Journal of Institutional Economics (2007), 3: 1, 1–31 Printed in the United Kingdom
© *The JOIE Foundation 2006* doi:10.1017/S1744137406000531

Endogenizing institutions and institutional changes*

MASAHIKO AOKI
Stanford University

Abstract: This paper proposes an analytical-cum-conceptual framework for understanding the nature of institutions as well as their changes. First, it proposes a new definition of institution based on the notion of common knowledge regarding self-sustaining features of social interactions with a hope to integrate various disciplinary approaches to institutions and their changes. Second, it specifies some generic mechanisms of institutional coherence and change – overlapping social embeddedness, Schumpeterian innovation in bundling games, and dynamic institutional complementarities – useful for understanding the dynamic interactions of economic, political, social, organizational, and cognitive factors.

1. Introduction

A consensus seems to have emerged among economists, as well as among other social scientists, that 'institutions matter' for understanding the differences in economic performances among various economies over time and space (e.g., Nelson and Sampat, 2001; the World Bank, 2002). But, if institutions are nothing more than codified laws, fiats, organizations, and other such deliberate human devices, why can't badly performing economies design (emulate) 'good' institutions and implement them? This question would naturally lead us to a more fundamental, ontological question of what institutions are. Institutions are customarily identified with the 'rules of the game' by economists and others (e.g., North, 1990: 5). However, within this thrust, their nature, origin, and relevance to economic analysis are seemingly treated in two different ways.

One treats various institutions as rules in a hierarchical order. According to this view, there are rules exogenously pre-determined outside the domain of economic transactions, such as legal rules and social norms, while economic institutions, such as contracts (markets), organizations, and hybrids, are regarded as rational transaction-cost-saving responses within those constraints (North, 1990; Williamson, 2000). Also, pure theorists working in the fields of social

*This paper is a revised version of an invited paper presented at the 2005 World Congress of the International Economic Association held in Morocco. It also draws on my previous book, *Toward a Comparative Economic Analysis*, 2001, MIT Press. I am grateful to referees of the paper for their stimulating and thought-provoking comments and criticisms.

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

choice and mechanism design (e.g., Gibbard, 1973; Hurwicz, 1996) examine a related normative question of whether the rules of game with some socially desirable properties can be designed prior to the operational playing of the game by players of diverse orientations.

The other way is to treat institutionalized rules as something spontaneously and/or endogenously shaped and sustained in the repeated operational plays of the game itself. Hayek (1976) and Schotter (1981) are thought of as pioneers in this regard among economists, while various recent game-theoretic approaches to identify institutions with some kind of equilibrium outcome are considered attempts to provide an analytical foundation for this view (e.g., Greif, 1997, 2006; Aoki, 2001; Young, 1998; Calvert, 1995; Dixit, 2004). Running the risk of oversimplification, let us refer to the first view as the exogenous view of institutions and the latter as the endogenous view.

The focus of this paper is on institutional change as was suggested at the beginning. It will try to examine how the two views deal with various issues related to it. Which view provides a more appropriate framework for understanding institutional change? Alternatively, are these two views reconcilable or complementary to each other, or should each of them be modified after taking account of other's merits?

In order to deal with these and other issues, we start out by making rather stark, static, game-theoretic characterizations of the exogenous and endogenous views of institutions. Then, we incorporate an aspect of bounded rationality of the agents to the endogenous view, formalizing the notion that the agents may not know details of other agents' intentions in the game in which they are in, but can hold a common perception about salient patterns of the ways by which the game is being played (Section 2). An institution thus conceptualized is essentially endogenous, but appears to be an exogenous constraint to the individual agents. However, one sharp contrast between the two views may still remain in that, while the exogenous view takes a dichotonomous approach to separate the rule-making game and the operational game,[1] the endogenous view takes an integrative approach. We deal with this difference in two steps. First, we introduce four prototypes of the domain of game in which social norms, political states, economic contracts, and organizational architecture may potentially evolve respectively as multiple equilibria, thus leading to a variety of institutions (Section 3); and then proceed to an examination of possible linkages and interdependencies, rather than a hierarchical ordering, among institutions across those domains. Specifically, we provide game-theoretic treatment of such intuitively attractive notions as institutional complementarities, social embeddedness, and institutionalized linkages – as instances of equilibrium phenomena (Section 4). However, these static treatments cannot confront squarely the question of how the agents form a common perception of the state of game, thereby leading to the selection of an associated institution out

1 For various treatments of the dichotomy approach, see Levi (1988), Hurwicz (1996), Amable (2003).

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

of the possible many. In order to consider this question, we then move on to dynamic considerations. We first discuss what factors are likely to trigger a crisis of an existing institution and then try to understand how the bounded rational agents can transit to a new institution, by focusing on the process in which they revise their individual expectations about how the game is to be played and eventually reach a modicum of common expectations with the help of ideological and entrepreneurial factors as well as past legacies (Section 5). Section 6 conceptualizes the dynamic counterparts of institutional linkages and interdependencies across domains – dynamic institutional complementarities, overlapping social embeddedness, and Schumpeterian bundling innovation. In these mechanisms, social, political, economic, and organizational factors interact rather than operate in unidirectional manner. For example, a consequence of the political-exchange game (i.e., a policy) may affect an institutional framework of the economic-exchange domain, but institutional changes in economic and organizational domains may conversely affect the institutional structure of the polity. Section 7 concludes.

2. Institutions as common knowledge at equilibrium

In order to make a conceptual distinction between the exogenous view and the endogenous view of institutions, first in sharp contrast and then as reconciled, we will introduce the notions of 'game-form' and game. The game-form is expressed as a pair of the domain and the consequence function defined on it. The *domain* is further composed of the set of agents mutually interactive in certain kinds of interactions (economic, political, organizational, or social) and the sets of activated action choices for each of them. With each profile of all the agents' action choices [and each state of nature uncontrollable by them], the *consequence function* specifies particular (physical) consequences of concern to some or all the agents [contingent on the state of nature].[2] The set of agents may include not only natural persons but also organizations such as the government, corporations, and the court, depending on context. The set of a particular agent's action choices can be conditioned by his/her mental state, personal development, acquired skills, and the like in the case of a natural person, and by the accumulated collective know-how, scope of collective actions, and so on in case of an organization. Formally, an agent's set of potential action choices can be thought as being infinite, but only finite ones of them are regarded as activated for a period of time. Other aspects of the individual agents' bounded rationality will be introduced later in this section and in Section 5. Until then let us be satisfied with saying that only a limited list of finite action choices is actively considered for choice by each

2 The terminology 'game-form' due to Gibbard (1973) is distinct from the classical notion of the game defined as 3-tuple of the set of players, the sets of action choices facing them, and the sets of pay-off functions. Indeed, a game-form is a game with no individual utilities yet attached to possible consequences and is meant to capture only the objectively identifiable parameters of the game.

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

agent. The consequence function can be conditioned by available technologies, formal rules with specific rewards or penalties to particular actions (such as laws and fiats), and other relevant external factors (such as institutional parameters in other domains as specified later). Thus, the game-form can be considered fixed for a period of time, but should be viewed as historically conditioned. As pointed out by Field (1981), it is not possible to construct a completely history-free game-form.

Suppose, given a game-form, each agent tries to choose an action (or more generally a plan of actions, each one contingent on an evolving state of the domain) that (s)he considers the most suitable according to his/her own preferences, given his/her expectations about others' choices and associated consequences. The 'maximization' of preference need not to be taken to imply that agent is exclusively 'self-interested' but that (s)he may also exhibit genetically programmed altruistic concerns, sense of appropriateness, and the like.[3] Thus, the agents in the domain can be conceived of as the players of the game with the *formal rules* of the game specified by the consequence function. From now on, therefore, we will use the word agents and players interchangeably. Viewing social interactions among agents as a 'game' should not be regarded as idiosyncratic to mathematical game-theorists who have developed elaborate analytical tools, however. Such a view, albeit informally, can be traced as far back as to the writings of Adam Smith (1759), as well as to those of prominent scholars of trans-disciplinary orientations, such as Hayek of later years (1988) and Braudel (1958).[4]

The reason why we introduced the utility-independent game-form prior to the explicit reference to the game is that it is useful for pinpointing the differences in the exogenous and endogenous views of institutions and the implied nature of institutional change in the literature. For example, we may distinguish:

- *An institution as a game-form:* As noted already, the neo-institutionalists like North (1990, 2005) identify institutions with formal rules such as constitutions, statutory laws, and contracts, as well as informal rules such as social norms. Those rules may be considered representable in a game-form, i.e. with specifications of the parameters of the consequence function as well as 'permissible' constraints on the sets of agents' action choices (Hurwicz, 1996).

3 The recent development of evolutionary psychology suggests the existence of genetically programmed altruism. For economists, Gintis (2000)and Field (2001)are good introductions. See also Gintis (2004) for a justification of the postulate of preference consistency from the viewpoint of gene-culture co-evolution theory.

4 Adam Smith refers to 'the great chessboard of human society (in which) every single piece has a principle of motion of its own, altogether different from that which the legislature might chuse to impress upon it' (1759: 234). For Hayek, see the *Appendix E* of Hayek (1988), in which he stated that '(t)he practices that led to the formation of the spontaneous order have much in common with rules observed in playing a game' (154). Braudel, the foremost scholar of the French *Annales* School of History, made an interesting suggestion on the possibility of 'qualitative social mathematics' as tools for comparative and historical analysis, as well as for crossing lines of different social science disciplines. He admitted that this idea was inspired by 'the games of Von Neumann and Morgenstern' (Braudel, 1958/1980: 38–52).

