Chapter 5

The Co-Evolution of Organizational Conventions and Human Asset Types

> One might ask, as one does frequently in the theory of the firm, why all firms do not have the same codes, so that training in the code is transferable? In the first place, in this combinatorial situation, there may easily be many optimal codes, all equally good, but to be useful in a firm it is important to know the right code. The situation here is very much that of the games of coordination which have been stressed so much by Schelling. If it is valuable for two people to meet without being able to communicate with each other during their trips, the meeting place must be agreed on before hand. It may not matter much where the meeting is to be. But a person who learned one meeting-place is not much use to an organization which has selected another.
>
> – Kenneth J. Arrow, 1974 *

In the last chapter we identified different types of organizational and quasi-organizational architecture. But, if there is a variety in organizational and quasi-organizational architecture, how is a selection of architectural structure is made from the many possibilities? Do firms adopt the best architecture only in response to evolving technological and market conditions? Recent theories of the firm seem to imply that an answer to this question could be, or at least ought to be, affirmative.[1] Technical details leading to these predictions need not concern us now, but the logical implications of technological determinism common to all of them do: organizational architecture ought to converge among firms in the same industry facing similar technological and market parameters. However, we often observe firms with different architectural and governance characteristics, as well as different degrees of integration, in the same industry across regions and economies. Why could this be the case?

There are undoubtedly elements of rational design in any particular organizational architecture, but there may also be an element of convention due to the "bounded-rationality" of agents and accordingly of firms. Once organizational conventions develop, there can be a tension between structural inertia and competitive pressure from the changing environment. Entrepreneurs try to experiment with new organizational design or emulate an organizational practice that evolved elsewhere with perceived superiority. However, the usual outcome of such experiments and emulations, even if they occur in a critical mass, is often neither a dramatic switch from one convention to another, nor a "chaotic" cohabitation of widely divergent organizational architecture. Rather, they are likely to result in a "modification," or a "ramification" of conventional organizational architecture that may significantly alter some characteristics of the existing conventions, yet retain other basic features. Alternatively, the emergence of new practices may occur as a clustering of new firms on the periphery of traditional industrial centers, as in the case of Silicon Valley.[2] Thus, as we observed in the previous chapter, references are often made to national or regional forms of organization, e.g., Silicon Valley firms, the "American system of manufacture" and the Fordist mode of production, the German participatory firms, the Japanese *keiretsu*, Italian industrial districts, etc. Some argue that inter-regional/inter-national differences in organizational architecture and implied organizational competence may explain the patterns of regional/national advantage in industry and trade.[3].

However, unlike the naturally endowed resources that constitute the source of Ricardian comparative advantage, the organization is a human contrivance. As such, shouldn't it be ultimately transplantable (mobile) across national economies? Or, even if the unbounded-rational entrepreneur cannot implement the optimal design of organizational architecture, will competitive selection eventually weed out inefficient organizational architecture which do not fit the technological imperatives of each industry? Therefore, as Alchian (1950) and M. Friedman (1953) argued some time ago, won't the competitive outcome look as if the optimal design is implemented? Especially, doesn't the growth of transnational firms and globalization of financial markets finally pave the way for the convergence of organizational architecture across economies toward industrial technological imperatives? These are the questions that we plan to deal with throughout this book, and this chapter begins our journey. For that purpose, we envision a relatively unstructured domain of the game, the organizational field, in which organizational architecture of firms may be simultaneously chosen.[4] We will construct a model of evolutionary game played on this domain which is too simple to be testable, yet which may be of heuristic value for considering the above questions.

We will first elaborate the point cursorily dealt with in the last section of the previous chapter that different types of organizational architecture require correspondingly different types of information-processing capacity – more conventionally, human assets or skills – from participating individual agents. In the previous chapter we compared the informational efficiency of three generic organizational modes under the assumption that the levels of the information processing capacities of agents are quantitatively comparable.[5] However, not only do the information-processing capacities of individual agents quantitatively differ, but there may be important qualitative differences as well. If so, since individual capacity for information processing is generically limited, individual needs to invest in a particular type and sink cost. When agents make this strategic human investment choice, their decisions may be conditioned by a prevailing organizational practice. On the other hand, when entrepreneurs make a strategic choice of organizational architectural design, their choices may be conditioned by available skill types. We will capture such reciprocal relationship as strategic complementarity in the domain of organizational field, analyze the possible co-evolution of a convention in organizational architecture and human assets types, using a simple evolutionary game model.[6] Contrary to the Alchian-Friedman conjecture, we will observe that a sub-optimal convention of organizational architecture can evolve in a closed organization-field under evolutionary selection pressure.

