Journal of Institutional Economics (2012), 8: 3, 327–349
© *Millennium Economics Ltd 2012* doi:10.1017/S1744137412000045
First published online 12 April 2012

Linking institutions to economic performance: the role of macro-structures in micro-explanations

CYRIL HÉDOIN*
Department of Economics, University of Reims Champagne-Ardenne, 57B Rue Pierre Taittinger, 51100 Reims, France

Abstract: Economists have been investigating the link between institutions and economic performance for several years. While econometric studies of this link have flourished, they are of limited use in understanding the causal mechanisms making some institutions responsible for economic performance. Several works using game theory and akin to the 'new institutional economics' have entertained the goal of developing micro-explanations of the institutions–performance link. Because game theory focuses on individuals' actions and beliefs, game-theoretic studies of institutions are thought to oppose more 'structuralist' explanations that downplay the role of individual agents and put more emphasis on the importance of social (or 'macro') structures. This paper demonstrates that this claim is misconceived, as the micro-explanations produced by game-theoretic models must assume already existing macro-structures. Institutions produce downward effects, shaping each agent's action. Moreover, in a game-theoretic framework, macro-structures are constitutive of individual agency since, without them, agents would often be unable to choose. I illustrate this claim with the example of Avner Greif's study of the role of cultural beliefs in the economic organisation of medieval societies.

1. Introduction

Economists are now all convinced that institutions are a major factor in shaping economic outcomes. However, while the profession has demonstrated a growing interest in measuring quantitatively the impact of economic and political institutions on several economic variables such as growth, unemployment or financial stability, econometric studies still fail to provide a full answer to important questions such as 'do institutions cause growth?' (Glaeser *et al.*, 2004). This statement has provided one of the motivations for the development of a complementary approach in modern institutional economics. It seeks to

*Email: cyril.hedoin@univ-reims.fr

This article was presented at the 'Performance and Institutions' conference, held in Reims, France, on 17 and 18 November 2011. I thank Geoffrey Hodgson for his comments on the previous version of this article. I also thank the two anonymous referees for their valuable remarks which have help to improve the overall quality of the paper. Finally, I am grateful to Sophie Delabruyère for her help to improve the quality of the English. All remaining errors and omissions are mine.

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

uncover the causal relationships between institutions and economic outcomes. While several frameworks have been developed, the main one finds its roots in the 'new institutional economics' (NIE). Despite the diversity of the NIE, at least two characteristics are always present when the problem at stake is the link between institutions and economic outcomes. First, theoretical analysis is always embedded in some form of historical narrative or complemented with empirical case studies. Second, the analysis is always committed in providing an explanation which focuses on the *micro-foundations* of the link between institutions and economic outcomes. Hence, the NIE typically seeks to provide *micro-explanations* of the working of institutions. Works by authors such as Masahiko Aoki (2001), Robert Bates *et al.* (1998), Avner Greif (2006) and Douglass North (1990) fall into that category.

The search for micro-explanations leads one to make use of the main tools of micro-economics, namely rational choice theory and game theory. Hence, major works in recent comparative institutional economics all make heavy use of game-theoretic models and reasoning.[1] The propensity of economists to use game theory to study institutions is partially explained by the fact that economics is committed to methodological individualism and rejects any kind of 'structuralist' explanation loosely defined. Indeed, game theory models the reasoning and the action of individual agents explicitly. This paper investigates the way that game theory helps to link institutions and economic outcomes by producing micro-explanations. More precisely, we evaluate the implicit claim that game-theoretic analyses conform to the requirements of methodological individualism. Using the concept of *social mechanism*, we claim that game-theoretic analyses of institutions *must* include structural features to produce complete micro-explanations. Hence, such analyses do not conform to methodological individualism, at least in its classical version well summarised by Jon Elster (1982: 453) as 'the doctrine that all social phenomena (their structure and their change) are in principle explicable only in terms of individuals – their properties, goals and beliefs'.

Our argument deals with controversial issues in social sciences and takes place in a fairly large literature about various topics such as the role of reductionism and micro-foundations in socioeconomic explanations, the methodological individualism/holism issue or the problem of emergence in the social world. The points we are making are more or less closely related to those made by authors such as Kenneth Arrow (1994), Jack Vromen (2010) or, in sociology, Ronald Jepperson and John Meyer (2011). The argument of the paper is particularly reminiscent of Alexander J. Field's critique of 'neoclassical

1 This is particularly the case for Acemoglu and Robinson (2005), Aoki (2001) and Greif (2006). North (1990) does not make use of game theory or any kind of formalisation but he still uses to a large extent the logic and concepts of rational choice theory.

