us roughly draw its whole picture here. Suppose that there are only two types of projects ($i = a, b$) present. A venture capitalist, referred to as VC henceforth, sets a limit to the number of start-up firms in each project and selects them by screening. Hereafter we use a "start-up firm" and its "entrepreneur" as interchangeable terms. Each of the selected entrepreneurs is then engaged in R&D activity that requires effort, which results in the observation of relevant environments. For each type of project, the VC holds a tournament game among entrepreneurs. Once the VC determines winners in the respective tournament games, and the VC and the winners choose their activity levels based upon their observations of the environments, the value of the whole product system will be

$$\begin{aligned} V(x, y_a, y_b) &= \gamma_s x + (\gamma_s + \gamma_a) y_a + (\gamma_s + \gamma_b) y_b \\ &\quad - \frac{A}{2} x^2 + D x (y_a + y_b) - \frac{K}{2} (y_a + y_b)^2 - \frac{L}{2} (y_a - y_b)^2 \end{aligned} \tag{2}$$

where x is the activity level chosen by the VC and y_i is that chosen by the winner in project type i. As in the previous section, γ_s and γ_i are stochastic parameters expressing the uncertainty in the systemic segment E_s and idiosyncratic engineering segment E_i of the environment. Different from Equation (1), in Equation (2) the systemic engineering environment is not present. This implies that E_e is relatively unimportant, i.e., its variance is low, where the Silicon Valley model is applied.[10] Although this value function is similar to that used in the previous section, the analysis in this section explicitly concerns the incentives of entrepreneurs.

A more detailed description of the stage game is as follows. At date 1, the VC chooses the number of start-up firms to fund in each project type, and screens many R&D projects proposed by cash-constrained, would-be entrepreneurs (*ex ante* monitoring). Let the number of selected entrepreneurs in project type i be denoted by n_i. The selected start-up firms are indexed by subscript ij, where i denotes the project type and $j (= 1, \cdots, n_i)$ indexes each firm in the same project type.

At date 2, each start-up firm funded by the VC is engaged in R&D activity that requires effort. The choice of effort level by start-up firm ij is denoted by e_{ij} and the associated cost by $c(e_{ij})$ that has the usual property of increasing marginal cost. The R&D effort of entrepreneur ij generates noisy one-dimensional information ξ_{ij} — research results – regarding E_i with the precision $\Pi_i(e_{ij})$. We assume that the higher the effort level, the higher the precision of the entrepreneur's posterior estimates regarding the environment that it faces; namely $\Pi_i(e_{ij})$ is increasing. The actual levels of effort exerted by the start-up firms may be inferred, but are not verifiable in the courts, so that they are not contractible. The fixed amount of funds provided to each entrepreneur in project i at this date is denoted by K_i, which covers only the cost of processing information at this date and is not sufficient for further product development.

At date 3, communication between the entrepreneur and the VC takes place. In this process the entrepreneurs and the VC mutually improve and assimilate their estimates of the systemic environment E_s, resulting in assimilated information ξ_s. Suppose that the precision of their assimilated information is an increasing function $\Pi_s(\cdot)$ of the VC's mediating effort level e_{VC}. The costs associated with the VC's mediating and monitoring efforts are represented by $\kappa(e_{VC})$, which has the usual property of increasing marginal costs.

At the beginning of date 4, the VC observes the value potentially created by entrepreneur ij with imprecision. Based on this observation, the VC estimates which combination of a product design from each type of project is expected to generate higher value, if the respective firms are offered to the public or acquired by an existing firm. According to this estimation, the VC selects one proposal from each type of project for implementation and allocates one unit of available funds to the winning entrepreneurs. The start-up firms that are not selected exit.

At the end of date 4, the selected projects are completed. Selected entrepreneurs and the VC choose their decisions based upon ξ_{ij} and ξ_s. Then all the environmental uncertainties are resolved. The VC offers the ownership of these firms to the public through markets or sells it to an acquiring firm. The realized value V is then distributed among the VC and the entrepreneurs. Suppose that the initial contract is such that at the time when winners are selected and the value is realized, a fixed share α_i is vested with the winning entrepreneur in project i and the unfunded entrepreneur forfeits any share. Let us denote the distributive share of the value to the VC by $\alpha_{VC} = 1 - \sum_i \alpha_i$. The payoff to the winning firm ij is then $\alpha_i V - c(e_{ij}) + K_i$ and that to the VC is $\alpha_{VC} V - \kappa(e_{VC}) - n_i K_i$.

Note that, at date 4, the VC and the entrepreneurs have the full incentives to choose their activity levels according to the second-best decision rules as derived in the previous section. This is because e_{ij} are already exerted at date 3, α_{VC} and α_i are fixed, and the expected payoffs to the winners in respective projects and that to the VC are increasing in V, which does not depend on any effort levels at this juncture.