However, as hinted above, the increasing ease of geographical mobility of human assets, inter-penetration of firms of different regional/national origins into others' traditional territories, availability of more precise information regarding better organizational practices elsewhere and the like may contribute to the viability of mutants within sub-optimal conventions having evolved in closed organizational fields. Will the integration or interactions of organizational fields then eventually weed out inefficient forms of organizational architecture in each industry and evolve an efficient configuration of organizational architecture across them? In other words, will it contribute to the realization of gains from organizational diversity in the sense that the most efficient form of organizational architecture prevails in each industry, depending on characteristics of goods and technology involved? In order to examine this issue, we analyze consequences of various types of interactions of organization-fields having evolved different conventions and examine its implications.

In the end of the chapter, we will comment on practical relevance, as well as a limit, of the evolutionary game approach which we will have employed. We will provide our answer for finding a way out of this limit in chapter 9.

5.1. Types of Mental Programs: Individuated vs. Context-Oriented Human Assets

As discussed in the previous chapter, any activity within an organization, including the most primitive physical work, has an aspect of information processing. It implies that when engaged in an organizational activity, the agent runs his/her own mental program – or cognitive mechanism – in order to recognize and interpret the state of relevant environments, predict the consequences of various alternative actions on the state (including other agents' reactions), and make a decision choice from among them to solve relevant problems. Such programs are composed of a bundle of "rules," usually in the form of "if, ... , then ...." For example, a doctor in a hospital may interpret a patient's situation and try to solve a problem by mobilizing stored rules, such as "if the X-ray film shows this image and the stethoscope examination detects that sonic pattern, then the patient suffers from bronchitis," "if this medicine is given, then he is likely to have an allergic reaction. If that case is highly probable, it should not be given to him," etc. Such rules are accumulated, revised over time, organized in the mind of the agent in a certain order, and "triggered" by a perceived situation. (Holland *et al* 1986). Thus, they have aspects of capital – human assets. To simplify the discussion, for now let us regard the mental program of the agent as composed simply of two kinds of rules: *cognitive rules* that are used by him or her to form interpretative representations of the relevant situations from cognitive inputs (e.g., digital data, written reports, icons, conversations, observed gestures and expressions, etc.); and *decision rules* that are used to make an individual choice from the set of feasible actions based on the interpretative representation of the situation.[7] Note that "decision rules" here refer to components of the individual, internal mental program, but not to organizational or administrative rules/procedures by which the collective decision-making process is arranged, although both are undoubtedly in constitutive relationships with each other.[8] We are concerned here with internal organizing principles for these rules, only to an extent necessary for a basic classification of the mental programs relevant to our discussion.

In the previous chapter the obvious fact was pointed out that the execution of an individual task in an organization involves information processing regarding the respective idiosyncratic environment. In that sense, any organization embraces the division of information-processing labor. However, what makes the firm an *organization* is the presence of the common segment of the environment whose state is simultaneously relevant to multiple constituent task units. Without it the division of information-processing labor can be performed just as well without organizational cooperation. And, it is precisely the ways in which that part of the environment is informationally processed that differentiates one type of organizational architecture from the others. We have observed that in this regard functional and horizontal hierarchies stand diametrically opposed.

In functional hierarchies, the environment relevant to an organizational domain is completely decomposed in a disjointed manner for the specialized division of information processing. Thus, the division of information-processing labor is complete. In terms of an individual mental program, this implies that the sole inputs to cognitive rules (the "if ..." part of the rules) mobilized for interpreting relevant environmental situations are messages that each agent processes directly from those segments of the environment, while the inputs to decision rules are his own interpretative representations of relevant situations, together with messages transmitted to him or her from other designated agents in the form of formalized codes, such as commands, briefings, announcements, letters, etc. In this sense mental programs are run separately by agents in different task units. The agent may polish, add, and reorganize his/her own store of cognitive and decision rules based on his/her own experiences and learning. If the agent faces the same (or a similar) task, however, the accumulated store may remain valuable in any functional hierarchy. A mental program embodying the types of cognitive rules and associated decision rules effective in functional hierarchies can thus easily be portable from one functional hierarchy to another, as long as the tasks the agent faces are designed similarly across functional hierarchies in terms of the range and object of information processing required. Thus, we may characterize this type of mental program as *individuated* (separated) and refer to the capacities and skills of individual agents to effectively run them as *individuated human assets*.