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

One well-discussed problem with this view is how these rules are enforced (see Greif, 1997, 2006). They may be enforced by particular organizations, such as the court, or social sanctions, but then the question can be raised as to how the enforcer(s) is motivated to enforce the specified rules, which leads to the infinite regression of who enforces the enforcer(s), who enforces the latter, ad infinitum. Another question related to the above, which we focus on in this paper, is that of who formulates the formal rules and how this is done. North considers that formal rules reflect the cultural/ideological belief system of agents, particularly that of the influential and powerful and it can be 'changed by fiat' (see, e.g., North, 2005: chapter 5). In other words, they visualize a kind of hierarchical ordering in which the political structure (and the social structure in the case of social norms) formulates rules for the economic domain (e.g., Williamson, 2000; Levi, 1988). But then how are the rules of the polity (and social norms) constructed? Are they formulated endogenously in the polity (and through social interactions)?

- *An institution as an endogenous equilibrium outcome of the game*: A clue to the solution to the problem of infinite regression, as noted above, can be provided by endogenizing the enforceability question. Suppose that all the players, including the enforcer of the rules, responds with the best action choices given their respective information regarding possible states [characterized by the state of nature and over-all pattern of action choices by all the agents]. If and only if agents' action plans and beliefs become mutually consistent and repeatedly implementable, then (salient features of) those plans may be regarded as a sustainable (enforceable) rule of the game, and thus as an institution. Such ideas have been entertained by various authors, using diverse specifications of game and associated concepts of equilibrium. However, how can agents, presumably bounded-rational in information processing, find mutually consistent choices? Do they need to know all the details of evolving states, each time when they arise, in order to arrive at mutually consistent choices? Not only that, is it necessary that each agent knows that every other agent knows that, that every other knows that, etc. ad infinitum? (Lewis, 1969). What are relationships between formal rules represented in the game-form and rules of game endogenously constructed? In the equilibrium approach, how can the notion of institutional change be entertained? By the gradual change of an equilibrium in response to changes in the parameters of a game-form? Or, by a qualitative jump of equilibrium a la Schumpeter? If so, what could cause a quantum jump in equilibrium?

The differences between the two views, as they stand, may appear sharp. However, let us try to gradually explore a way to reconcile them. We will do this step by step by considering the interactions and interdependencies of games across domains and over time. As a first step, we introduce a concept of an institution that we are going to rely on, modifying the equilibrium view by incorporating an aspect of the bounded rationality of individual agents: the limit of their capacity to analyze the objective structure of game.

In order for all the players' action choices to become mutually consistent and sustainable (thus in equilibrium), each player need not know the details of the

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

other players' intentions and choices. In addition to relatively fine, idiosyncratic information relevant to their own choices, such as 'the knowledge of particular circumstances of time and space' (Hayek, 1945), it is sufficient for the agents to know only the salient features of the ways the game is being repeatedly played via a certain medium. We assume that agent's information structure is represented by an information partition of the set of possible profiles of agents' action choices. Let us call each of such profiles a state of game or simply a state.[5] Then, if an agent cannot distinguish one state from others, those states may be grouped into the same cell (information set) of his/her partition. We allow that agents may have different partitions. That is, some agents may be more (less) informative than others' in the sense that their partitions are finer (coarser). Also, some of them may be more informative about a certain category of states, being able to distinguish possible states more finely in that category, but not so in others. Although agents' information partitions may be thus various and intersecting with others in intricate ways, we assume that there can be knowledge common to all at the stable (equilibrium) state of the game; that is, what even the least informed agent in the domain knows about (technically speaking, it is the meet of agents' information sets containing the actual state of the game, i.e. the finest partition which is at least coarser than any of the agents' information sets]. As such, the common knowledge captures the salient features of the state of the game and would be invariant and self-sustaining when the outcomes of the game are kept fairly stable, only subjected to marginal changes in response to minor changes in the underlying game-form and other possible disturbances. If it remains stable, common knowledge may have a meaningful linguistic/symbolic representation, such as rules, as we will discuss in more detail shortly. It may also provide each agent on the domain with rough expectations regarding what the other agents on the domain are doing and how they will react to what s/he does. In the spirit of the game theory, such expectations may be referred to as 'behavioral beliefs'.

We thus propose the following conceptualization of an institution.

> *An institution is self-sustaining, salient patterns of social interactions, as represented by meaningful rules that every agent knows and are incorporated as agents' shared beliefs about how the game is played and to be played.*

A few caveats are immediately due. Albeit not explicit in this conceptualization, by referring to the 'self-sustaining' patterns of social interactions, we implicitly assume the existence of a particular equilibrium path allowing for (endogenous formation of) common knowledge. Technically, an institution is then said to consist of common knowledge among the players regarding a particular equilibrium path of the game from the many possible paths. Although

5 We are adopting here the Aumann's event-based approach to modeling the information structure. See Fagin *et al.* (1995: Chs 1–2). Osborne and Rubinstein (1994: Ch.5).

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

we do not specify a particular equilibrium concept at this level of abstraction, its requirement may be conceptualized as a Nash equilibrium relative to the distribution of information among agents.[6] As a result, multiple equilibria are more likely to ensue. However, this is not problematic for a theory of institution. Institutions are humanly devised constructs and could be diverse, not being just a mechanical transformation of natural factors conditioned prior to game.[7]

The common knowledge regarding the self-sustaining, salient features of social interactions as technically defined may not by itself have an intrinsic linguistic or symbolic representation. However, by being known to everybody, relevant features of the state of games (states) are normally likely to have such an expression. For example, such knowledge may take the rule-form: 'If somebody acts in such and such a way, then the typical other (say, the court, government, etc.) will act in such and such a way (so that such and such a consequence will fall on him/her)'. This conditional sentence may be regarded as an expression of expectations common to all the agents as regards the equilibrium play at a sub-game (state) in which an agent moves in the described way. Some agents, like professional lawyers, may have more detailed knowledge about possible consequences (e.g., the terms of punishment), while some others may be deterred from being engaged in the described action, even with a much less specified expectation regarding its consequence (e.g., even a simple scare image of the jail might do). Thus, there can be differences in the fineness of information sets among the agents, but there can be a primitive proposition or a rule summarizing them all as far as their information sets are mutually intersected and sustained.[8] And what is important is that in order to have a common knowledge, it is sufficient and necessary that every agent knows that such a proposition or rule

6 More technically, we can perceive that each agent assigns subjective probability distribution to its partition and that an institution is defined as the meet of all the agents' information cells with positive probabilities (or substantive probabilities) assigned. If subjective probabilities are Bayesian relative to the individual information partition with prior being common, then equilibrium may be characterized as 'correlated equilibrium' (Aumann, 1987). The set of correlated equilibrium contains the set of mixed strategy Nash equilibrium. The mixed strategies may be interpreted as reflecting the uncertainty of agents' conjectures about other agents' choices, although each agent always chooses a definite pure strategy. In this setting, subjective probabilities may differ among agents depending on their information processing competence and experiences, yet there arises an institution defined as shared aspects of agents' beliefs. The common prior may be interpreted as reflecting the deep generic structure of the game (the so-called Harsanyi doctrine).

7 The multiplicity was emphasized by Sugden (1986) as a precondition of the equilibrium conceptualization of conventions.

8 We rely on the Aumann's event-based approach to modeling common knowledge. There is another logic-based approach, called the Kripke approach, in which a set of 'primitive propositions' is associated with possible events, of which truth the agents may or may not know. But the two approaches are shown to be essentially equivalent and thus it can be reasonable to assume that some primitive proposition may be associated with events (subset of game-states) of interest. See Fagin *et al.* (1996:41)

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

is true and that everybody else knows that it is true (Aumann, 1976).[9] Thus, as the philosopher Searle observed, 'collective linguistic and symbolic acceptance' may be thought of as being the essential element of institutions (1995, 2005). However, as we will discuss shortly, we emphasize that such linguistic and symbolic representations are substantiated, and need to be reconfirmed, by repeated equilibrium plays of game, but not as an *a priori* entity. Also, the above example suggests that, although some economists identify institutions with some organizations as such, an organization can appear as an institutional reality (e.g., the court, government, etc.) as a subject in an institutionalized proposition or rule.[10]

Once an institution acquires a linguistic or symbolic representation recognized by every agent, it may be regarded to exist as an objective reality. Its validity can be tested by an actual choice. For example, the objectivity of the rule that smuggling will be punished can be tested and experienced by actually violating the rule. On the other hand, unless it also constitutes an internal belief shared by all the agents, any social rule may be irrelevant to their action choices and thus may not be taken as institutionalized. For example, even when the objective existence of a statutory law in the books is unquestionable, if nobody believes it to be implementable or enforceable, it will not prevail as an institution. This indicates the difficulty of changing an institution just by enacting a law or issuing a fiat. A law may certainly change agents' expectations, but whether they will yield a sustainable outcome consistent with the original intention of the legislature cannot be taken for granted, unless it constitutes a sort of collective mind-set. An epistemically objective rule needs to, and can, coordinate subjective beliefs of agents and generate cognitive mechanisms of their own involving some common elements.[11] In a highly significant paper, Denazau and North argued that agents construct 'mental models' to interpret and produce expectation about the environment, and some type of them are 'shared intersubjectively' partly through institutions (as well as by ideologies and a culturally provided set of categories). They posit that '[t]he mental models are the internal representations that individual cognitive systems create to interpret the environment; the institutions are the external (to the mind) mechanisms