4.2 Incentive Impacts of Governance by Tournament

Let us now analyze the stage game. At date 4, the VC and the winning entrepreneurs will coordinate their decisions according to the second-best decision rules. As was shown in the previous section, the second-best decision rules for the VC and the entrepreneurs turns out to be linear in the precision of processing information, $\Pi_s(e_{VC})$ and $\Pi_i(e_{ij})$. Specifically, the second-best decision rule of the VC is $x^* = \frac{D+K}{AK-D^2}\Pi_s(e_{VC})\xi_s$, and that of entrepreneur ij is $y^*_{ij} = \frac{D+A}{2(AK-D^2)}\Pi_s(e_{VC})\xi_s + \frac{1}{K+L}\Pi_i(e_{ij})\xi_{ij}$ in this mode of V-mediated information encapsulation. The resultant expected value if the VC selected entrepreneur j from project a and k from b would be

$$\frac{2D+K+A}{2(AK-D^2)}\sigma^2_{\gamma_s}\Pi_s(e_{VC}) + \frac{1}{2(K+L)}(\sigma_a^2\Pi_a(e_{aj}) + \sigma_b^2\Pi_b(e_{bk})) \tag{3}$$

Since this expected value is an additively separable function in the effort levels by the VC and the winning entrepreneurs in both types of projects, and the winning entrepreneurs receive a fixed share of the value, the incentive effect on the entrepreneurs of the VC tournament game can be examined by considering only the tournament game within each project, to which we now turn.

Since we can now restrict our attention to a tournament within a fixed project, we henceforth suppress subscript i. Suppose that n start-up firms are selected in this project at date 1. Let $e_j \geq 0$ be the effort level exerted by the j-th entrepreneur $(j = 1, \cdots, n)$ at date 2 and $c(e_j)$ be its associated cost function. In order to assure that a unique interior solution exists, we assume $c(e_j)$ is increasing and convex and $c'(0) = 0, c''(\infty) = \infty$. Let the share of the winning entrepreneur in this project type be $\alpha \in (0, 1)$. Consider the portion of Equation (3) that is affected by an entrepreneur's effort level. For entrepreneur j engaged in project i, this is $\frac{1}{2(K+L)}\sigma_i^2\Pi_i(e_{ij})$. Suppressing the subscript i and rewriting σ_i^2 as β, we rewrite the relevant expected value as $g(e_j, \beta)$. Note that β represents the degree of uncertainty involved in the R&D activity. Assuming that $\Pi_i(\cdot)$ is differentiable, $g(e_j, \beta)$ is such that $\partial g/\partial e_j > 0$, $\partial g/\partial \beta > 0$, and $\partial^2 g/\partial e_j \partial \beta > 0$ hold.

As an expert in estimating market values of firms, the VC evaluates the potential market value of each firm with some imprecision, and then selects the entrepreneur with the highest value as a winning entrepreneur. Suppose that the VC observes a random variable $y_j = g(e_j, \beta) + \epsilon_j^{VC}$ for each j with effort level e_j at the beginning of date 4. Here ϵ_j^{VC} comprises the VC's observation error with respect to $g(e_j, \beta)$, as well as the uncertainty in the potential market value of the entrepreneurial firm j that is different from and independent of the technological uncertainty to be resolved at the end of date 4. We assume that $\epsilon_j^{VC} \sim N(0, \sigma_{VC}^2)$ for all j and *i.i.d.* Thus, in the present model, a large value of σ_{VC} may be a result of either low precision of the VC's observation (as measured by $1/\sigma_{VC}^2$) or high marketing uncertainty. In what follows, we will resort to either interpretation depending on the context. The resultant expected value created in this project is $max_j\{g(e_j, \beta) + \epsilon_j^{VC}\}$. The winner receives share α of the realized value and the VC the share α_{VC} after all the technological uncertainty unfolds at the end of date 4.

First suppose that the VC has already chosen $n(\geq 2)$, and consider the game held at date 2 where entrepreneurs choose their level of R&D effort. Since the situation each entrepreneur faces is the same, we restrict our attention to the symmetric Nash equilibrium of this game. Let e^* be the equilibrium level of effort. Then the j-th entrepreneur's problem is described as choosing e_j to maximize

$$\alpha E[(g(e_j, \beta) + \epsilon_j^{VC}) Prob\{g(e_j, \beta) + \epsilon_j^{VC} > g(e^*, \beta) + \max_{k \neq j} \epsilon_k^{VC}\}] - c(e_j) \tag{4}$$

where the expectation is taken with respect to $\epsilon^{VC} = (\epsilon_1^{VC}, \cdots, \epsilon_n^{VC},)$ and $\max_{k \neq j} \epsilon_k^{VC}$ is the maximum order static of a sample of size $n-1$ (Galambos, 1984). Denoting the pdf and cdf of ϵ_k^{VC} by f and F, the pdf and cdf of the maximum order statistics of a sample of size $n-1$ are $(n-1)f(x)F(x)^{n-2}$ and $F(x)^{n-1}$ respectively. So rewriting (4) yields

$$\alpha \int_{-\infty}^{\infty} (g(e_j, \beta) + x) F(g(e_j, \beta) - g(e^*, \beta) + x)^{n-1} f(x) dx - c(e_j) \tag{5}$$

By differentiating (5) and letting $e_j = e^*$, the symmetric Nash equilibrium condition becomes

$$\alpha\frac{\partial g(e^*,\beta)}{\partial e_j}[\frac{1}{n}+g(e^*,\beta)\int_{-\infty}^{\infty}(n-1)f(x)^2F(x)^{n-2}dx$$
$$+\int_{-\infty}^{\infty}x(n-1)f(x)^2F(x)^{n-2}dx]=c'(e^*) \quad (6)$$

The first term in the parentheses on the left-hand side is the probability of winning, which turns out to be $1/n$. The second term is the expected payoff times the marginal increase in the probability of winning. The third term is the marginal expected value resulting from the marketing uncertainty. The next proposition should be intuitively obvious.