The situation is different with respect to mental programs which operate in a horizontal hierarchy (nested information assimilation). In this type of organizational architecture, the information-processing labors of the agents are not entirely decomposed in a non-overlapping manner. Regarding the segment of the environment common to them, the agents in a horizontal hierarchy process information in such a way as to assimilate their interpretations about it and to construct a common basis for decision-making. At the basic biological level, agents individually perceive various messages from the environment just like their counterparts running individuated mental programs do. But, in order to arrive at an assimilated cognitive interpretation and to form a common basis for decision-making, the agents need to utilize, not only their own individually-processed messages from the relevant environment, but also messages - tacit and explicit - from the relevant others regarding their cognition of the same environment, that is, one needs to take into account how the relevant others perceive and interpret the common environment). Borrowing an expression from M. Polanyi (1958), we might as well say that the agents must "indwell" with each other in a common environmental situation.

This is an entirely different type of information sharing from the one we observe in a network-integrated, functional hierarchy in which agents draw the same digitalized information from a formal communication network as input to organizationally-designed decision rules. To arrive at an assimilated interpretation as a basis for decision-making, the agents may possibly discard individually idiosyncratic elements of cognitive inputs as organizationally irrelevant. Thus, cognitive rules and the associated decision rules used by individual agents may involve more dialectic, rather than analytical, reasoning.[9] In this way, some potentially valuable information may be sacrificed, but as we have seen, it may serve the organizational goal better when technological and/or attribute complementarities between tasks are high, the information processing capacities of agents are relatively equal, and the environments of task units are highly correlated and fairly stable.[10]

The type of mental program operating in a horizontal hierarchy is thus made qualitatively distinct from the individuated ones in that it incorporates the mechanism of cognitive and evaluative assimilation as an essential element. As such, the mental programs of this type within an organization are mutually connected. Thus we may characterize this type of mental program as *context-oriented*, and the capacities or skills of agents to effectively run them as *context-oriented human assets*. The essential characteristic of this type of mental program would not by itself preclude its portability across organizations. The required skill for running the mechanism of interpretative and evaluative assimilation, once acquired, can remain valuable if it is put to work with others of the same type. However, the effectiveness of context-oriented human assets may be enhanced by their continued connectedness. In that sense, context-oriented human assets may be more organization-specific than individuated human assets.

So far we have implicitly assumed that appropriate types of human assets are available for both functional and horizontal hierarchies. But this premise can now be questioned. If environmental parameters change, a type of organizational architecture that has enjoyed an efficiency advantage may lose its position to another type. But, can the new architectural structure then quickly replace the original one? If this replacement requires a basic shift in the types of required human assets from context-oriented to individuated, or the reverse, the transition may not be automatic. As defined, the actual mental programs of any individual at any point in time are composed of a bundle of cognitive and decision rules. These rules may be revised, refined, and incrementally restructured over time by individual learning, both formal and on-the-job, and the organizational experiences of the agent. In this sense, any mental program may be simultaneously individuated, context-oriented, and path-dependent in varying combinations. However, a meta-rule for organizing these rules determining whether or not the mechanism of assimilated cognition is incorporated as an essential element of the mental program may not be so flexibly altered once formed.

Then, the situation is seen where agents face the following choices prior to entry into any organization: whether to invest in an individuated human asset useful in a specialized task across organizations, or to invest in a relatively more general, communications and problem-solving skill, expecting that it will be molded into a more context-specific skill after being matched with a particular organization. However, an actual individual choice of which type of mental program to develop is greatly affected, consciously or unconsciously, by the societal institutions of education and training, the cultural meanings attached to an appropriate self, and, more economically, which type of organizational architecture is dominant in the organizational fields to which agents expect to have access. We note that recent works in cultural psychology on this topic are theoretically consistent with, and empirically supportive of, our categorization of individuated vs. context-oriented.[11] Scholars in that discipline argue that while a cultural view may be transformed into either of the attributes of the self, the latter supports and reproduces the prevalent pattern of individual psychology. Thus, psychological processes and a cultural system are mutually constitutive.

Similarly, we submit that a type of human asset (mental program) and a type of organizational architecture may be mutually constitutive and co-evolve. While the formation of individual mental programs useful in an organization may be affected by the dominant organizational architecture, the design of organizational architecture by the entrepreneurs may also depend on the distribution of available human asset types in the population. Thus there is an important complementarity between their choices in the domain of the organizational field. If individuals or entrepreneurs (organizers of firms) deviate from the dominant pattern of types, they may face a higher risk of mis-matching with a wrong type of organizational architecture or pool of human assets. Thus, even if one type of information architecture has a better fit than others with the technological environments of some particular industries, it is not certain whether that architectural structure will dominate everywhere in those industries. If the distribution of types of human assets in the population is highly skewed toward one fitting a particular type of organizational architecture, that type may persist as a convention in the organization field, even if technologically sub-optimal. An organizational convention can thus be viewed as a coordination device by which the choice of investment strategy (human assets formation) by agents is aligned in one way or another in the organization field, thus helping them avoid the risk of costly mis-matching. We formalize this intuition in terms of an evolutionary game in the next section.