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

institutional economics' (Field, 1981, 1984).[2] Field convincingly demonstrates the impossibility surrounding the project of providing a truly micro-economic, rational choice theory-based explanation of institutions. We reach the same conclusion as his but while Field built his demonstration through the study of several examples of neoclassical models, we provide a more fundamental and general objection based on a game-theoretic framework. Like Field, we suggest that some institutions (or macro-structures) must be taken as exogenous in any micro-explanation; but, moreover, we show that this is so in principle because the epistemic properties of an institution are always related to a larger set of macro-structures.

The paper is organised as follows: in a second section, we set up our basic framework and define some essential concepts. Notably, we define the concept of social mechanism and introduce 'Coleman's diagram'. The third part argues that any game-theoretic explanation of the relationship between institutions and economic outcomes involves some kind of downward effect from non-individual entities toward individual agents. The fourth part goes further by indicating that game-theoretic explanations have to make appeal to macro-structures (i.e. non-individual entities) because macro-structures are constitutive of individual action. A fifth part illustrates the preceding argument by taking Greif's (2006) comparative study of the role of cultural beliefs as an illustration. Finally, a sixth part briefly concludes in arguing that any game-theoretic explanation of the link between institutions and economic outcomes is necessarily incomplete and must involve some kind of complementary narrative.

2. Institutions, game theory and social mechanisms

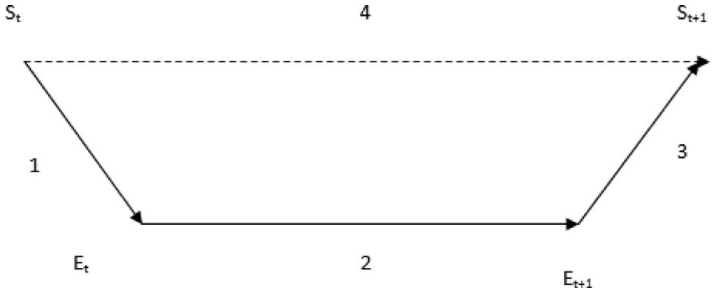
This section provides some definitions for the main concepts to be used below. Also, it sets up the framework we make use of to investigate the nature of the micro-explanations of the link between institutions and economic outcomes produced by game-theoretic analyses.

Surprisingly, the concept of 'institution' still does not have a widely accepted definition in economics. North's (1990) definition of institutions as 'the rules of the game' individuals have defined has had some influence but its shortcomings have been exposed many times (see Aoki, 2001; Amable, 2003; Greif, 2006; Hodgson, 2006). Geoffrey Hodgson (2006: 18) provided recently an encompassing definition: 'Institutions are systems of established and embedded social rules that structure social interactions'. He also gave a useful definition of the more general concept of *social structures* as including 'all sets of social relations, including the episodic and those without rules, as well as social institutions' (Hodgson, 2006: 17). Hence, an institution refers to any set of

2 We are grateful to a referee to have pointed out to us the relatedness of Field's thesis with the arguments made in this paper.

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

Figure 1. Coleman's diagram



social relations that are organised by one or several social rules. Greif (2006: 30) provides a similar definition: 'An institution is a system of rules, beliefs, norms, and organizations that together generate a regularity of (social) behavior.' This last definition can be divided into two components. First, any institution has an observable component through the regularity of behaviour. This can be seen as the *implication* of any institution and it has a straightforward game-theoretic interpretation: any institution must translate into an (Nash) equilibrium in a specific game. Second, the *nature* of an institution consists in observable or non-observable entities: (formal) rules, beliefs, (tacit) norms and organisations. All these entities have in common that they are irreducible to the individual level, since they necessarily involve social relations between several individuals.[3]

Econometric studies of the relationship between institutions and economic performance generally take the observable component (or its measurable implications) as the proxy for the institution under study. This is so because the entities constitutive of an institution (such as a convention or a social norm) are hard to measure or even to observe. However, the studies exploring the micro-foundations of the link institutions–economic performance will generally first identify the set of norms, rules and beliefs constituting the institution and then will work to demonstrate the causal link between this set and the economic outcome of interest. A useful graphical illustration of this idea can be made thanks to the so-called Coleman diagram first proposed by the sociologist James Coleman (1990).