Proposition 3 *Consider the symmetric Nash equilibrium of the subgame in which entrepreneurs choose their effort levels. The equilibrium level of effort is strictly decreasing in the number of selected entrepreneurs ($n \geq 2$), strictly decreasing in the variance of the VC's observation error and/or the marketing uncertainty, and strictly increasing in the uncertainty involved in R&D activity.*

Proof. Since $c'' > 0$, for the first and second parts, it suffices to show that the expression enclosed by the parentheses on the left-hand side of the Nash equilibrium condition is strictly decreasing in n and σ_{VC}^2. Suppose $n \geq 2$. By Lemma 2 in the appendix, the coefficient of the second term is decreasing in n. By Lemma 3 in the appendix, the sum of the first and the third terms is shown to be strictly decreasing in n. Thus the expression in the parenthesis on the left-hand side is strictly decreasing in n.
Let us denote the pdf and cdf of the standard normal distribution by ϕ and Φ respectively. Then the coefficient of the second term is expressed as

$$\frac{(n-1)\int_{-\infty}^{\infty}\phi(x)^2\Phi(x)^{n-2}dx}{\sigma_{VC}}$$

which is strictly decreasing in σ_{VC}. The third term can be rewritten as

$$\int_{-\infty}^{\infty}x(n-1)\phi(x)^2\Phi(x)^{n-2}dx$$

Thus the expression in the parenthesis on the left-hand side is decreasing in σ_{VC}.
Finally observe that the left-hand side of the Nash equilibrium condition is obviously strictly increasing in β. ∎

Proposition 3 shows that increasing the number of entrepreneurs engaged in the same modular component lowers their incentives to make efforts in the symmetric equilibrium. It also shows that higher precision of the VC's observation and/or the lower marketing uncertainty as well as higher R&D uncertainty induce higher effort levels from the tournament participants. Also note that when there is only one entrepreneur in the project, the effort level he/she chooses is determined by $\alpha \frac{\partial g(e,\beta)}{\partial e} = c'(e)$. Keeping the other parameters constant, we see that the VC tournament can elicit higher efforts from entrepreneurs if and only if $\frac{g(e,\beta)\int_{-\infty}^{\infty}\phi(x)^2\Phi(x)^{n-2}dx}{\sigma_{VC}} + \int_{-\infty}^{\infty} x\phi(x)^2\Phi(x)^{n-2}dx > \frac{1}{n}$. This holds when σ_{VC} is small and $g(e,\beta)$ is large. Thus a large prize for the winner, the high precision of the VC's monitoring, and/or a low level of marketing uncertainty are essential for this tournament scheme to work well.

Now we are in a position to turn to the VC's problem of choosing the optimal number of tournament participants. The expected payoff of this project to the VC when he/she chooses n (we omit additional financing cost at date 4), denoted $\pi_{VC}(n, \sigma, K)$, is

$$\pi_{VC}(n, \sigma_{VC}, K) = \alpha_{VC}[g(e^*(n, \sigma_{VC}, \beta), \beta) + \int_{-\infty}^{\infty} nxf(x)F(x)^{n-1}dx] - nK \tag{7}$$

where e^* is the equilibrium level of effort, and K is the cost of start-up financing. Now consider the VC's problem of maximizing (7) over the set of integers with $n \geq 2$. $g(e^*(n, \sigma_{VC}, \beta), \beta)$ is strictly decreasing in n, because Proposition 3 shows e^* is strictly decreasing in n and $g(\cdot, \beta)$ is strictly increasing in the first argument. The second term in the parentheses is the effect of running n experimentations in parallel, which is mathematically the expected value of the maximum order statistic of a sample of size n and turns out to be strictly increasing in $n \geq 1$ by Lemma 1 in the Appendix. Thus we have the next proposition.

Proposition 4 *Consider the VC's problem of maximizing (7) over the set of integers with $n \geq 2$. The set of the solution is nonempty. Furthermore, if $K' < K''$, $n' \in \arg\max_{n:n\in\mathbb{N}, n\geq 2} \pi_{VC}(n, \sigma_{VC}, K')$ and $n'' \in \arg\max_{n:n\in\mathbb{N}, n\geq 2} \pi_{VC}(n, \sigma_{VC}, K'')$, then $n'' \leq n'$.*

Proof. Let $W(n) = U(n-1) = \int_{-\infty}^{\infty} nxf(x)F(x)^{n-1}dx$. By Lemma 1, $W(n)$ is strictly increasing in $n \geq 1$, and $W(n+1) - W(n)$ is strictly decreasing in $n \geq 1$. Also observe that $\lim_{n\to\infty} W(n) - W(n-1) = 0$. Thus the maximum of $\alpha_{VC}W(n) - nK$ is attained. Since $\alpha_{VC}g(e^*(n, \sigma_{VC}, \beta), \beta)$ is strictly decreasing in n, the relevant domain for the maximization problem is obviously bounded above, and thus finite. This proves the existence of the solution.