5.2. The Evolutionary Dynamics of Organizational Conventions [12]

Imagine a domain, called the organization-field, consisting of a large number of population of agents in which a firm is organized by the matching of two agents. Each agent invests for his/her life-time either in individuated human assets or in context-oriented human assets. Agents die or quit work and are replaced by their children at a certain rate. If two agents who have invested in individuated human assets are matched to form a firm, its organizational architecture will be a functional hierarchy (functional specialization), whereas if two agents who have invested in the context-oriented type are matched, it will be a horizontal hierarchy.[13] Suppose further that there are two types of industries, B and D. Functional hierarchies have an efficiency advantage in industry B (e.g., because of substitutability of constituent tasks), whereas horizontal hierarchies have an efficiency advantage in industry D (e.g., because of high complementarity of constituent tasks).[14] If a mis-matching of two different types of human assets occurs, that organization will be the least efficient in both industries. This assumption can be represented by two matrices, one for each industry, showing the costs of unit-production contingent on the matching of human assets types. Denote the human assets types by I (=individuated) and C (= context-oriented), and the unit output cost of the industry B (alternatively D) by b_{jk} (alternatively d_{jk}) , when the matching of human assets types j and k (=I or C) occur. We have

$$\boldsymbol{B} \equiv \begin{bmatrix} b_{II} & b_{IC} \\ b_{CI} & b_{CC} \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{D} \equiv \begin{bmatrix} d_{II} & d_{IC} \\ d_{CI} & d_{CC} \end{bmatrix}$$

where $b_{II}< b_{CC}< b_{IC}= b_{CI}$ and $d_{CC}< d_{II}< d_{IC}= d_{CI}$. Agents work in one of the two industries, $i = B$ or D. Let us denote the distribution of the population over human assets-types and industries at a particular moment in time as $\boldsymbol{m} = (m_{CB}, m_{CD,} m_{IB,}, m_{ID})$, where $m_{Hi,}$ represents the proportion of the population who choose industry i with type of human assets H so that $m_{CB}+ m_{CD}+ m_{IB}+m_{ID} =1$. A firm can produce two units of output at any moment with a specified unit-cost contingent on the mode of matching, i.e., type of organizational architecture. The revenue of the firm, net of production cost is equally shared between the two agents forming the firm.

Regarding matching technology, assume the following. Agents equipped with individuated human assets are mobile between the two industries at any moment in time. On the other hand, context-oriented human assets sink some costs in a particular industrial-organizational context. Therefore, agents who are invested in context-oriented human assets are not as easily mobile across industries (organizations). Because of their relative immobility, agents with human assets of that type select matching partners more carefully. I assume that the probability of agents equipped with context-oriented human assets being matched with the same type in industry *i* is given by:

$$\pi_{Ci} = \left[\frac{m_{Ci}}{m_{Ci} + m_{Ii}}\right]^{\gamma}, \qquad i = B, D$$

where $0 < \gamma \leq 1$. If $\gamma = 1$, matching is random, and if it is less than one, matching is positively assortative and, as γ decreases, the probability of proper matching increases. I will state shortly how the perfect mobility of an individuated type can be formulated.

All agents in the organization-field domain have identical consumption tastes and spend their incomes on the product of the B and D industries in the proportions β and Δ, where $\beta + \Delta = 1$. Recalling that each firm is producing two units of the product, the total outputs of industries *B* and *D* are $m_{CB} + m_{IB}$ and $m_{CD} + m_{ID}$ respectively. Prices of products *B* and *D* are then given by the unit-elasticity inverse demand function

$$p_B = \frac{\beta}{m_{CB} + m_{IB}}, \quad p_D = \frac{\Delta}{m_{CD} + m_{ID}}$$

respectively.

The expected pay-offs of agents possessing context-oriented human assets and working in industries $j = B, D$ is then given by

$$u_{CB} = \frac{\beta}{m_{CB} + m_{IB}} + \pi_{CB} b_{CC} + (1 - \pi_{CB}) b_{CI}$$
$$u_{CD} = \frac{\Delta}{m_{CD} + m_{ID}} + \pi_{CD} d_{CC} + (1 - \pi_{CD}) d_{CI}$$

For example, an agent who enters the *B*-industry has the probability, π_{CB}, of being matched with an agent of the same type, and consequently bearing the cost b_{CC}, while having the probability $1-B_{CB}$ of being mis-matched with an agent of the different type of human assets and consequently bearing the larger cost, b_{CI}. He receives an equal share of the net revenue with his partner, whoever she may be.