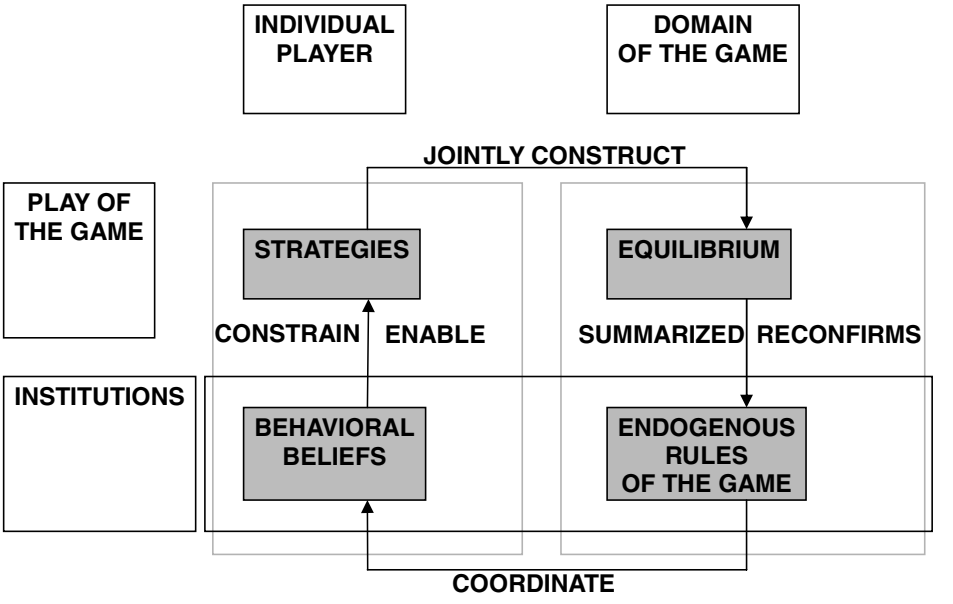
9 It is unnecessary to use the Lewis' definition of common knowledge, which involves an infinitary intersection of information sets, such as: the event that each knows, that all other agents know it, each knows it, and so on ad infinitum (Lewis, 1969). See Faigen *et al.* (1995: Proposition 2.5.1)

10 Nelson identify institutions with prominent organizations such as 'industry associations, technical societies, universities, courts, government agencies, legislatures, etc.' (Nelson, 1995: 57). Greif (2006) also identifies 'organizations' as elements of institutions together with 'rules, norms, and beliefs'. On the other hand, North (1990) opposes such views, regarding that an organization is a player of a game, but not an institution as in the rules of game. My position in the text may be regarded as closer to that of North.

11 Some sociologists (e.g., Meyer and Rowan, 1977; Powell and DiMaggio, 1994; Scott, 1955) regard a shared cognitive framework the essential element of institutions, such as the 'schema' to be relied on in the process of information processing or the 'scripts' guiding what agents do in certain circumstances.

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

Figure 1. Institutions as rules cum shared beliefs



individual create to structure and order the environment' (1994: 4). We may however regard the shared aspects of mental models as constituting internal representations of institutions, and thus institutions characterized by objective–subjective duality, as shown in the lower row of Figure 1.

The duality of an institution just mentioned leads to another enabling/constraining duality, indicated by the flow from the lower-left corner to the upper-left in the figure. An institution constrains each agent's action choices through the beliefs implied by it. Indeed, North once defined institutions simply as 'humanly devised constraints that shape human interaction' (1990: 3). However, an institution also enables the bounded-rational, information-processing-ability-constrained agents to arrive at mutually consistent choices in an information-saving manner. This is somewhat analogous to the situation in which perfectly competitive equilibrium market prices are supposed to summarize information regarding the preferences and technologies facing market participants in the most information-efficient way for sustaining Pareto-efficient equilibrium (see Hayek, 1945; Koopmans, 1957; Hurwicz, 1960). Each market participant only needs to know competitive prices, which dimension equals the number of goods minus one (with one particular good serving as the numeraire). A difference between this approach and ours exists, however, in that what is implied in an institution is not a summary representation of exogenous data of the game, such as technology and preferences, but a summary representation (rules cum beliefs) regarding how the game is being played. Each player may

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

collect information and form expectations regarding other players' choices and intentions in a manner idiosyncratic to his own choices. Therefore, as already said, there may be wide differences and variety in how finely their information sets are partitioned in the space of action profiles. However, each player cannot, and need not, know the choices of all other players in their entirety. It may be sufficient for them to share some rough ideas regarding how the game is repeatedly played in terms of institutions in addition to some localized individual-relevant knowledge of the game. Being guided by such information, bounded-rational agents can economize on information processing and still arrive at mutually consistent choices, although there is no guarantee whatsoever that its outcome is the most efficient one as a Walrasian equilibrium or even a Nash equilibrium.[12]

Finally, even if an institution is crystallized as an epistemically objective rule, in order for it to be sustained and relevant, it needs to be consistent with, and confirmed by, salient features of evolving states of games via agents' repeated choices. For example, agents can be deterred from importing legally prohibited goods, if they believe that 'if I smuggle, I am likely to be caught by the enforcement officer and penalized according to the law'. However, if it is widely observed that many agents are able to import the goods by bribing the law enforcer and thus escape a punishment, this belief will not be sustained. The unenforceable law is hardly qualified to be an institution, but the practice of bribing corrupt law enforcers does qualify. In this simple example, there may be at least two equilibria and accordingly two institutions: an institution of the rule of law and an institution of bribery. If the rule-of-law institution prevails, then compliance with the law would appear to be imperative for an individual agent to avoid a penalty. An attempt to smuggle legally prohibited goods by bribing the enforcement officer would appear to be futile (except as a random drift tending to be absorbed into an equilibrium). Alternatively, in the bribery institution, an attempt by an honest enforcement officer to enforce the law may be frustrated. Thus, either of the two institutions, once established, would appear as an external constraint to individual agents. However, in order to be sustained and viable as an institution, corresponding (*de facto*) rules and associated beliefs need to be continually reconfirmed and reproduced through relevant strategic plays of the players. Thus an institution may be characterized by still another duality, i.e. the exogenous–endogenous duality.[13]

Now we have completed a circular tour, starting from (the common knowledge of) equilibrium to a rule (as its linguistic representation), then to shared behavioral beliefs, and finally back to equilibrium. Thus, equilibrium, rule, and beliefs are considered to be mutually reinforcing and interdependent.

12 As implied by note 6, correlated equilibrium may not necessarily be Nash equilibrium.

13 This exogenous/endogenous duality was the focal point of the phenomenological approach to the sociology of knowledge by Berger and Luckmann (1966).

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

But then a problem arises as regards where and how the circular movement starts.[14] This question cannot be dealt with within a stationary framework, however, but calls for an explicit treatment of institutional change, i.e. the process of equilibrium displacement and its reconstruction. We will take this up later in Section 5.

3. Prototypes of domain and associated proto-institutions

So far we have imagined the generic game played in an abstract domain and conceptualized in an institution as an endogenous rule cum shared behavioral beliefs without specifying its content. However, actual institutional dynamics appear to involve interactions of economic, organizational, political, and social factors. As a first step toward the analytical treatment of such an interactive process, this section attempts to specify prototypes of domains that may capture some minimally essential elements of each of those factors that may arise prior to interactions.[15] The following section will then try to capture the nature of the interactions among those factors in the context of games across those domains.

Let us consider the following four prototypes of the domain distinguished by types of players and types of interactive choices. We note for each of them a unique challenge that must be solved in order to sustain interactions (exchanges) and the proto-types of institutions that may arise in response to such a challenge. Although we attempt to formulate each of the prototype domains in a primitive game-form as much as possible and identify possible institutional elements as something endogenously emergent, it is inevitable, as we noted already, that some primitive institutional elements are already implicitly presupposed in the formulation of game-form.

- *The economic exchange domain*: This is the domain in which transactions of private goods take place.[16] The most primitive type is the domain composed

14 This problem may be considered as a variant of the common knowledge problem (see Lewis, 1969).

15 More generally, one may conceive of types of domains that 'intersect' with other domains and generate cultural values and beliefs, although the ways in which such domains can be analyzed in terms of a game with a broader concept of 'strategies' have hardly been explored. See Hayek (1976) who referred to the need to learn from an insufficiently appreciated work by the historian J. Huizinga (1949) about 'the role of play in the evolution of culture' (71 and n.10). For an interesting non-game-theoretic exploration on the co-evolution of culture and civic society, see Ikegami (2005: chs 1 and 2). Greif (1993) analyzes relationships between cultural values and contract-enforcing institutions with reference to the game-theoretic framework. In Binmore (2005), culture is assigned a very important role for selecting a 'social contract' (equivalent to our 'institution' in being 'the set of common understandings that allow the citizens of a society to coordinate their efforts' (2005: 3). He emphasizes that 'culture means more . . . than the fact that it is common knowledge that everybody is rational in whatever sense is deemed appropriate. It includes *historical* data' (1994: 140).

16 The institution of (customary) property rights is already presumed in making the exchange of private goods the object of analysis, even if its third party enforcement is not. A still more primitive domain than the economic exchange domain could be that of the common resource pool where private ownership is

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

of only two agents who can potentially repeat the transactions over time. As Hicks once noted, even the simplest exchange of this type is essentially a 'contract': making the agreement to exchange, delivery one way, and delivery the other. 'Trading is trading in promises' (Hicks: 1969). How can this bilateral promise be assured and fulfilled? This is essentially the problem of contract enforcement. As is well known, the simplest institution that can arise in response to this problem is the bilateral reputation mechanism, called trust: the mutual belief that default on a contract will be penalized by the other party who will refuse to exchange in the future. However, transactions may not necessarily be repeated between a fixed pair of players but among a large number of (mutually unknown) players, and in this case an effective self-sustainable reputation mechanism may not be feasible, because the information necessary for identifying and penalizing past cheaters may become increasingly difficult to disseminate. A solution to this problem can be facilitated by the emergence of a third party, acting as mediator in information dissemination and/or enforcing contracts across multiple exchange domains, leading to the notion of linked games to be discussed in the next section. There can be a variety of such third parties, ranging from Law Merchants (Milgrom, North, and Weingast, 1990) to the state (Olson, 1993) to digital rights management (e.g., auction websites, certification authorities), and not necessarily limited to the state (see Aoki, 2001: Chapter 4; Greif, 2005).