It is easy to see that the objective function has strictly increasing differences in $(n, -K)$. By the well-know theorem of monotone comparative statics (Topkis, 1998, p.79), $K' < K''$, $n' \in \arg\max_{n:n\in\mathbb{N},n\geq 2} \pi_{VC}(n, \sigma_{VC}, K')$ and $n'' \in \arg\max_{n:n\in\mathbb{N},n\geq 2} \pi_{VC}(n, \sigma_{VC}, K'')$ imply that $n'' \leq n'$. ∎

Proposition 4 shows that increasing the number of tournament participants will induce increased benefits from parallel experiments but it is at the cost of decreased incentives for tournament participants. The model developed in this section can be regarded as an extension of the VC tournament model in Aoki (2001), as well as of the model of the "substitution operator" by Baldwin and Clark (2000). In the former, the number of entrepreneurs in each R&D project is fixed at two, while the latter model abstracts from the effects of increasing the number of competitors on the entrepreneurs' incentives.[11] The results shown above suggest that incentive consideration can limit the effectiveness of the substitution operator.

We can derive some interesting implications from the above results concerning the dot.com bubble in which most entrepreneurial firms were engaged in so-called e-commerce businesses. Although the dot.com bubble and crash might have been caused primarily by erroneous expectations regarding profitability, we can say that the number of entrants into the Internet/Web services was very large because their start-up costs were low. Proposition 3 in turn suggests that the incentives of entrepreneurs might have been adversely affected by such a large number of entrants. As stated in Section 1, the technology involved in those businesses was not strikingly innovative, and only new business models had to be contrived. Thus it can be inferred that most e-commerce businesses had low technological uncertainty as well as high marketing uncertainty, which might also have affected the entrepreneurs' incentives adversely.

5 Conclusion

In this paper, we argue that a novel institutional arrangement for product system innovation has emerged in Silicon Valley. We have tried to capture the innovative nature of the "Silicon Valley model." We suggest that it is crucial to take a look at multifaceted relationships between venture capitalists, on the one hand, and a cluster of entrepreneurial firms, on the other. Our focus is on the information structural relationship and the governance relationships between them.

According to the analysis of the comparative R&D organizations, application of the Silicon Valley model may be limited to domains in which a product system design can be partitioned into modular products by standardized interfaces, thus the technological complementarity between them is reduced. On the other hand, the analysis of a governance relationship between a venture capitalist and entrepreneurial firms suggests that the Silicon Valley model may be effective when successful developmental projects are expected to yield extremely high values in markets, where there are venture capitalists who are capable of monitoring, and where the uncertainty in markets is low and technological uncertainty is high. It also shows that an increase in the number of entrants affects the incentives of entrepreneurs adversely and that the optimal number of entrants is decreasing in the amount required for start-up financing.

The identification of conditions for the informational efficiency of information encapsulation may have broader implications for corporate organizations in general. Because of the development of communications and transportation technology, even mature products (e.g. automobiles) are increasingly decomposed into modules, for which production and procurement become less integrated in comparison to traditional hierarchical firms (as represented by traditional American firms of a decade ago) or interactive firms (as represented by Japanese firms). This tendency renders compact modular organizations (either in the form of independent firms or subsidiaries) increasingly more efficient and viable. Various innovations in corporate governance appear to be evolving even in existing firms, somewhat emulating the Silicon Valley model, such as governing subsidiaries with flexible coupling and decoupling and less operational intervention, but with tournament-like financial discipline, which will be the subject of another paper.

Appendix

We first present the second-best decision rules and expected payoff for each type of R&D organization. The second-best decision rules are linear in the pieces of information available as is shown by Marschak and Radner (1972, p.168). This enables us, say in the case of hierarchical R&D organizations, to let $x = \lambda_s \xi_s$, $y_i = \lambda_{si}(\xi_s + \epsilon_{si}) + \lambda_{ei}(\xi_e + \epsilon_{ei}) + \lambda_i \xi_i$ $(i = a, b)$ and then to solve for the coefficients, $\lambda_s, \lambda_{si}, \lambda_{ei}$, and λ_i that together maximize the expected payoff. The derivation method is the same across the following types of organizations.

Hierarchical R&D organization In this case, the second-best decision rules turn out to be

$$x = \frac{K + D\Pi_{se}}{AK - D^2\Pi_{se}} \Pi_s^H \xi_s$$
$$y_i = \frac{D + A}{2(A\frac{K}{\Pi_{se}} - D^2)} \Pi_s^H (\xi_s + \epsilon_{si}) + \frac{1}{2K} \Pi_e^H (\xi_e + \epsilon_{ei}) + \frac{1}{K + L} \Pi_i^H \xi_i \quad (i = a, b)$$

where

$$\Pi_s^H = \frac{\sigma_{\gamma_s}^2}{\sigma_{\gamma_s}^2 + \sigma_{\epsilon_s}^2}$$
$$\Pi_{se} = \frac{\sigma_{\gamma_s}^2 + \sigma_{\epsilon_s}^2}{\sigma_{\gamma_s}^2 + \sigma_{\epsilon_s}^2 + \sigma_{se}^2}$$
$$\Pi_e^H = \frac{\sigma_{\gamma_e}^2}{\sigma_{\gamma_e}^2 + \sigma_{\epsilon_e}^2 + \sigma_{ee}^2}$$
$$\Pi_i^H = \frac{\sigma_i^2}{\sigma_i^2 + \sigma_{\epsilon_i}^2} \quad (i = a, b)$$