Likewise, the expected pay-offs of agents with individuated human assets working in industries *B* and *D* are given by:

$$u_{IB} = \frac{\beta}{m_{CB} + m_{IB}} + \pi_{IB} b_{II} + (1 - \pi_{IB}) b_{IC}$$

$$u_{ID} = \frac{\Delta}{m_{CD} + m_{ID}} + \pi_{ID} d_{II} + (1 - \pi_{ID}) d_{IC}.$$

where B_{IJ} *(j= B, D)* is the probability of agents with individuated human assets being matched with others of the same type of human assets in industry i. These probabilities can be determined by the labor market clearing conditions:

$$(1 - \pi_{Ci}) m_{Ci} = (1 - \pi_{Ii}) m_{Ii} \qquad i = B, D.$$

That is, even if agents miss a correct-matching, they are matched with an agent of the different type and there is no unemployment. The agents with individuated human assets do not select the same type of human assets in an assortative manner, but they are more flexible in choosing the industry because of their functional specialization useful in either industry. I assume that they can instantaneously choose the industry in which they can expect to earn a higher income, so that the agents with the individuated human asset type are allocated between the two industries in such a way as to equalize their expected incomes.

With this, we have completed the specification of the state of the organization-field domain at any moment in time characterized by the distribution of population $\boldsymbol{m} = (m_{CB}, m_{CD,} m_{IB} + m_{ID})$. The next task is to describe a dynamic process of the organization-field domain over time, along which the distribution of the population evolves, and inquire into the nature of the equilibria of such a process. Though I do not explicitly model the dynamics, a brief description of a process which might underlie such a model may be stated as follows.

At each moment in time a fraction of the population is replaced by a new generation of individuals, most of whom mimic the strategies of their parents. A small fraction of them, however, choose their strategies to mimic the existing strategy with the highest expected return (I will introduce later the possibility that an even smaller fraction experiments with random choice). As a result, only the most successful type will increase its relative share in the population. Such dynamics is called the *best response evolutionary dynamics*. An *equilibrium* of these dynamics is any population distribution $\boldsymbol{m}^* = (m_{CB}^*, m_{CD}^*, m_{IB}^*, m_{ID}^*)$ at which the distribution of population across human asset types and industries becomes stationary. An equilibrium is said to be an *evolutionary equilibrium* if it is locally asymptotically stable (D. Friedman 1993). That is, all states near an evolutionary equilibrium will eventually converge toward it.[15]

There are nine equilibria for the best-response evolutionary dynamics, all of which are Nash equilibria. They are shown diagrammatically in Figure 5.1 which depicts the distribution of the equilibria in the three-dimensional simplex representing $\boldsymbol{m} = (m_{CB}, m_{CD,} m_{IB,} + m_{ID})$ and their dynamic properties. Among them, *P, I, C* and *L* are evolutionary (*IP, IL, CP* and *CL* are saddle points and *W* is a source) .

Figure 5.1 about here

Of the nine equilibria, the *P-equilibrium* is the unique Pareto optimal equilibrium in which an optimal diversity of types of organizational architecture (the most efficient matching) is realized across both industries: that is, functional hierarchies in B-industry and horizontal hierarchies in D-industry. In the *I-equilibrium* and *C-equilibrium* all the agents adopt a uniform choice strategy regarding human assets type, either individuated or context-oriented, regardless of industry and thus functional hierarchies or horizontal hierarchies prevail as sole organizational architectural type. Once these two equilibria are established historically, it would be difficult to upset them in spite of their sub-optimality, because the deviation of a small group of agents from the corresponding equilibrium strategy would be heavily penalized by the larger risk of mis-matching. The adoption of the prevailing human asset type would then become a *convention*, and as a result functional or horizontal hierarchies would be established as organizational conventions regardless of industries. Since neither of these organizational conventions has an absolute advantage but only a relative one, different organizational fields internalizing different organizational conventions may enjoy advantages only in particular industries. The *L-equilibrium* is a pathological equilibrium in which less efficient matching prevails in both industries by historical accidence.

5.3. The Interactions of Organizational Fields and Gains from Diversity

The above model indicates that the Pareto efficient industrial structure involves a diversity of organizational architecture, contingent on the technological and market parameters of each industry, whereas an organization-field in which some type of organizational convention prevails cannot achieve the same level of efficiency. The efficiency gains from the diversity of organizational architecture are referred to below as the *gains from organizational diversity.* The model so far has not predicted which evolutionary equilibria will be likely to emerge, except that it depends solely on the initial condition. As already suggested, however, any economy (national or local) tend to be more or less characterized by the relative uniformity of organizational architecture, although it may be preceded by a period of cohabitation of diverse organizational experiments.