- *The organizational exchange domain and organizational field.* The organization may be a player of the game in an economic–exchange domain.[17] At the same time, the organization itself may be regarded as emerging as an institution in the domain of work collaboration. Indeed, H. Simon noted that 'the term *organization* refers to the complex pattern of communications and other relations in a group of human beings. This pattern provides to each member of the group much of the information, assumptions, goals, and attitudes that enter into his decisions, and provides him also with a set of stable and comprehensible expectations as to what the other members of the group are doing and how they will react to what he says and does' (1957: xvi). This definition has a striking similarity to our definition of an institution. A peer-team may emerge as a most simple form of collaboration, but, as a seminal contribution by Alchian and Demsetz (1972) argued, the presence of a third party may soon become essential for more complex work collaboration. The challenge for such an asymmetric organizational form is not limited to the need to monitor moral hazard problems, but it also involves the question of how to process, assimilate/distribute, and use information

not defined yet. The problem unique to this type of domain is the 'tragedy of commons' due to the over-consumption of the goods and/or the shirking of efforts to sustain them. If it is technologically feasible to exclude particular players from the domain, then self-sustainable institutions such as the establishment of customary property rights (Demsetz, 1967), collective norms for consumption, and efforts of maintenance (Ostrom, 1990) and the like, may cope with the problem.

17 North (1990, 2005) emphatically argues that the organization is a player of the game, but not an institution defined as the rules of the game. See note 11 above. To reconcile the two views (the organization as a player vs. an institution), see Hodgson (2006).

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

collectively. In this regard, organizational architecture as an information system could take on various proto-type forms, not limited to a simple hierarchy but including those combining modes of information-sharing and information-encapsulation among units, vertically as well as horizontally, in a distinct manner (Aoki, 2001: Chapter 4). When a certain mode of organizational architecture arises and becomes institutionalized, we may refer to it as an organizational convention.[18] Theoretically, diverse organizational conventions may be seen to arise as multiple evolutionary equilibria in the domain called the 'organizational field' (DiMaggio and Powell, 1983). In this domain, organizations are formed through the matching of agents that have developed certain cognitive orientations, thus leading to distinct clustering of collective cognitive frameworks (Aoki, 2001: Chapter 5).

The two prototype domains introduced above have traditionally been the objects of study in economics, business economics, and organization theory, while political and social factors have been taken as the given environments in those analysis. However, recently there has been a growing awareness among social scientists that there are actually important interactions between economic domains, on the one hand, and social and political factors, on the other. In order to capture these interactions and possible institutional linkages between them, we first conceptualize the following two prototype domains:

- *The politcal-exchange domain.* Let us presuppose that this domain in its prototype is composed of two types of agents: the government and multiple private agents. This asymmetric structure is somewhat similar to that of the prototype organization emerging in the collaborative work domain. It is different, however, in that in the organization the members have the option to participate or not, but in the politcal-exchange domain the exit option is not open to private agents. The government can provide public goods to the private agents (the protection of property rights, rights to live, collective security, and so on) in exchange for the extraction of costs in the form of taxes, issuing of money, etc. But the fact that the government has such power may also imply that it may have the power to transgress the various rights of the private agents (the so-called 'fundamental dilemma of political economy' due to Weingast, 1997). The private agents may respond by supporting/resisting/submitting to the government's choice (protect/transgress) with/without mutual coordination among themselves. Even in this simple game structure a variety of different equilibria can arise, depending on the ways in which coalitions between the government and particular private agents, as well as those among the private agents, are formed. These equilibria can be identified as institutions of the 'state' (Aoki, 2001: Chapter 6).[19] The English words 'stable', 'state', and 'institution' are all said to have been derived from the same Latin word 'status' (standing condition). Thus, it seems to make sense semantically to identify the 'state'

18 Sociological organizational theorists, like DiMaggio and Powell (1983), refer to the phenomenon as organizational isomorphism.

19 Greif (2005) refers to equilibria in the political domain as 'coercion-constraining institutions'.

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

as a political institution with a stable equilibrium in the politcal-exchange domain. Aoki (2001: Chapter 6) derives a variety of prototype states, such as the liberal democratic state, corporatist state, developmental state, bureau-pluralistic bargaining state, and predatory state as multiple equilibria from a structure-wise isomorphic game-form.

- *The social exchange domain.* This domain may be conceptualized as the one in which social symbols (languages, rituals, gestures, gifts, etc.) directly affect the payoffs of players, such as esteem, emotional rejection, sympathy, benign neglect, and so on, and are unilaterally delivered and/or exchanged with 'unspecified obligations to reciprocate', sometimes accompanied by gift-giving (Blau, 1964). Institutions that arise in this type of domain are identifiable with social customs and norms enforced by the threat of social ostracism from the domain, gradational rankings (stratification) of prestige/social status among the agents, etc. (See Coleman, 1990; Aoki, 2001: Chapter 8). Norms are taken as exogenous rules for the 'economic exchange' game in the North–Williamson framework, but their production and reproduction may be susceptible to game-theoretic analysis, which we will discuss in greater depth shortly.

We have identified four prototypes of domains. However, an equilibrium of the game, and consequently an institution, may neither arise nor be sustained in a single domain independently of other domains. Also, institutions arising in different domains may not be hierarchically aligned in such a way that social norms precede a political institution, while decisions made in the context of a political institution determine the game forms in economic and organizational domains. Rather, institutions may arise encompassing different domains, and/or institutions in different domains may co-evolve through complementary relationships, leading to a complex structure of over-all institutional arrangements. We now turn to this structure.

4. Analyzing institutional linkages across the domains

In this section we discuss possible institutional inter-linkages across domains using game-theoretic apparatus. Indeed, one of the important advantages of the game-theoretic approach to institutions lies in the possibility that the intuitively appealing and plausible notions of institutional interdependencies, coherency and path-dependence are made analytically tractable rather than being presented as an *ad hoc* presumption. From the game-theoretic perspective, there can be two types of equilibrium linkages and thus institutional linkages: linked games and strategic complementarities. They expand equilibrium possibilities as well as make multiple, non-Pareto-optimum equilibria possible, thus leading to a diversity of over-all institutional arrangements.

Linked games: Games are 'linked' if one or more players coordinate his/her own choices of strategies across more than one domain so that the sum of payoffs is greater than would be possible from playing separately in each of

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

these domains. The reason for this is that by doing so these players can benefit from externalities, such that possible gains in one domain can be transferred to another to sustain some strategic profile that would not be profitable in isolation. Thus, equilibrium possibilities can be expanded. For example, suppose there is a common resource pool domain where agents who misuse the commons (abuse, shirking of maintenance efforts, and the like) cannot be excluded from using them for technological reasons so that the reputation mechanism cannot be implemented. However, if the members of the common resource pool domain all belong to a social exchange domain where large social surpluses can be created by cooperation in rituals, festivities, assistance in times of private hardships, and the like, then misbehavior in the common resource pool domain can be credibly punished by ostracism in the social exchange domain (Aoki, 2001: Chapter 2.2). This is an instance where the reputation mechanism may become self-enforcing by linking games in different domains, even if players are short-sighted and/or cannot be excluded from one of those domains. Essentially the same mechanisms are found in a variety of situations, for example the quality of the natural environment is protected by the community of citizens who share a same values; open-source software is developed by free contributions from individual engineers who aim to enhance their professional reputations, etc.

This type of linkage mechanism corresponds to the sociological notion of 'social embededness' due to Granovetter (1986). It is particularly worth noting that the author made an explicit reference to the endogeneity of norms as well as to their strategic nature from a sociological perspective. He argued that 'agents in markets and organizations in the modern society generate trust and discourage malfeasance by being embedded in concrete personal relations and structures (networks)'. However, the norms and values are not a one-time influence but an ongoing process, continuously needing to be constructed and reconstructed through interactions. In other words, values and norms may be perceived as exogenously received by individuals, but actually they are endogenously shaped by them 'in part for *their own strategic reasons*' (p. 57; italics by the present author). As mentioned, linking games generally expands equilibrium possibilities. Thus, the commonsense notion, that differences in social norms lead to diversity of institutional arrangements, does indeed have a logical foundation.

Another type of linked games of institutional relevance can be found in the *bundling* of multiple similar or disparate domains. For example, suppose an employer cannot credibly threaten a single worker with dismissal in the event of shirking because the costs of recruitment and training of a replacement worker would be more than the gains from honest work and efficiency. Under such situation a likely outcome could be the repetition of one-time Nash, i.e. the exchange of shirking with low wage. However, suppose that the employer bundles multiple employment contracts that include the threat of replacement of a worker who shirks. If there is no coordination among the workers to shirk together, then the threat could become credible because the benefits of

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

the honest and efficient work from the other workers can exceed the cost of replacement of a single worker (Murdock, 1996). From this perspective, the emergence of the factory system may be regarded as an institutional response to the problem of worker discipline, rather than to technological exigencies. The economic exchange domains thus bundled are then transformed into a prototype of hierarchy in an organizational domain.

A somewhat similar example is found in the financing of multiple entrepreneurial projects of a similar type by a venture capitalist, with the arrangement of not continuing to finance those projects judged to be performing unsatisfactorily relative to others. In spite of multiple financing costs, bundling of multiple contracts may be beneficial to the financier, not only because they can broaden future options in the presence of high developmental uncertainty (Baldwin and Clark, 2002), but also because it can elicit higher efforts from entrepreneurs through the threat of termination, or equivalently through the enhancement of probability of continued finance due to better performance. But in order for this to be true, the identification of badly performing projects must be made precisely, while the reward from a successful project must be very high. The possibilities of the option value and externalities created by the tournament-like competition are considered two fundamental institutional features of entrepreneurial competition as observed in Silicon Valley (Aoki and Takizawa, 2002).