By substitution, the maximized expected payoff is calculated as

$$\frac{2D + \frac{K}{\Pi_{se}} + A}{2(A\frac{K}{\Pi_{se}} - D^2)} \sigma_{\gamma_s}^2 \Pi_s^H + \frac{1}{2K} \sigma_{\gamma_e}^2 \Pi_e^H + \frac{1}{2(K + L)} (\sigma_a^2 \Pi_a^H + \sigma_b^2 \Pi_b^H)$$

Interactive R&D organization For this case, the second-best decision rules are

$$x = \frac{D + K}{AK - D^2} \Pi_s^I \xi_s$$
$$y_i = \frac{D + A}{2(AK - D^2)} \Pi_s^I \xi_s + \frac{1}{2K} \Pi_e^I \xi_e + \frac{1}{K + L} \Pi_i^I \xi_i \quad (i = a, b)$$

where

$$\Pi_s^I = \frac{\sigma_{\gamma_s}^2}{\sigma_{\gamma_s}^2 + \sigma_{\epsilon_s}^2}$$
$$\Pi_e^I = \frac{\sigma_{\gamma_e}^2}{\sigma_{\gamma_e}^2 + \sigma_{\epsilon_e}^2}$$
$$\Pi_i^H = \frac{\sigma_i^2}{\sigma_i^2 + \sigma_{\epsilon_i}^2} \quad (i = a, b)$$

By substitution, the maximized expected payoff is

$$\frac{2D + K + A}{2(AK - D^2)}\sigma_{\gamma_s}^2 \Pi_s^I + \frac{1}{2K}\sigma_{\gamma_e}^2 \Pi_e^I + \frac{1}{2(K+L)}(\sigma_a^2 \Pi_a^I + \sigma_b^2 \Pi_b^I)$$

V-mediated information encapsulation Here the second-best decision rules are

$$x = \frac{D + K}{AK - D^2}\Pi_s^V \xi_s$$
$$y_i = \frac{D + A}{2(AK - D^2)}\Pi_s^V \xi_s + \frac{1}{(K - L)\Pi_e^V + (K + L)}\Pi_e^V(\xi_e + \epsilon_{ei}) + \frac{1}{K + L}\Pi_i^V \xi_i$$

where

$$\Pi_s^V = \frac{\sigma_{\gamma_s}^2}{\sigma_{\gamma_s}^2 + \sigma_{\epsilon_s}^2}$$
$$\Pi_e^V = \frac{\sigma_{\gamma_e}^2}{\sigma_{\gamma_e}^2 + \sigma_{ee}^2} \quad (i = a, b)$$
$$\Pi_i^V = \frac{\sigma_i^2}{\sigma_i^2 + \sigma_{\epsilon_i}^2} \quad (i = a, b)$$

By substitution, the maximized expected payoff is

$$\frac{2D + K + A}{2(AK - D^2)}\sigma_{\gamma_s}^2 \Pi_s^V + \frac{1}{(K - L)\Pi_e^V + (K + L)}\sigma_{\gamma_e}^2 \Pi_e^V + \frac{1}{2(K+L)}(\sigma_a^2 \Pi_a^V + \sigma_b^2 \Pi_b^V)$$

Proof of Proposition 1 By assumption, σ_{γ_s}, σ_{γ_e}, $\sigma_i (i = a, b)$, and constant parameters are all equal across types of R&D organizations and $\Pi_s^H = \Pi_s^I = \Pi_s^V$, $\Pi_e^H = \Pi_e^I = \Pi_e^V$, $\Pi_i^H = \Pi_i^I = \Pi_i^V$. First observe that the only difference in the maximized expected payoff between hierarchical R&D organization and Interactive Organization lies in the coefficient of $\Pi_s^t (t = H, I)$. Since $\frac{2D+K+A}{2(AK-D^2)}$ is decreasing in K and Π_{se} is less than 1, the coefficient in the maximized expected payoff for the hierarchical R&D organization is less than that for the interactive R&D organization. This establishes the second half of the statement of Proposition 1. Now it suffices to make a comparison between interactive R&D organization and V-mediated information encapsulation.

In this comparison, the only difference lies in the coefficient of $\Pi_e^T (T = I, V)$.

$$\frac{1}{(K-L)\Pi_e^V + (K+L)} > \frac{1}{2K}$$

if and only if

$$K - L > 0$$

This completes the proof.∎

Proof of Proposition 2 Letting the maximized expected payoff for V-mediated information encapsulation be greater than that for interactive R&D organization,

$$\frac{1}{2(K+L)}\left[\sigma_a^2(\Pi_a^V - \Pi_a^I) + \sigma_b^2(\Pi_b^V - \Pi_b^I)\right] > \sigma_{\gamma_e}^2\left[\frac{\Pi_e^I}{2K} - \frac{\Pi_e^V}{(K-L)\Pi_e^V + (K+L)}\right]$$

Since $\Pi_i^V > \Pi_i^I$ for $i = a, b$, $\Pi_e^I > \Pi_e^V$, the above inequality holds for sufficiently large σ_i's and sufficiently small σ_{γ_e}.∎

The following lemmata are used in the proofs of Proposition 3 and 4 in Section 4. In this appendix, let $f(x)$ and $F(x)$ be the generic pdf and cdf of a normal distribution $N(0, \sigma^2)$ respectively.