Suppose that at an initial stage of the market economy a more primitive organizational mode, say classical hierarchical decomposition, prevailed, in which operational tasks were served by a simple skill type under commands of the proprietor-entrepreneur. Imagine that multiple organizational experiments subsequently emerged, which relied on the information-processing capacity of the workers at the operational task level. These can be functional hierarchies based on individuated, task-specific human assets, or team-oriented horizontal hierarchies based on context-oriented human assets. In any case suppose that both of them are more efficient than classical hierarchies in both industries, as they can make better use of *ex post* information evolving at the operational level. But suppose that both of them has an absolute productivity advantage vis-a-vis the other in only one industry, as specified in the cost matrix introduced in the beginning of the last sub-section, provided that corresponding human assets type were available.

In the beginning, there might be competition among the new organizational experiments, but suppose that either one gains a momentum due to relatively abundant supply of the fitting human asset type. Then, even if the other one were potentially more efficient in some industry, evolutionary pressure might make the sustainability of the latter increasingly harder. Because of the fear of the higher risk of mismatching, it becomes ever less advantageous for the new generations to invest in the type of human assets tailored to the less dominant organizational architecture (this is an instance of strategic complementarity). Thus the presence of evolutionary pressure suggests that organizational diversity, in the sense of cohabitation of diverse organizational architecture in a single organizational field may not be taken for granted. However, the possibility of multiple equilibria also suggests that the evolution of different organizational conventions across different fields (e.g., national economies and localities) may occur. Given such a possibility of cross-field organizational diversity, let us now consider several possible avenues for exploiting the gains from diversity through interactions or integration of two organizational fields that have evolved different conventions.[16]

The rest of the section omitted

5.4. The Relevance and Limits of the Evolutionary Game Model.

One of major objectives of this chapter has been to see how different organizational conventions could arise in different economies/localities and become a source of relative industrial advantage/ disadvantage, even if potential technologies and tastes were the same everywhere. The reason for the evolution of multiple, sub-optimal organizational conventions is not increasing returns as focused on recently in economics, but complementarity among the strategic choices of agents. If a large proportion of the population is adopting a certain strategy, it becomes the best response for agents to adopt the same strategy. The apparent difference from the contract theory of the organization which prescribes/predicts a (second) best response of the principal to exogenous parameters arises from the fact that the principal-agent theory treat outside options open to the principal and agents as exogenously given, while in evolutionary models, alternatives open to each agent are determined endogenously as a result of the strategic interplay of agents.

The model presented in this chapter is extremely simple: there are only two human asset types and two types of organizational architecture. Technological and market conditions are parametrically fixed. Although the setting of the model thus remains at an extremely abstract level, it hopefully captures some fundamental factors which underlie the observed diversity of organizational architecture across economies and the co-evolution of a diversity of human asset types. More specifically, we submit that if one compares two, (possibly polar) cases of organizational evolution in North America and Japan, one can't fail to notice the striking relevance of the distinction made in this chapter (and later) between functional and horizontal hierarchies, as well as individuated and context-oriented human assets types.

Think of some of the notable examples of innovation in the area of work organization which were initiated and institutionalized in North America, such as: the afore-mentioned "American manufacturing system" developed in the last century in New England machinery industry (Rosenberg 1963, Pine 1993); the Taylorist scientific management movement (whose innovative nature has often been misunderstood —see Wredge and Greenwood 1991); and the subsequent development of management hierarchies, bureaucratization of the personnel administration (Jacoby 1985, Baron *et al* 1986) and associated job controlled unionism in the 1930's and 40s. All these examples have a common characteristic in either having introduced or institutionalized a new method of combining individuated human assets at progressive degrees of maturity — human assets that became embodied in individual workers, engineers, and managers through professional and vocational training, even though some elements of organizational contextuality cannot be entirely ignored. The organizational innovation which has recently taken place in Silicon Valley may appear a radical departure from traditional functional hierarchies in which bureaucratic control of highly segmented jobs was the norm. However, an excellent description and analysis of Baldwin and Clark (2000) shows, it may be also considered as an evolutionary outcome from the centralized control over modular tasks as represented by the design of IBM 360. We will analyze later in chapter 10.2 and 13 in what respects the Silicon Valley model differs from, as well as conforms with in some respects, the traditional functional hierarchies and why it evolved in a periphery of industrial America.