Figure 1 depicts the essence of Coleman's idea about the relationships between patterns or regularities observed at the macro-level and elements or properties

3 This claim may be contested in the case of beliefs. Any person can form individual beliefs independently. However, the existence of institutions involves *social relations* where agents' beliefs, though sometimes formed independently, are dynamically interdependent. Indeed, social relations will lead generally to an adjustment in agents' beliefs. Aoki's (2001) own definition of institutions as an equilibrium between players' beliefs emphasises this fact. Moreover, as we argue in section 3, it seems more relevant to consider that, when an institution exists, agents' beliefs *do not* form independently.

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

responsible for them at the micro-level.[4] The diagram describes the process through which a system characterised by a state S_t at time t reaches a state S_{t+1} at time $t+1$. The core of Coleman's argument is that while social sciences are essentially committed to providing explanations of the relation between these two macro-states (relation 4 in the diagram), they cannot do so directly because relation 4 is not causal but only statistical. Hence, the argument goes, the explanation must engage with the micro-elements E_t and E_{t+1} responsible for the transition from state S_t to state S_{t+1}. In the social sciences, these micro-elements are constituted by individuals' actions and their interactions, even though on a more reductionist reading one can search for more fundamental (that is, psychological, biological or even physical) elements. Three relations must then be accounted for: the macro-to-micro relation (relation 1), which indicates how individuals' actions are influenced by the macro-state of the system; the micro-to-micro relation (relation 2), through which the micro-elements act and interact; and finally the micro-to-macro relation (relation 3), which accounts for the way the interactions between micro-elements aggregate into some new pattern or regularity at the macro-level.

Building on this basic framework, analytical sociology (Hedström, 2005; Hedström and Bearman, 2009a, Hedström and Swedberg, 1998a) has developed the fundamental concept of *social mechanisms* to formalise the processes behind relations 1, 2 and 3. Though several definitions are available (see Hedström and Bearman, 2009b: 6), a social mechanism can be basically characterised as referring to 'a constellation of entities and activities that are organised such that they regularly bring about a particular type of outcome (...) we explain an observed outcome by referring to the mechanism by which such outcomes are regularly brought about' (Hedström and Bearman, 2009b: 5). The respective definitions of a social mechanism and of an institution overlap in an interesting way since both include an emphasis on the regularity of the outcome. We observe an institution through the recurrent pattern it gives rise to. That means that an institution is a special case of social outcome that analytical sociology takes as the *explanandum* in the explanation through social mechanisms.

Meanwhile, Greif's definition of institutions clearly suggests a causal link between the pattern or the regularity of behaviour and the set of norms, rules, beliefs and organisations that gives rise to it. Hence, in Figure 1, the state S_t can be characterised by the set of norms, rules, beliefs and organisations while state S_{t+1} is defined by the regularity of behaviour. The methodological and theoretical

4 Coleman (1990) made use of this diagram in the context of a discussion of the famous Weberian thesis about the causal link between the religious ethic of Protestantism and the rise of modern capitalism. Specifically, Coleman argued that Weber failed to provide a satisfactorily account of the way Protestantism has orientated individuals' actions through norms and beliefs (relation 1 in the diagram) and how individuals' actions have generated together the rise of the institutions of modern capitalism (relation 3).

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

emphasis on the need for micro-explanations in the NIE therefore parallels analytical sociology's defence of explanations through social mechanisms. Still, institutions are peculiar in that they generate their own reproduction: the agents' behaviour reinforces the set of norms, rules and beliefs that produced it. The theoretical endeavour of institutional economics to link a set of institutional elements (i.e. of rules, norms and beliefs) to a regularity of behaviour (and its associated economic performance) leads to the search of a specific kind of social mechanism – which we will call an *institutional mechanism* – that frames together[5] relations 1, 2 and 3 to produce a stable outcome. Even a quick look at the literature in the NIE using game theory suggests that this is actually what is done (we substantiate this claim in the fifth section). Then, an institution appears to be a *set of rules, norms, beliefs and organisations that, through an institutional mechanism, generates its own reproduction and produces a specific social outcome*. We contend that this definition approximates in a satisfactory manner the way the concept is used in the NIE.