Lemma 1 *Let n be a nonnegative integer. The expected value of the maximum order statistic of a sample of size $n+1$, denoted as $U(n) = \int_{-\infty}^{\infty} x(n+1)f(x)F(x)^n dx$, is strictly increasing in n. Furthermore $U(n+1) - U(n)$ is strictly decreasing in n.*

Proof. Suppose $n \geq 0$.

$$
\begin{aligned}
\frac{U(n+1)}{n+2} &= \int_{-\infty}^{\infty} x f(x) F(x)^{n+1} dx \\
&= \left[(F(x)-1) x F(x)^{n+1}\right]_{-\infty}^{\infty} \\
&\quad - \int_{-\infty}^{\infty} (F(x)-1)\left(F(x)^{n+1} + x(n+1) f(x) F(x)^{n}\right) dx \\
&= \int_{-\infty}^{\infty} (1-F(x)) F(x)^{n+1} dx + \int_{-\infty}^{\infty} (1-F(x)) x(n+1) f(x) F(x)^{n} dx \\
&= \int_{-\infty}^{\infty} (1-F(x)) F(x)^{n+1} dx + \int_{-\infty}^{\infty} x(n+1) f(x) F(x)^{n} dx \\
&\quad - \int_{-\infty}^{\infty} x(n+1) f(x) F(x)^{n+1} dx \\
&= \int_{-\infty}^{\infty} (1-F(x)) F(x)^{n+1} dx + U(n) - \frac{(n+1)U(n+1)}{n+2}
\end{aligned}
$$

Thus

$$
U(n+1) - U(n) = \int_{-\infty}^{\infty} (1-F(x)) F(x)^{n+1} dx > 0
$$

Therefore $U(n)$ is strictly increasing in n, $U(n+1) - U(n)$ is strictly decreasing in n.∎

Lemma 2 *Let n be an integer with $n \geq 2$ and $T(n) = \int_{-\infty}^{\infty} (n-1) f(x)^2 F(x)^{n-2} dx$. Then $T(n)$ is decreasing in n for $n \geq 2$, and strictly decreasing in $n \geq 3$.*

Proof. For $n \geq 2$,

$$
\begin{aligned}
\frac{T(n+1)}{n} &= \int_{-\infty}^{\infty} f(x)^2 F(x)^{n-1} dx \\
&= \left[(F(x)-1) f(x) F(x)^{n-1}\right]_{-\infty}^{\infty} \\
&\quad - \int_{-\infty}^{\infty} (F(x)-1)\big(f'(x) F(x)^{n-1} + (n-1) f(x)^2 F(x)^{n-2}\big) dx \\
&= \int_{-\infty}^{\infty} (1-F(x)) f'(x) F(x)^{n-1} dx + \int_{-\infty}^{\infty} (1-F(x))(n-1) f(x)^2 F(x)^{n-2} dx \\
&= \int_{-\infty}^{\infty} (1-F(x)) f'(x) F(x)^{n-1} dx + \int_{-\infty}^{\infty} (n-1) f(x)^2 F(x)^{n-2} dx \\
&\quad - \int_{-\infty}^{\infty} (n-1) f(x)^2 F(x)^{n-1} dx \\
&= \int_{-\infty}^{\infty} (1-F(x)) f'(x) F(x)^{n-1} dx + T(n) - \frac{(n-1)T(n+1)}{n}
\end{aligned}
$$

Thus

$$\begin{aligned}
T(n+1)-T(n) &= \int_{-\infty}^{\infty}(1-F(x))f'(x)F(x)^{n-1}dx \\
&= \int_{-\infty}^{\infty} f'(x)F(x)^{n-1}dx - \int_{-\infty}^{\infty} f'(x)F(x)^{n}dx \\
&= \int_{-\infty}^{\infty} -\frac{x}{\sigma^2}f(x)F(x)^{n-1}dx - \int_{-\infty}^{\infty} -\frac{x}{\sigma^2}f(x)F(x)^{n}dx \\
&= \frac{1}{\sigma^2}\left[\int_{-\infty}^{\infty} xf(x)F(x)^{n}dx - \int_{-\infty}^{\infty} xf(x)F(x)^{n-1}dx\right] \\
&= \frac{1}{\sigma^2}\left[\frac{U(n)}{n+1} - \frac{U(n-1)}{n}\right] \\
&= \frac{1}{\sigma^2}\left[\frac{W(n+1)}{n+1} - \frac{W(n)}{n}\right]
\end{aligned}$$

where the third equality follows from $f'(x) = -(x^2/\sigma^2)f(x)$ and $W(n) = U(n-1)$. Now it suffices to show $\frac{W(3)}{3} = \frac{W(2)}{2}$, and $\frac{W(n)}{n}$ is strictly decreasing in $n \geq 3$. For the first part,

$$\begin{aligned}
\frac{W(2)}{2} - \frac{W(3)}{3} &= \frac{U(1)}{2} - \frac{U(2)}{3} \\
&= \int_{-\infty}^{\infty} xf(x)(1-F(x))F(x)dx \\
&= \int_{-\infty}^{\infty} xf(x)(\frac{1}{4} - G(x)^2)dx \\
&= -\int_{-\infty}^{\infty} xf(x)G(x)^2dx \\
&= 0
\end{aligned}$$

where $G(x) = F(x) - \frac{1}{2}$ is an odd function, and the fifth equality follows because the integrand is also an odd function.