If we turn to historically known examples of organizational practices and innovations which affected the evolutionary path of organizational practices in Japan, a striking contrast to the American path is immediately discernable. For example, think of such examples as: the design of the seniority and bonus payment system by advanced factories at the beginning of the twentieth century as means of restraining excessive quits of skilled workers; collective, ad hoc problem-solving self-organized by the workers on the shop floor in response to the scarcity of tools and materials during the second world war; the transformation of the American-born, engineer-led quality-control system into shop-floor level work-group practices; the evolution of the "kanban" system in the 1950s which partially emulated an inventory restocking method used by American super-market firms.[17] In contrast to the American case, reliance on information channels shared by workers within the context of a particular organization is a distinct attribute, even though more recent innovations seem to increasingly accommodate elements of the workers' individuated information processing skills. For example, the now famous "kanban system" was not able to be implemented without the ability of individual workers to cope with emergent events on the spot, such as breakdown of machines, or spotting defective parts . But even these abilities were backed up by the workers information networking channels (e.g, mutual help in teams, the sharing of engineering and production knowledge through job rotation and cross functional meetings, company-specific training programs). It is worth noting that, when the idea of the "kanban" system was brought back to America in the 1980's as a lesson from the so-called "Japanese management," it was transformed into something different, yet consistent with the American path — the "lean production method" (Warmack *et al* 1990) which would reduce inventories and hierarchical-layers of production-control by flexible matching (outsourcing) with specialized suppliers.

The brief observations in the last two paragraphs also indicate the limits of applying the results of a "stationary" evolutionary model for interpreting real phenomena, however. By "stationary" we specifically refer to the assumption that the strategic choice set of each agent is exogenously given and fixed. Therefore, in our theoretical model, "innovation" in one economy can occur only in the form of a shift from one equilibrium to the one characterized by diversity. Competition among economies often induces an attempt to emulate an organizational convention prevailing in a foreign economy which may be perceived to be a superior practice. However, as just mentioned above regarding the evolution of the "kanban" system and its recycling into North America, learning from a foreign practice may not result in a simple transplant or a hybrid, but to the enrichment or adaptation of indigenous organizational practice. Also, entrepreneurial experiments that eventually lead to the rise of a new type of organizational architecture may actually be a new type of bundling of existing skills, as in the case of Silicon Valley firms. Both possibilities alter and enlarge the agents' sets of strategy choices.

Thus actual organizational evolutionary processes involving innovation may not then be characterized as a mere shift from one convention (such as *C*- or *I*- equilibrium in Figure 5.3) based on one type of human assets to a diversity merely mixing the two (such as *P*- equilibrium). One conjecture could be that the organizational evolutionary process can be characterized by successive equilibria, at each of which one type of human asset is sequentially enriched by learning and intentional design. In this regard, the following comment made by the pioneer of the evolutionary game theory in biology, Maynard-Smith, is highly suggestive: 'Whenever an optimisation or game-theoretic analysis is performed, an essential feature of the analysis is *a specification of the set of possible phenotypes* from among which the optimum is to be found. This specification is identical to a description of developmental constraints.'(1982:5. Italics added) In other words, the analysis of optimization (i.e., evolutionary selection) can explain the phenomena of multiple organizational conventions once developmental constraints (i.e., possible types of human assets) are given, but cannot explain how developmental constraints themselves are determined. In order to understand the mechanism of organizational change, therefore, we need to go beyond the scope of "stationary" evolutionary modeling as presented in this chapter. A preliminary attempt toward such direction will be attempted in chapter 9.

CHAPTER 5

* Arrow (1974:56).

1. For example, contract theory regards questions, such as whether multi-tasks are jointly or separately assigned to workers, as basically answerable depending on technological parameters. For example, if technological inter-relatedness among tasks in the sense of statistical correlation is high (alternatively low), it is incentive-wise preferable to separately (respectively jointly) assign workers to those tasks (Holmstrom and Milgrom 1991; Itoh 1992). Transaction cost theory submits that [explicitly considers the choice of organizational architecture between unitary and multi-divisional forms of internal organization on the efficiency criteria. The increasing complexity of communicational, technological and market conditions warrant the adoption of the multi-divisional form (Williamson 1975). On the other hand – remove], if the specificity of physical assets is reduced, a non-hierarchical governance mechanism may be adopted to mitigate its inherent disincentive impacts, but outside of the organization as relational-contracting (Williamson 1985). See Williamson (1995) for a defense of the one-dimensional transaction cost criteria.

2. Pagano (1999a, 1999b) applies the biological concept of "allopatric speciation" to account for such phenomena..

3. For example, see Aoki (1986, 1990), Dosi, Pavitt and Soete (1990), Saxenian (1994), Teece *et al* (1994).

4. The notion of the "organizational field" is due to new institutional sociologists. According to them, it refers to "those organizations that in the aggregate constitute a recognized area of institutional life."(DiMaggio and Powell 1983:143) It connotes "the existence of a community of organizations that partakes of a common meaning system and whose participants interact more frequently and fatefully with one another than with actors outside of the field (Scott 1994: 207-208). We will define it as the domain in which an convention in organizational architecture may be formed as an outcome of an evolutionary game played by the agents. For a sociological model of the evolutionary approach to organizational population dynamics, see Carroll and Hannan (1989) and Hannan and Carroll (1992).