Examples in the above paragraph are about bundling by a single player internal to each of the bundled domains of a similar type (i.e., the employer, the financier).[20] Bundling may also be institutionalized by a third party external to domains. Suppose, for example, that the reputation mechanism cannot sustain honest exchanges (mutual contract compliances) between two anonymous traders because they are not expected to meet again. However, if multiple domains of this sort are bundled with an intermediary who can disseminate information regarding the past contractual compliances of the agents, the two-person reputation mechanism can be effectively replicated, provided that honest information processing and dissemination by the third party can be motivated by his/her own reputation concerns. The Law Merchants (Milgrom, North, and Weingast, 1990), credit bureaus, escrow services, on-line certification authorities, and auction sites are examples of such third parties. It is important to note that third parties in bundling are by themselves strategic players and they should be treated as such in an analysis of institutionalization. Still more complex linked games exist between domains of different types. As suggested, the organizational exchange domain may tend to generate a particular mode of information-systemic architecture internally; but it may not be sustained in isolation. The members of the organization may also be active players in other domains (such as financial, labor, and politcal-exchange domains) and coordinate their own

20 An example of bundling of different domains by a single player internal to all of them is also found in 'linked contracts' in land-leasing and financing in the developing economy. See Bardhan (1977).

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

internal and external strategies. As a result, a complex institutional structure, known as corporate governance, may evolve across those domains (Aoki, 2001: Chapters 11–14).

Institutional complementarities: In linked games each agent or a particular agent coordinates his/her own strategic choices across domains and generates a single institution (equilibrium) therein. Alternatively, we can conceive of the possibility that, even if agents may not consciously coordinate their own choices across domains, they regard an institution in another domain as a parameter and choose strategies in their own domains accordingly, and vice-versa. In such situations, institutions evolving in each of these domains may become interdependent and mutually reinforcing. This intuition can be game-theoretically warranted. Suppose simply that x' and x'' are two alternative institutions (equilibrium outcomes) in domain $\boldsymbol{X}$, while z' and z'' are two alternative institutions in domain $\boldsymbol{Z}$. Suppose that the pay-off difference $U(x') - U(x'')$ increases for all the players in domain $\boldsymbol{X}$ (they do not need to have the same pay-off function), when z' rather than z'' prevails in domain $\boldsymbol{Z}$. By the same token, suppose that the pay-off difference $V(z') - V(z'')$ increases for all the players in domain $\boldsymbol{Z}$ (they may be partially or totally overlapped with those in domain $\boldsymbol{X}$), when x' rather than x'' prevails in $\boldsymbol{X}$. Then the games in $\boldsymbol{X}$ and $\boldsymbol{Z}$ are said to be super-modular, and x' and z' (alt. $x''$and z'') are said to complement each other. If the super-modular condition holds, then an equilibrium combination, and a viable over-all institutional arrangement, can be either (x', z') or (x'', z'') (Topkis, 1978; Milgrom and Roberts, 1990; Aoki, 1994, 2001: Chapter 8). Further, even if one of them is less efficient in terms of Pareto-ranking, it may still prevail as an equilibrium, once it is achieved (Aoki, 2001: 225–9).

This is a powerful and useful analytical tool for institutional analysis. First, as just mentioned, it explains why there can be a variety of over-all institutional arrangements across economies, even if the economies face the same types of domain characteristics (such as technologies or common markets connecting them), as well as why a sub-optimal over-all institutional arrangement can persist in some economies while a better institutional arrangement is viable in others. Second, institutional complementarities are not necessarily conditional on consensus among agents in domain $\boldsymbol{X}$ regarding the absolute ranking of x' versus x'' (i.e., it is not required that $U(x') - U(x'') > 0$). Only a weaker agreement in the direction of change in their pay-offs associated with a parametric change in z matter. Thus, an over-all institutional arrangement may emerge and become sustainable even if there is a conflict of interests among the agents about the absolute preference for a component institution in isolation. For example, suppose that the manager prefers retaining the exclusive rights of managerial prerogatives over works within his/her own firm, but wages are set by an industrial collective bargaining agreement and its results are enforced over the entire industry by the corporatist state, as in Germany, rather

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

than through individual/firm-specific contracts. Then the manager's dislike of workers' participating in work control through co-determination and/or the works council may be mitigated, because yielding partial control of rights to the workers may substitute for the missing pecuniary incentives. Then the work council in the organizational-exchange domain and the corporatist state in the political-economic exchange domain would become institutionally complementary to each other, while the combination of exclusive managerial prerogatives and the liberal democratic state (in the sense of government non-interference in private employment wage contracts) would become another possibility (Aoki, 2001: Chapter 11.2).

5. How do bounded rational agents change institutions

Even if the nature of the over-all institutional arrangements can be understood in equilibrium terms, it does not mean that institutions will not change. Change will occur when there is a substantial equilibrium shift. However, as we have conceptualized an institution as a summary representation (i.e., common knowledge) of the salient features of equilibrium, gradual changes in equilibrium as a passive response to continual changes in the parameters of the game-form may not immediately be reflected as an institutional change.[21] In that sense, institutions can be robust and inertial even if the environments of the game, as well as individual agents' choices, change to a certain extent. On the other hand, institutional changes, i.e. large changes in agents' knowledge, are not to be explained or characterized as having the so-called 'sunspot equilibria' property either: those induced by purely 'extrinsic uncertainties' leaving economic fundamentals intact (Prescott and Shell, 2002). Large exogenous changes in initial endowments are sometimes cited as an instance of such extrinsic uncertainties, but implicit in the distribution of initial endowments is an institution of property rights so that its change needs to be endogenously and/or explicitly explained. According to our conceptualization, a change in institution is to be characterized by a quantum shift in the equilibrium constellation of agents' strategies such as to generate, as well as be induced by, substantive changes in agents' shared behavioral beliefs. Then, when and how does such a quantum change in equilibrium occur? In this section, we consider the generic nature of institutional change. We will argue that aspects of the agents' bounded rationality play essential roles in this process (otherwise, equilibrium change will be gradual or only have sun-spot properties.) Some basic mechanisms of institutional change involving multiple domains are discussed in more concrete settings in the next section.

21 Technically, even if information partitions of the space of strategic choices by individual agents change marginally, the meet of information sets (common knowledge) may not be altered, or change at most marginally.

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

First, suppose there is a fairly stable situation of plays of the game, i.e. one in which an institution as the shared behavioral beliefs is unchanged while the game is being played. Namely, although the objective game-form may be continually changing in terms of the composition of interacting agents, their activated choice sets, and the parameters of the consequence function (formal laws, technologies, environments of the game, etc.), the agents only marginally adjust their action choices in the belief that the over-all characteristics of the ways in which the game is being played will remain the same and their experiences do not refute their beliefs. Also 'mutant' players, i.e. players satisfied with sub-optimal behavior, may exist who, mistakenly or intentionally, deviate from equilibrium play. But as long as such drifts remain in a sufficiently close neighborhood of equilibrium (technically, in the basin of attraction of evolutionary equilibrium), the summary representation of equilibrium common to all the players, thus an institution, will remain intact.

Recall that the situation in which institutions are stabilized is not characterized by agents' complete knowledge regarding the game structure, i.e. the game-form and other agents' preference structures. For each possible choice from the activated menu of his/her own action choices, each agent can expect a certain consequence with the help of institutions summarizing others' equilibrium responses, in addition to localized knowledge of the game-form intimately relevant to his/her own pay-offs. Then (s)he may make a choice that would produce the highest satisfaction to him/her within this institution-induced, individualized information structure. In this process, even a large portion of the set of objectively feasible action choices may be automatically removed from active consideration for choice. The individually tailored, truncated game-forms complemented with the common knowledge of institutions may be referred to as the *institution-induced, individual game models*.[22] They are simplified, personal versions of the objective game structure (game form) as well as the actual state of game, but can be more informative to individual agents in the sense of specifying only information sufficient for individual choices in the currently prevailing situation. They are not exogenous data of the game structure available for agents who are capable of super-rational calculations to achieve Nash equilibrium. They already contain salient features of the endogenous outcome of the game, but in turn serve as a support for continual reproduction of them.

Now consider changes in the objective game-form explicit. They may be changes in the activated sets of individual choices due to the accumulation and

22 The institution-induced game-model is roughly a renamed version of what I called the 'subjective-game model' in Aoki (2001: Chapter 8). But as the term 'subjective game' is used in other contexts by Kalai and Lehrer (1995), I adopted the new term. Note that it roughly corresponds to the notion of 'mental models' in the induction theory of Holland *et al.* (1986). They conceive of mental models as models of the problem space 'that cognitive systems construct, and then mentally "run" or manipulate to produce expectations about the environment' (ibid.: 12). But our conceptual framework emphasizes the inter-linkage of those models among agents through their common knowledge of events via institutions.

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

development of skills, learning, innovation-induced new action possibilities, and so on. Or they may be changes in technological and environmental conditions that result in different physical consequences for the same action choices. They may be new laws or fiats which are enacted as consequences of the game in the political domain, but appear as exogenous changes in the parameters of the consequence functions in other domains. Note that, although all these changes have been conveniently conceptualized as parametric shifts in the game-form so far, they may be actually regarded as endogenous in the sense of being induced through the repeated plays of the game and thus patterned by them. Therefore Greif (2006) aptly calls the parameters of the game-form the 'quasi-parameters' (although he does not employ the concept of the game-form explicitly).