For the second part, the proof proceeds by induction. First observe that $W(2)/2 = W(3)/3 > W(4)/4$, because $W(4) - W(3) < W(3) - W(2)$. Suppose $W(2)/2 = W(3)/3 > \cdots > W(k-1)/(k-1) > W(k)/k$, which implies

$W(k) - W(k-1) < W(k)/k$.

$$\begin{aligned}\frac{W(k+1)}{k+1} &= \frac{1}{k+1}(W(k+1) - W(k)) + \frac{k}{k+1}\frac{W(k)}{k} \\ &< \frac{1}{k+1}(W(k) - W(k-1)) + \frac{k}{k+1}\frac{W(k)}{k} \\ &< \frac{W(k)}{k}\end{aligned}$$

where the first inequality follows from $W(k+1) - W(k) < W(k) - W(k-1)$ and the second inequality follows from $W(k) - W(k-1) < W(k)/k$. Thus $\frac{W(n)}{n}$ is strictly decreasing in n for $n \geq 3$.∎

Lemma 3 *For $n \geq 2$, let $S(n) = 1/n + \int_{-\infty}^{\infty} x(n-1)f(x)^2F(x)^{n-2}dx$. Then $S(n)$ is strictly decreasing in n.*

Proof. First observe that

$$\int_{-\infty}^{\infty}(n-1)f(x)(1-F(x))F(x)^{n-2}dx = \int_0^1 (n-1)(1-y)y^{n-2}dy = \frac{1}{n}$$

By substitution

$$\begin{aligned}S(n) &= \int_{-\infty}^{\infty}(n-1)f(x)(1-F(x))F(x)^{n-2}dx + \int_{-\infty}^{\infty}(n-1)xf(x)^2F(x)^{n-2}dx \\ &= (n-1)\int_{-\infty}^{\infty} f(x)(1-F(x)+xf(x))F(x)^{n-2}dx\end{aligned}$$

For $n \geq 2$,

$$
\begin{aligned}
\frac{S(n+1)}{n} &= \int_{-\infty}^{\infty} f(x)(1 - F(x) + xf(x))F(x)^{n-1}dx \\
&= \int_{-\infty}^{\infty} [F(x) - 1]'(1 - F(x) + xf(x))F(x)^{n-1}dx \\
&= \left[(F(x) - 1)(1 - F(x) + xf(x))F(x)^{n-1}\right]_{-\infty}^{\infty} \\
&\quad - \int_{-\infty}^{\infty} (F(x) - 1)[(n-1)f(x)(1 - F(x) + xf(x))F(x)^{n-2} + xf'(x)F(x)^{n-1}]dx \\
&= \int_{-\infty}^{\infty} (1 - F(x))(n-1)f(x)(1 - F(x) + xf(x))F(x)^{n-2}dx \\
&\quad + \int_{-\infty}^{\infty} (1 - F(x))xf'(x)F(x)^{n-1}dx \\
&= \int_{-\infty}^{\infty} (n-1)f(x)(1 - F(x) + xf(x))F(x)^{n-2}dx \\
&\quad - \int_{-\infty}^{\infty} (n-1)F(x)^{n-1}(1 - F(x) + xf(x))f(x)dx \\
&\quad + \int_{-\infty}^{\infty} (1 - F(x))xf'(x)F(x)^{n-1}dx \\
&= S(n) - \frac{n-1}{n}S(n+1) + \int_{-\infty}^{\infty} (1 - F(x))xf'(x)F(x)^{n-1}dx
\end{aligned}
$$

Thus

$$
\begin{aligned}
S(n+1) - S(n) &= \int_{-\infty}^{\infty} (1 - F(x))xf'(x)F(x)^{n-1}dx \\
&= -\frac{1}{\sigma^2} \int_{-\infty}^{\infty} (1 - F(x))x^2 f(x)F(x)^{n-1}dx \\
&< 0
\end{aligned}
$$

which completes the proof. ∎

References

Aoki, Masahiko. *Towards A Comparative Institutional Analysis.* Cambridge, MA: MIT Press, 2001.

Aoki, Masahiko, and Rosenberg, Nathan, "The Japanese Firm as an Innovating Institution." in Takashi Shiraishi and Shigeto Tsuru, Eds., *Economic Institutions in a Dynamic Society,* pp.137-154. New York: Macmillan, 1989.

Baldwin, Carliss Y., and Clark, Kim B. *Design Rules - vol.1: The Power of Modularity.* Cambridge, MA: MIT Press, 2000.

Baldwin, Carliss Y., and Clark, Kim B, "Modularity after the Crash." in Raghu Garud, Arun Kumaraswamy, and Richard Langlois, Eds., *Managing the Modular Age: Architectures, Networks and Organizations,* Oxford, UK: Blackwell Publishers, 2001 (forthcoming).

Cremer, Jacques, "A Partial Theory of Optimal Organization of a Bureaucracy." *Bell Journal of Economics.* **11**, 683-693, 1980.

Cremer, Jacques, "Common Knowledge and the Co-ordination of Economic Activities." in Masahiko Aoki, and Bo Gustafsson, and Oliver E. Williamson, Eds., *The Firm as a Nexus of Treaties*, pp.53-76. London: Sage Publications, 1990.