It is worth noting that this definition implies two related but different issues: first, how a given outcome is repeated through time – what we can call the 'institutionalisation process'; second, how the specific outcome is determined – the 'determination process'. The institutionalisation process is embedded in the game-theoretic framework we develop below; indeed, it is embedded in any game-theoretic account of institutions. The reason for this is that an institution is identified as an equilibrium in a given game and the very (mathematical) property of an equilibrium is to be a fixed point. Hence, the mechanism whereby the institution reproduces itself is explained by the very same institution. The determination issue refers to the problem of multiple equilibria: when there are multiple equilibria, several outcomes are amenable to an institutionalisation process. How and why a specific outcome is selected is a key question in any institutional game-theoretic analysis. We do not confront this issue directly (in particular because a general answer to it probably does not exist) but the fourth section indicates that the ability for an institution to solve the problem of the multiplicity of equilibria requires a set of macro-structures (in particular, other institutions) shared by the members of a community.

The rest of the paper investigates some of the implications of seeing institutional analysis in economics through the lens of the framework exposed

5 Hedström and Swedberg (1998b: 22,23) elaborate a typology of social mechanisms relatively to the relation they refer to in Coleman's diagram. They distinguish 'situational mechanisms' (relation 1), 'action-formation mechanisms' (relation 2) and 'transformational mechanisms' (relation 3). Despite its relevance, this typology is not useful here since game-theoretic explanations of institutions typically encompass the three types of mechanisms. Notably, the structure of the game as it is constructed by the analyst implies a situational mechanism. Action-formation mechanisms are captured by the various hypotheses made on the rationality of the players and on the formation of their beliefs. Finally, transformational mechanisms are generated by the type of solution concept adopted, since a solution concept allows one to 'predict' the outcome of the game.

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

above. We are essentially concerned here with the possibility of giving a truly micro-explanation of institutions. Hence, we are mainly interested with relation 1 in Coleman's diagram. The nature of this relation is controversial because it seems to imply some causal relationships that need to be specified. In the following, we argue that in the game-theoretic framework used by institutional economists, the macro-to-micro relation is primarily of a constitutive nature. This makes any pure micro-explanation of institutions impossible.

3. Institutions and (belief-mediated) downward effect in game-theoretic analysis

At first sight, the arrow depicting relation 1 in Coleman's diagram seems amenable to a purely causal interpretation such as 'the macro-properties of the state S_t causally affect the micro-elements E_t'. However, this reading may be too quick. Coleman himself and other rational-choice-oriented sociologists are selective in the mechanisms they use to illustrate this relation. Indeed, relation 1 is more akin to Mark Granovetter's (1985) concept of embeddedness according to which individual action always takes place in a network of relationships, occupations and opportunities. This network defines constraints on individual choice but is not a cause of it properly speaking. Interestingly, according to Lars Udehn (2001), Coleman's diagram in general is compatible with 'structural individualism' which he characterises as follows:

> In structural individualism, (...) actors are occupants of positions, and they enter relations that depend upon these positions. The situations they face are interdependent, or functional related, prior to any interaction. The result is a structural effect, as distinguished from a mere interaction effect (Udehn, 2001: 304).

Structural individualism is a loose version of methodological individualism recognising that individual action always takes place in a prior network of interactions and positions. Structural individualism does not contradict methodological individualism but adds that an essential component of social phenomena is the structured (i.e. non-random) interactions between agents. Crucially, it does not imply any ontological commitment regarding the nature of the relationship between the structure of interactions (the network) and individual action. Indeed, most analytical sociologists reject the claim that the social structure has its own ontological power (e.g. Hedström, 2005: 70–74). Hence, relation 1 in Coleman's diagram should not be necessarily read as 'social structure S *causes* individual action A'.

The idea that a macro-property can have a causal influence on micro-elements is known in philosophy as 'downward causation'. However, this concept is subject to dismissive or sceptical claims in numerous fields, including the philosophy of biology (e.g. Craver and Bechtel, 2007), the philosophy of mind (e.g. Kim, 2005) and the social sciences (e.g. Vromen, 2010). Many philosophers

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

and some social scientists argue that relation 1 cannot be causal but refers to what is known in philosophy of mind as *supervenience*.[6] More generally, the very idea of 'inter-level cause' is deemed to be problematic because it embodies several contradictions with most of the reasonable concepts of causation entertained by scientists and philosophers: while intra-level relations are causal, inter-level relations are compositional (one also finds 'mereological' in the literature) because they relate parts to wholes. Hence, a whole (a system) is *defined* by its parts. Conversely, parts are defined by the fact that they are parts of a specific whole.