Also, as the patterns of play structured by institutions repeat themselves, the cumulative consequences of repeated plays may start to generate internal inconsistencies and/or endanger their sustained compatibility with the environments. This may be evidenced by the depletion of the natural environment and stock of natural resources, the obsolescence of some skills, the revelation of attractiveness of under-utilized skills, increasing frictions within the institutional orders of external domains (including foreign markets), and the like. In such situations, mutant strategies that have been suboptimal thereto may exhibit an increasing viability. Also, searches for, and experiments with, new types of strategies may be initiated. In other words, some agents may try to expand their sets of activated choices with greater energy. When deviations from the existing patterns of playing occur beyond a certain threshold, hither-to-held individual perceptions about the ways in which the game is, and should be, played become problematic and may not be taken or granted any more. Shared behavioral beliefs become de-stabilized, signaling the crisis of institutions.

In such crisis situations, competition can arise between searches for new types of playing on the one hand and efforts to preserve the existing pattern of playing on the other. Or, even if an overt conflict does not arise, agents may not be in agreement over the direction to which the play of the game should be changed. In either case, mutual bonds through institutional stability can be lost. If such a crisis situation persists, political contests may become harsher in the political-exchange domain, economic depression may hit a certain segment of the economic-exchange domain, and a sense of anomie may spread in the social-exchange domain. In order to restore political stability, economic development opportunities and social bonds, institutions of shared behavioral beliefs need to be reconstructed. How can this be accomplished? Especially, given the many unforeseen alternatives, how can one particular institution out of the many possible eventually be selected?

Shared beliefs as social constructs may sometimes be generated spontaneously or in a decentralized manner, but often by the creation of a 'focal point' (Schelling, 1960) around which new strategic profiles and behavioral beliefs can coalesce. The well-know problem of common knowledge (i.e., how the

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

players acquire shared behavioral beliefs) can be resolved by reckoning that the salient features of equilibrium that form an institution may have precedents in the *ex ante* beliefs of some agents during the process of institutional crisis and reconstruction. For example, in the organizational domain, entrepreneurial agents may be engaged in creating a novel mode of bundling various productive activities with their own predictive beliefs in its viability and profitability, emulated by the others later. Alternatively, it may be a charismatic leader or a political organization in the political-exchange domain that advocates a particular normative belief and associated policies. Here, the entrepreneurial 'predictive' beliefs and political 'normative' beliefs can be distinguished from endogenous, shared 'behavioral' beliefs that constitute an institution because the former are beliefs held *ex ante* or are advocated only by a partial set of agents prior to the formation of an equilibrium. The competition among them in impacting the patterns of agents' action choices characterizes the transitional process. However, it may not be a straightforward process that a particular 'belief' becomes a focal point for the convergence of behavioral expectations and thus implied patterns of behavioral choice become prevalent and equilibrated. The process may be highly complex, depending on how learning, emulation, adaptation, reinforcement, resistance, and inertia interact across economic, political, organizational, and social exchange domains. For example, even though entrepreneurs experiment on a new model of organizational architecture that may fit a new technology, it may not become viable, unless complementary changes occur in labor and financial markets. Even if a change in the law is introduced as a consequence of a political leader's active initiatives in the political domain, the ways in which it impacts other domains may not be straightforward and an unintended outcome may result. An existing pattern of social norms may facilitate or deter the emergence of new patterns of behavior in other domains by embedding them. In next section we will discuss these interactive mechanisms that may condition agents' search for new action choices to converge in a certain direction of institutional change.

The transitional process converges when and only when: (i) with the help of a system of predictive and/or normative beliefs that have guided agents' learning, a new pattern of plays of the game emerges and becomes collectively recognized as the way the game is now being played; and (ii) agents' new action choices based on such expectations generate satisfactory pay-offs to them without a big surprise. That part of behavioral expectations common to all the agents then emerges as a new institution. The transitional juncture of institutional evolution comes to an end and another spell of relative institutional stability will be initiated. As a remark in the previous section regarding institutional complementarities suggests, even at the time of institutional emergence, a prevailed advocacy need not be unanimously accepted as the best social choice so that social harmony will be achieved. Disagreement may remain among agents, but once a pattern of action choices consistent with the advocacy becomes

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

prevalent, the agents will find themselves better-off acting in compliance with it.[23]

6. Mechanisms of cross-domain institutional change

Toward the end of the last section, we hinted at a few possible mechanisms of institutional change. In this section we will try to make explicit the nature by which they induce a qualitative equilibrium-shift and examine the ways in which they are susceptible to analytical understanding. Specifically, we examine the dynamic counterparts of the linked games and institutional complementarities, as discussed in Section 4, i.e.: (1) Schumpeterian innovation in bundling, (2) social embeddedness encompassing sequentially arising domains, and (3) dynamic institutional complementarities, or the co-evolution of institutions across domains. These mechanisms may operate in such a way that the shift of equilibrium moves in a certain direction *vis-à-vis* the other possible many. For each of these, a few examples are given. They may appear anecdotal, but are susceptible to rigorous game-theoretic analysis in terms of context-specific game models.

Schumpeterian innovation in bundling: As suggested already, bundling can arise when there is a possibility of rents accruable to the bundling agent, either through new information or the externalities he or she creates. In the dynamic context a new type of bundling becomes viable when an old established type disintegrates because of an internal crisis, or, as is more likely, a new type destroys the rents accrued to the old type. Thus, the institutionalization of a new type of bundling may not necessarily be smooth. A vigorous thrust of Schumpeterian entrepreneurship may be required to pursue the creative destruction of old combinations (Schumpeter, 1911/1934).

One example is the emergence of the clustering of small entrepreneurial start-up firms in Silicon Valley and other high technology centers. They have emerged as a result of the unbundling of business activities integrated within the organizational architecture of market-dominant IBM. In traditional integrated firms such as IBM, GM, a comprehensive set of business activities, stretching from design to manufacturing and marketing, were bundled under the umbrella of a single corporate headquarters. Design activities themselves were modularized and organized in a hierarchical manner, starting from the centralized conceptual design of an integrated product system at a higher level, to an analytical design, to the detailed designs of modular parts, to a manufacturing process design, to the manufacturing of a pilot product, and, finally, to its improvement at a lower level. In contrast, Silicon Valley start-up firms tend to be specialized in modular

23 North argues that '[w]hen conflicting beliefs exist, the institutions will reflect the beliefs of those (past as well as present) in a position to effect their choices' (2005: 50).

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

product design and possibly in the production of pilot products targeted for particular niche markets. Other activities, such as large-scale manufacturing and marketing, are beyond their immediate scope. As already pointed out, these entrepreneurial start-up firms are engaged in tournament-like competition. The winner in design of a modular product is likely to be acquired by leading firms in markets of a broader range (such as Cisco Systems, Intel, and Microsoft), whose aim is to create a market-dominant, complex product-system. Thus, from the viewpoint of product-system innovation, the system integration of component technologies is evolutionarily realized by the *ex post* bundling of selected modular products developed by entrepreneurial firms, in contrast to the *ex ante* bundling of comprehensive design activities within the framework of centralized planning of a single integrated firm (e.g. the development of IBM/System360). Clearly, the former type of bundling, referred to as A&D (acquisition and development) in contrast to the traditional intra-organizational R&D (research and development), is more flexible in keeping alternative options open until the uncertainty involved in viable system-design is reduced. The cost of running this Schumpeterian-type process is the duplication of development efforts by multiple competing entrepreneurial firms before interim and *ex post* selection is complete. We have already pointed out some technological and institutional conditions under which the expected benefits (the option value and tournament value) from this new mode of organizational architecture become more than enough to compensate for the costs.

Overlapping social embeddedness: Let us imagine that the potentially possible action sets of agents may be large (possibly of infinite dimensions), but that they actually activate only small subsets (of finite dimensions) as a 'repertoire' of actual choices at any point in time (Dosi and Marengo, 1994). In general, agents can change their repertoire over time by adding new choice possibilities and deleting those that have become obsolete, in response to changes in their own physical and mental states shaped by training, experience, perception of the external world, etc. However, the speed of these changes can vary depending on the domain. The choice possibilities open to agents can change relatively slowly in the social-exchange domain, while those in the organizational exchange domain may change more flexiblly, because the organizational architecture is susceptible to design competition among entrepreneurs subject to constraints imposed by complementary institutions (such as labor and capital markets). Thus, it is possible that the same pattern of choice profiles in the social-exchange domain can become linked with different choice profiles in other domains. In other words, the same type of social norm can embed different types of domains over time in an overlapping manner.[24]

24 Interactions between slow-changing institutions and fast-changing institutions discussed by Roland (2005) can be analyzed in terms of over lapping linked games.

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

For example, consider the opening of exchange opportunities with outside merchants facing a rural community in the pre-capitalist period, where the social norm of cooperation in the management of the commons, such as the irrigation system, prevailed. Until the resurgence of institutional economics in the late twentieth century, various social science disciplines, including economics, sociology, and anthropology, tended to draw a sharp line between community relationships in the pre-capitalist economy and market relationships in the capitalist economy (e.g., Polanyi, 1944). But the presence of community norms as an institutional device to promote and sustain cooperation in the rural community can facilitate, or deter, its transition to a market economy, contingent on its prevailing historical conditions (Aoki and Hayami, 2001). In some cases, the presence of a community norm can serve as a transitory mechanism of contract enforcement in the underdevelopment of the rule of law governing the economic exchange domain. It does so by facilitating collective punishment of breaches of contracts by merchants from the outside, while restraining their own members' breaches through peer monitoring in order to preserve their collective reputation to outside markets.

Another example is the evolution of small firms clustering in the Italian industrial districts. They were begun by small enterprises in the garment industry that emerged after the highly integrated textile companies failed because of high wages and labor disputes in the 1960s. Skilled workers released from large companies were encouraged to establish their own enterprises, often by purchasing equipment from large companies that were closing (Barca *et al.*, 1999). The types of transactions and coordination that quickly developed among these firms, such as the reciprocity of subcontracting and the sharing of productivity-enhancing knowledge, would not have been feasible without mutual trust as an essential governance mechanism. Their development was possible because the transaction domain was embedded in a pre-existing social exchange domain. The new owners of these small firms had already invested a significant amount of social capital – social surplus derived from their individual reputations as members of the civic community and/or the labor organization that had confronted the old integrated companies.