Galambos, Janos, "Order Statistics." in P.R. Krishnaiah and P.K. Sen, Eds., *Handbook of Statistics 4: Nonparametric Methods*, pp.359-382, Amsterdam: North-Holland, 1984.

Gompers, Paul A., and Lerner, Josh "The Use of Covenants: An Empirical Analysis of Venture Partnership Agreements." *J. Law Econ.* **39**, 2:463-98, October 1996.

Grossman, Sanford J., and Hart, Oliver, "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *J. Polit. Economy* **94**, 691-719, 1986.

Hart, Oliver, and Moore, John, "Property Rights and the Nature of the Firm." *J. Polit. Economy* **98**, 1119-1158, 1990.

Hellmann, Thomas F., "Symantec Corporation: Acquiring Entrepreneurial Companies," Teaching Note S-SM-27. Stanford: Stanford Business School. Nov. 1998.

Kaplan, Steven N., and Strömberg, Per, "Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts," Working Paper 7660. Cambridge: National Bureau of Economic Research. Apr. 2000.

Klein, Stephen, and Rosenberg, Nathan, "An Overview of Innovation." in Ralph Landau, and Nathan Rosenberg, Eds., *The Positive Sum Strategy: Harnessing Technology for Economic Growth*, pp.275-306. Washington, D.C.: National Academic Press, 1986.

Marschak, Jacob, and Radner, Roy. *Economic Theory of Teams.* New Haven, CT: Yale University Press, 1972.

Milgrom, Paul, and Roberts, John, "Complementarities and Fit: Strategy, Structure, and Organizational Change in Manufacturing." *J. Acc. Econ.* **19**, 179-208, 1995.

Prat, Andrea, "Shared Knowledge vs Diversified Knowledge in Teams." *J. Japanese Int. Economies* **10**, 181-195, 1996.

Rajan, Raghuram G., and Zingales, Luigi, "Power in a Theory of the Firm." *Quart. J. Econ.* **113**, 2:387-432, 1998.

Sahlman, William A., "The Structure and Governance of Venture Capitalist Organizations." *J. Finan. Econ.* **27**, 473-521, 1990.

Saxenian, Annalee. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128.* Cambridge, MA: Harvard University Press, 1994.

Topkis, Donald M. *Supermodularity and Complementarity.* Princeton, NJ: Princeton University Press, 1998.

Footnotes

1. Acknowledgement

2. Rajan and Zingales (1998) is an attempt to generalize the basic model of the property rights approach. They point out that the original property rights models put too great and exclusive an emphasis on the ownership of physical assets as a source of power. They assert that power can come from allocation of *access* to various kinds of critical resources, such as specialized machines, good ideas, and talented people.

3. Similar argument is found in Baldwin and Clark (2000), which uses "DSM(design structure matrix)" and "TSM(task structure matrix)" to describe dependencies among design parameters and design tasks respectively.

4. This payoff function may be thought of as a second-order Taylor series approximation of a general payoff function around the optimal values of x and y_i's with respect to the prior distribution of the stochastic parameters. We also normalize the payoff so that the expected payoff is zero when there is no *ex post* information other than the priors.

5. Namely its Hessian matrix is negative definite.

6. This proposition can be seen as an extension of a theorem due to Cremer (1990). In the hierarchical R&D organization, the communication is one-directional and thus involves communication errors, while in the interactive R&D organization information is completely shared. This is the reason why the interactive R&D organization is informationally more efficient than the hierarchical R&D organization. Considering the cost saved by one-directional communication would change the result. However we will not be concerned about the comparison between the hierarchical R&D organization and the interactive R&D organization hereafter.

7. For more detailed comparative statics results in the same framework, see Aoki (2001, Ch.4 and 14). As the crux of the present paper lies in the analyses of the VC tournament game in the next section, the results stated in Proposition 1 and 2 are sufficient for our purposes.

8. Baldwin and Clark (2000) regard "modularization-in-design" as rationalization in the process of designing a complex product system. When they demonstrate how to modularize a product design by using a Design Structure Matrix, modularization is primarily to contrive an ideal hierarchical information system within the whole design process. Once this is done, or at the same time this is done, other aspects of modularization, such as reduced complementarity between different design tasks, information encapsulation etc., are supposed to come together immediately. In this sense, our approach is more analytical. It may also be said that we are deriving a second-best organizational arrangement with technological parameters given. Such a difference in the approach may make a somewhat subtle difference between our argument and theirs. According to our analysis, the practice of V-mediated information encapsulation is not realizable if there is indispensable complementarity between project teams or a systemic environment is necessarily very important. Some sorts of product system may not be modularized because of such difficulties.

9. The repeated nature of the game mainly concerns the incentive of venture capitalists. See Aoki (2001, Ch.14.3) for the impact of repeatedness on venture capitalists' incentives.

10. See Proposition 2 and the subsequent argument.

11. In the model of Baldwin and Clark (2000), the result of R&D activity in the current period is adopted if it turns out to be superior to the old one. Namely they regard the result of R&D activities in modular designs as "real options." They suggest that the greater the number of parallel experiments, the greater the value of real options,

which they call the "value of substitution." Although our model does not explicitly model the value of the VC tournament as real options, the same increasing property can be obtained. Namely increasing the number of entrepreneurs has the effect of mounting more parallel experiments.