5. The underlying framework was Bayesian where agents form posterior beliefs on the state of the environment by revising prior beliefs on the basis of observed samples containing certain errors. Then the agent's information-processing capacity was defined as the ratio of the prior variance of observed variables to the variance of sample error.

6. As for analytical works dealing with the co-evolution of the social order and

individual preferences, see Kuran (1991) and Bowles (1996). Also Pagano (1993) and Pagano and Rowthorn (1994) analyzes the codetermination of property rights arrangement (capitalist vs. labor rights regimes) and technology (easy-to-monitor capital vs. easy-to-monitor labor) as “organizational equilibria” (Nash equilibria) under strategic complementarities, where "production managers" choose that technology that maximizes profits given the existing property rights system and "financiers" arrange that property rights regime that maximizes ownership rent given the existing technology. Then they argue that the simultaneous “homogenization” (i.e., conventionalization) of property rights and technology regimes occur because of “network externalities” (i.e., economy-wide strategic complementarities).

7. We will propose a more complex model of mental program called the “subjective game model” in chapter 9.

8. Thus defined, mental programs or human assets in this chapter are seen as a close analogue of the “mental models” conceptualized in Denzau and North (1994). The relationships of both will be explored more in chapter 9.

9. As noted below, recent theoretical and empirical achievements in cultural and social psychology provide ample support for the distinction made here. For example, Peng and Nisbett (1999) demonstrated through their experimental studies that, when

presented with contradictory accounts of a situation, Chinese subjects attempted to find a "middle way" in which both accounts had some validity, whereas American subjects tended to reject one side entirely in preference to the other. See Fiske *et al* (1998) for a survey on recent psychological studies on analytic versus holistic modes of thought.

10. However, when the state of the environment becomes volatile and highly uncertain, the cognitive representations by agents may not easily converge. In such a situation, the information-assimilation mode may not function. We consider this issue and its implications in chapter 9.

11. For example, Markus and Kitayama summarize a series of their works in an attempt to integrate cross-cultural differences in many different psychological domains, such as cognition, emotion, and motivation, as follows:

> "Western, especially European American middle class cultures are organized according to meanings and practices that promote the independence and autonomy of a self that is separate from other similar selves and from social context. ... Those in Western cultures may then be motivated to discover and identify positively valued internal attributes of the self.... In contrast, many Asian cultures do not highlight the explicit separation of each individual. These cultures are organized according to meanings and practices that promote the fundamental connectedness among individuals within a significant relationship

(e.g., family, workplace, and classroom). The self is made meaningful primarily in references to those social relations of which the self is a participating part. Those in Asian cultures may then be motivated to adjust and fit themselves into meaningful social relationships." (Kitayama, Markus, Matsumoto and Norasakkunikit, 1997, p.1247)

12. This and next sections draws on Aoki (1998). Although all proofs of propositions are omitted and relegated to the article, these sections are still relatively technical. Those who are not particularly concerned with the formalization of the evolutionary thinking beyond what was stated in the last section may choose to skip this chapter.

13. An alternative modeling strategy could be to consider a matching game defined on two populations instead of one: the population of entrepreneurs whose strategy is composed of the choices of organizational architecture and industry, and the population of workers whose strategy set is composed of human asset types and industry choices. More realistic though this alternative modeling may appear, doing it will add little substance to the results obtained in this chapter. Also, in assuming that the matching of individuated human assets will lead to a functional hierarchy, I do not treat explicitly the disparity of information processing capacity between them which is an essential characteristic of a functional hierarchy (recall Proposition 4.3). However, this "quantitative" disparity within the class of individuated human assets is a secondary

importance in comparison to the qualitative difference between individuated and context-oriented human assets in the current context and thus ignored.

14. Recall Proposition 4.1.

15. In the setting of the current model, any evolutionary equilibrium of it is evolutionarily stable strategies in the sense of M. Smith (1982).

16. For evolutionary models dealing with interactions of different economies and their consequences, see Boyer and Orlean (1992) on firms, Matsui and Okuno-Fujiwara (1997) on culture, and Matsuyama, Kiyotaki and Matsui (1993) on currencies.

22. See for example, Okazaki (1993, 1997), Aoki (1996b), Okazaki and Okuno-Fujiwara (1998), Moriguchi (1998), Fujimoto (1999).