Dynamic institutional complementarities: The concept of static institutional complementarities has a natural dynamic version formulated in the Momentum Theorem by Milgrom, Qian, and Roberts (1991). Liberally rephrased, one version of it holds that, even if the initial level of human competence in domain $\boldsymbol{X}$ that is conducive to the support of potential institution x' is low, the presence of complementary institutions in other domains may amplify the impact of a policy intended to induce x', and that once momentum is initiated x' may gradually evolve as a viable institution. Conversely, even if laws are introduced to induce institution x', the absence of complementary institutions and supporting

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

competence in this and other domains can make its realization difficult (Aoki, 2001: 267–269).

One example is the role played by the institutional infrastructure of contract implementation in Hong Kong in the transition of the Chinese economy to a market economy. It is now well-recognized that China's remarkable growth since the initiation of the reform and open policy in the late 1970s was largely driven by foreign direct investments and commodity exports. By 1977, however, the Chinese economy was virtually closed to the world market economy, while the domestic economy was almost completely governed by a command system. How could such an economy attract massive foreign investments in spite of the infancy of its institutional infrastructure for market exchange? How could it be expected to entertain export contracts in the absence of an effective rule of law governing domestic market exchanges? Actually large portions of capital inflow and commodity exports were mediated through Hong Kong, where the legal infrastructure of contract implementation and enforcement was relatively well developed and the associated competence of individuals in law, accounting, consulting, transportation logistics, and foreign languages was already in place or was relatively easy to recruit from abroad.[25] Once the Chinese economy became involved in exchange relationships with outside markets, it was able to gradually develop its own market competence through complementary reinforcement by the infrastructure in Hong Kong.

Another possible version of the momentum theorem is the following: Suppose that changes in the parameters of the game-form (e.g., the introduction of new public policy, a change in legal rule, the accumulation of competence, and so on) conducive to a possible alternative institution occurs in a domain, but its initial impact in isolation is still too small. Suppose further that similar parametric changes also occur in a complementary domain (i.e., one in which the endogenous strategic profile is complementary to that in the other). Then, even though an institutional change may not immediately occur in either domain, if parametric changes are sustained thereafter in both domains, their cumulative impact on endogenous strategic choices in respective domains, together with mutually reinforcing impacts of evolving strategic choices across the domains, can eventually lead to the co-evolution of new institutions in both domains. This mechanism is particularly interesting for examining how a change in the political game-form (formal rules in the polity) and/or its consequence (policy change) can impact institutional changes in other domains. The causation may not be unidirectional from the polity to the other domains, however, but can be bi-directional in that changes in policy, and more fundamentally a political

25 The proportion of China's exports via Hong Kong in the 1990s exceeded 40 percent of its total exports and that of capital imports reached 50 percent. See Shinohara (2003) for statistical compilation and analysis of the role of Hong Kong in the rapid development of the Chinese economy in the last two decades.

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

institutional change, can also be induced by, and co-evolve with, changes elsewhere.[26]

6. Concluding remarks

One of the major objectives of this paper was to propose a unified, analytical, and conceptual framework for understanding the roles of social, political, economic, and organizational factors, as well as the nature of their interdependencies in the process of institutional change. Different social science disciplines have tended to focus on only one, or a few, of those factors as endogenous in the process of institutional change, while viewing the others as exogenous data. In order to develop a more integrative approach, we borrowed the essence of the game-theoretic apparatus and modified it somewhat liberally.

In the beginning, we made a sharp distinction between the game-form (the exogenous rules of the game) and the equilibrium outcome of repeated strategic plays within it and identified institutions with the shared behavioral beliefs of the agents summarizing the salient features of the latter. Later, we tried to refine the notion of agents' bounded rationality further by introducing the concept of institution-induced, individual game-models: beliefs that individual agents form via their knowledge of institutions. An institutional change was seen to be triggered by a situation in which these individual game-models were no longer to be taken for granted because of quasi-endogenous changes in the objective game-form. The notion of an institution as a summary equilibrium representation and that of institution-induced, individual game-models were helpful for understanding the roles of public discourse; entrepreneurial, political, or charismatic leadership; and legacies of social norm in the process of institutional change in which bounded rational agents participate.

One of the reasons why we introduced the equilibrium notion of institutions was to evade an often employed practice in the literature of making a distinction between operational games (with an institution given) and rule-making, social-choice games (or meta-games). One serious problem associated with this approach is the infinite regression, i.e. how the rules for the latter games are determined. Instead, we conceptualized institutions as the common knowledge of 'salient features' of repeatedly played games without making a distinction between operational and rule-making institutions.

In order to understand how social, political, economic, and organizational factors interact in the formation, sustenance, and evolutionary changes of overall institutional arrangements, we explicitly considered interdependencies of

26 See Aoki (2007) for an illustration of such mechanism, in which the traditional Japanese institutional arrangements in corporate governance (the so-called relational contingent governance) and political governance (the so-called bureau-pluralistic bargaining state) are shown to be decaying in mutually reinforcing manner in the last decade.

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

equilibrium formation across the domains involving each of those factors as a principal characteristic. More specifically, we specified Schumpeterian bundling innovation, overlapping social embeddedness, and dynamic institutional complementarities as mechanisms of endogenous, interactive institutional changes. While the Schumpeterian bundling innovation can bring new design elements into the process by destroying old modes of bundling, the overlapping social embeddedness can impact it with elements of past legacies. Dynamic complementarities capture, among other things, the role of policy in the process of institutional change, which may operate only gradually and interactively with changes occurring elsewhere, sometimes with an unintended institutional outcome. Those mechanisms are conceptually distinct, but are likely to operate simultaneously and in an interactive manner in the process of institutional change. Innovation-oriented, Schumpeterian re-bundling has to take place in the milieu of complementary institutions/competence inherited from the past and embedded in persistent social norms. But, overlapping social embeddedness and institutional complementarities may not necessarily deter institutional innovation. On the contrary, the former can facilitate the emergence of a new domain by embedding it with the reproduction of time-honored patterns of social exchanges. The latter can create a momentum for the emergence of new over-all institutional arrangements once worn-out links between the parameters of the game-form and the endogenous strategic choices break down simultaneously across complementary domains. On the other hand, Schumpeterian bundling innovation may make a new bundling of business activities independently viable with the generation of its own endogenous rents, thus leading to the erosion of an old social norm.

Putting the three mechanisms in a unified perspective also suggests that the process of institutional change may not necessarily be hierarchically ordered in terms of causation such that social norms are the most robust, while institutional change in the economic domain is primarily initiated in the polity. Certainly, there are many instances in which policy and legislative actions in the politcal-exchange domain appear to trigger institutional changes in other domains. But even in this case, the process can be characterized as an instance of dynamic institutional complementarities in which policy change (a consequence of the play of the game in the polity domain) is initiated in response to the emergence of concomitant endogenous changes in other domains, while policies reinforce complementary interactions across those domains.

Thus, institutions in the past and in the future are mutually interlinked in a complex manner. Institutions generated endogenously at one point in time become exogenous constraints and/or enabling facilitators for further institutional dynamics in their own domains as well as beyond. There are spiral movements for the newly born to eventually become the established, on which basis further institutional evolution can be molded ad infinitum. No theory can escape from the problem of infinite regression. However, we have sought to focus on infinite regression directed toward the historical past rather than the logical construct of

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

the meta-game or the dominance of the political-exchange domain in institutional formation. In that sense, 'history matters' as well as 'institutions matter'.

References

Alchian, A. and H. Demsetz (1972), 'Production, Information Costs, and Economic Organization', *American Economic Review*, **62**: 777–795.

Amable, B. (2003), *The Diversity of Modern Capitalism*, Oxford: Oxford University Press.

Aoki, M. (1994), 'The Contingent Governance of Teams: An Analysis of Institutional Complementarity', *International Economic Review*, **35**: 657–676.

Aoki, M. (2001), *Toward a Comparative Institutional Analysis*, Cambridge, MA: MIT Press.

Aoki, M. (2007), 'Whither Japan's Corporate Governance?', to appear in M. Aoki, G. Jackson, and H. Miyajima (forthcoming), *Corporate Governance in Japan: Institutional Change and Organizational Diversity*, Oxford: Oxford University Press.

Aoki M. and Y. Hayami (eds) (2001), *Communities and Markets in Economic Development*, Oxford: Oxford University Press.

Aoki, M. and H. Takizawa (2002), 'Information, Incentives, and Option Value: The Silicon-Valley Model', *Journal of Comparative Economics*, **30**: 759–786.

Aumann, R. J. (1976), 'Agreeing to Disagree', *Annals of Statistics*, **4**: 1236–1239.

Aumann, R. J. (1987), 'Correlated Equilibrium as an Expression of Baynesian Rationality', *Econometrica*, **55**: 1–18.

Baldwin, C. Y. and K. B. Clark (2000), *Design Rules: The Power of Modularity*, **Vol. I**, Cambridge, MA: MIT Press.

Barca, F., K. Iwai, U. Pagano, and S. Trento (1999), 'The Divergence of Italian and Japanese Corporate Governance Models: The Role of Institutional Shocks', *Economic Systems*, **3**: 35–61.

Bardhan, P. (1977), 'Variations in Forms of Tenancy in a Peasant Economy', *Journal of Development Economics*, **4**: 105–118.

Berger, P. and T. Luckmann (1966), *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*, NY: Doubleday Anchor.

Binmore, K. (1994), *Playing Fair: Game theory and social Contract*, Cambridge MA: MIT Press.

Binmore, K. (2005), *Natural Justice*, New York: Oxford University Press.

Blau, P. (1964/1998), *Exchange and Power in Social Life*, with new introduction, Brunswick, NJ: Transaction Publishers.

Braudel, F. (1958/1997), 'History and the Social Sciences: The *Longue Duree*', in S. Mathews (tr.), *On History*, Chicago: University of Chicago Press, pp. 24–55.

Calvert, R. L. (1995), 'Rational Agents, Equilibrium, and Social Institutions', in J. Knight and I. Sened (eds), *Explaining Social Institutions*, Ann Arbor: University of Michigan Press, pp. 57–93.

Coleman, J. (1990), *Foundations of Social Theory*, Cambridge MA: Harvard University Press.

Demsetz, H. (1967), 'Towards a Theory of Property Rights', *American Economic Review*, **57**: 347–359.

Denzau, A. and D. North (1994), 'Shared Mental Models: Ideologies and Institutions', *Kyklos*, **47**: 3–31.

DiMaggio, P. and W. Powell (1983), 'The Iron Cage Revisited: Institutional Isomorphisms and Collective Rationality in Organizational Fields', *American Sociological Review*, **48**: 147–160. Revised and reprinted as Chapter 3 of Powell and DiMaggio (1991): 63–82.

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

Dixit, A. K. (2004), *Lawlessness and Economics*, Princeton, NJ: Princeton University Press.
Dosi, G. and L. Marengo (1994), 'Some Elements of an Evolutionary Theory of Organizational Competence', in R. W. England (ed.), *Evolutionary Concepts in Contemporary Economics*, Ann Arbor: University of Michigan Press, pp. 157–178.
Fagin, R., J. Y. Halpern, Y. Moses, and M. Y. Vardi (1995), *Reasoning about Knowledge*, Cambridge, MA: MIT Press.
Field, A. J. (1981), 'The Problem with Neoclassical Institutional Economics: A Critique with Special Reference to the North-Thomas Model of pre-1500 Europe', *Exploration in Economic History*, **18**: 174–198.
Field, A. J. (2001), *Altruistically Inclined? Evolutionary Theory, the Behavioral Sciences, and the Origin of Complex Social Organization*, Ann Arbor: University of Michigan Press.
Gibbard, A. (1973), 'Manipulation of Voting Schemes: A General Result', *Econometrica*, **41**: 587–601.
Gintis, H. (2000), 'Strong Reciprocity and Human Sociality', *Journal of Theoretical Biology*, **206**: 169–179.
Gintis, H. (2004), 'Modeling Preferences', Santa Fe Institute.
Granovetter, M. (1985), 'Economic Action and Social Structure: The Problem of Embeddedness', *American Journal of Sociology*, **91**: 480–510.
Greif, A. (1993), 'Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition', *American Economic Review*, **83**: 525–548.
Greif, A. (1997), 'Micro Theory and Recent Developments in the Study of Economic Institutions through Economic History', in D. Kreps and K. Wallis (eds), *Advances in Economics and Econometrics: Theory and Applications*, **Vol. 2**, New York: Cambridge University Press, pp. 79–113.
Greif, A. (2005), 'Commitment, Coercion and Markets: The Nature and Dynamics of Institutions Supporting Exchange', in C. Menard and M. M. Shirley (eds), *Handbook of New Institutional Economics*, Dordrecht: Springer, pp. 727–786
Greif, A. (2006), *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*, New York: Cambridge University Press.
Hayek, F. A. (1945), 'The Use of Knowledge in Society', *American Economic Review*, **35**: 519–530.
Hayek, F. A. (1976), *Law, Legislation and Liberty*, Vol. II: *The Mirage of Social Justice*, London: Routledge & Kegan Paul.
Hayek, F. A. (1988), *The Fatal Conceit: The Errors of Socialism*, Chicago: University of Chicago Press.
Hicks, J. (1969), *A Theory of Economic History*, Oxford: Oxford University Press.
Hodgson, G. M. (2006), 'What are Institutions?', *Journal of Economic Issues*, **40**: 1–25.
Holland, J., K. Holyoak, R. Nisbett, and P. Thagard (1989), Induction: *Processes of Inferences, Learning and Discovery*, Cambridge, MA: MIT Press.
Hurwicz, L. (1960), 'On the Dimensional Requirements for Non-wasteful Resource Allocation Systems', in K. J. Arrow, S. Karlin, and P. Suppes (eds), *The Mathematical Methods in the Social Sciences 1959*, Stanford: Stanford University Press.
Hurwicz, L. (1996), 'Institutions as Families of Game-forms', *Japanese Economic Review*, **47**: 13–132.
Huizinga, J. (1949), *Homo Ludens. A Study of Play Element in Culture*, London: Routledge & Kegan Paul.
Ikegami, E. (2005), *Bonds of Civility: Aesthetic Networks and the Political Origins of Japanese Culture*, New York: Cambridge University Press.

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

Kalai, E. and E. Lehrer (1995), 'Subjective Games and Equilibria', *Games and Economic Behavior*, **61**: 1019–1045.
Koopmans, T. (1957), *Three Essays on the State of Economic Sciences*, New York: McGraw Hill.
Kripke, S. (1963), 'A Semantical Analysis of Modal Logic', *Acta Philosophica Fennica*, **24**: 83–94.
Levi, M. (1988), *Of Rule and Revenue*, Berkeley, CA: University of California Press.
Lewis, D. (1969), *Conventions, A Philosophical Study*, Cambridge, MA: Harvard university Press.
Meyer, J. W. and B. Rowan (1977), 'Institutionalized Organizations: Formal Structure as Myth and Ceremony', *American Journal of Sociology*, **83**: 340–63. Reprinted in W. W. Powel and P. DiMaggio (1991: 41–62).
Milgrom, P., D. North, and B. Weingast (1990), 'The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs', *Economics and Politics*, **2**: 1–23.
Milgrom, P., Y. Qian, and J. Roberts (1991), 'Complementarities, Momentum, and the Evolution of Modern Manufacturing', *American Economic Review*, **81** (Papers and Proceedings): 84–88.
Milgrom, P. and J. Roberts (1990), 'Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities', *Econometrica*, **59**: 1255–1277.
Murdock, K. (1996), 'The Role of Institutions in Enforcing Implicit Contracts: Analysis of Corporate and Government Policy', Ph.D. Dissertation, Stanford University.
Nelson, R. (1995), 'Recent Evolutionary Theorizing about Economic Change', *Journal of Economic Literature*, **33**: 48–90.
Nelson, R. and B. N. Sampat (2001), 'Making Sense of Institutions as a Factor Shaping Economic Performance', *Journal of Economic Behavior and Organization*, **44**(1): 31–54.
Nelson, R. and S. Winter (1982), *An Evolutionary Theory of Economic Change*, Cambridge, MA: Harvard University Press.
North, D. (1990), *Institutions, Institutional Change and Economic Performance*, Cambridge, UK and New York: Cambridge University Press.
North, D. (2005), *Understanding the Process of Economic Change*, Princeton, NJ: Princeton University Press.
Olson, M. (1993), 'Dictatorship, Democracy, and Development', *American Political Science Review*, **83**: 567–876.
Osborne, M. J. and A. Rubinstein (1994), *A Course in Game Theory*, Cambridge, MA: MIT Press.
Ostrom, E. (1990), *Governing the Commons*, Cambridge, UK: Cambridge University Press.
Polanyi, K. (1944), *The Great Transformation*, New York: Farrar Rinehart.
Powell P. P. and P. DiMaggio (eds) (1991), *The New Institutionalism in Organizational Analysis*, Chicago: University of Chicago Press.
Prescott, E. C. and K. Shell (2002), 'Introduction to Sunspots and Lotteries', *Journal of Economic Theory*, **107**: 1–10.
Roland, G. (2005), 'Understanding Institutional Change: Fast-moving and Slow-moving Institutions', http://emlab.berkeley.edu/users/groland/pubs/gr3.pdf.
Schelling, T. (1960), *Strategy of Conflicts*, Cambridge, MA: Harvard University Press.
Schotter, A. (1981), *The Economic Theory of Social Institutions*. Cambridge, UK and New York: Cambridge University Press.

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press

Schumpeter, J. (1912/1934), *The Theory of Economic Development*, Cambridge, MA: Harvard University Press; English translation of *Theorie der Wirtschaftlichen Entwicklung* by R. Opie, 1912.
Scott, W. R. (1955), *Institutions and Organizations*, Thousand Oaks: Sage Publications.
Searle, J. R. (1995), *The Construction of Social Reality*, London: Allen Lane.
Searle, J. R. (2005), "What is an Institution?" *Journal of Institutional Economics*, **1**: 1–22.
Shinohara, M. (2003), *The Mega-growth of the Chinese Economy and Hong Kong* (in Japanese), Tokyo: Keiso-Shobo.
Simon, H. A. (1997), Administrative Behavior, 2nd edn, NY: Free Press.
Smith, A. (1759/1976), *The Theory of Moral Sentiments*, Oxford: Oxford University Press.
Sugden, R. (1986), *The Economics of Rights, Cooperation and Welfare*, Oxford: Basil Blackwell.
Topkis, D. (1978), 'Minimizing a Submodular Function on a Lattice', *Operations Research*, **26**: 305–321.
Weingast, B. (1997), 'The Political Foundations of Democracy and the Rule of Law', *American Political Science Review*, **91**: 245–263
Williamson, O. E. (2000), 'The New Institutional Economics: Taking Stock, Looking Ahead', *Journal of Economic Literature*, **38**: 595–613.
World Bank (2002), *World Development Report 2002: Building Institutions for Markets*, New York: Oxford University Press.
Young, H.P. (1998), *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, Princeton, NJ: Princeton University Press.

https://doi.org/10.1017/S1744137406000531 Published online by Cambridge University Press