In the social sciences, the basic problem is to specify the nature of the relationships between some properties of the macro-state, i.e. social structures, and the elements at the micro-level, i.e. individuals' actions and beliefs. Economists generally dismiss 'structuralist' arguments on the account that social structures do not have a life on their own: they are the product of individual action and they would cease to exist if individuals stopped acting. Therefore, to argue that a social structure S has caused a set of individual actions A' is problematic: if we define S as a mere aggregate of a set of actions A (S supervenes upon A), it follows that we only need A to explain A'. This claim is merely a basic restatement of the supervenience argument. Despite some contrary views, it *does not imply*, however, that a social structure S cannot have an effect of its own, even if it is not a causal effect. Carl Craver and William Bechtel (2007) argue that the top-down effect between wholes and parts takes place through a hybrid combination of intralevel causal relations and inter-level compositional relations.[7] In the context of the study of an institutional mechanism, we have to determine the way the components of an institution S (say, a convention) can affect the behaviour A of agents while acknowledging that the convention supervenes from the behaviour of the very same agents (see Figure 2).

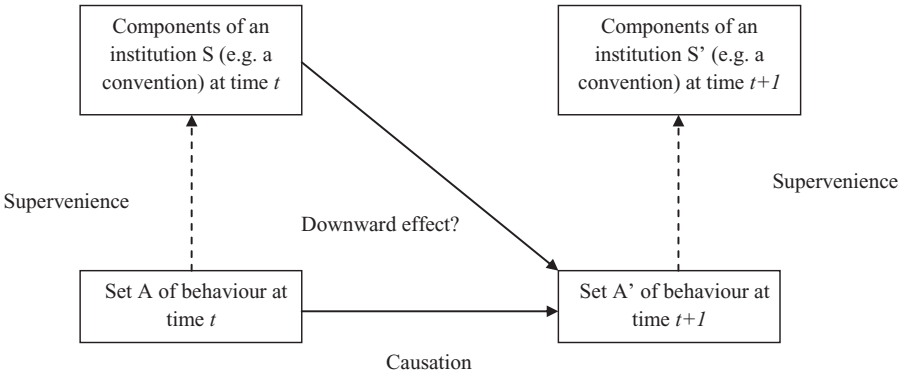
The challenge is thus the following: accepting the premise that downward causation from institutions toward individual action is problematic, we must qualify first the way institutions change agents' behaviour and then demonstrate that it is impossible *in principle* to make an account of an institutional mechanism without introducing in the *explanans* some macro-structures, including some exogenous institutions. It is important, however, to note that our argument does not rely so much on the premise that downward causation is a dubious

6 A macro-state M is said to supervene upon a set of properties P if two systems endowed with the same set of properties P are necessarily in the same macro-state M. Supervenience is a functional (or compositional) relation between two levels of properties or elements since it indicates that a higher-level property is defined by the lower-level properties that constitute it.

7 Craver and Bechtel (2007) call this combination a 'mechanism'. Their definition overlaps with the definition of a mechanism in analytical sociology: mechanisms are 'entities and activities organized such that they exhibit a phenomenon'. Hence, institutional mechanisms are a special kind of Craver and Bechtel's mechanisms. Note, however, that contrary to analytical sociologists, these authors take for granted the existence of higher-level causes (relation 4 in Coleman's diagram).

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

Figure 2. Supervenience and downward effect in an institutional mechanism



concept. Were we to accept that there is downward causation from institutions toward individual action, we would still have to demonstrate that an institutional mechanism is irreducible to individual actions and their interactions. But as the supervenience thesis presented above indicates, we would make the task unnecessarily harder. Hence, the point is that even if there could be causation from institutions toward individual action, the crucial thing about the institutions–agency relationship is that it is primarily constitutive.[8]

To take up this challenge, we have first to show how a change in an institution induces a change in the agents' behaviour, especially in a game-theoretic analysis. The concept of a *correlated equilibrium* is useful to formalise the *downward effect* of institutions on agency in a game-theoretic framework.[9] The point is to make explicit the connection between the existence of what we call a correlating device (here, an institution) and agents' beliefs regarding both the state of the world and the actions of other agents. An institution can then be interpreted as a 'choreographer' (Gintis, 2009), sending signals to all agents regarding what they have to do in a given situation. As will appear below, one of the advantages of conceiving institutions as correlated equilibria is to lead us to make explicit the epistemic features of the social interaction the agents are part of. Once we have

8 There are many complex philosophical and technical issues around the causation concept and it is beyond the scope of the paper to discuss them, less well to settle them. Our argument relies on the claim that macro-structures are constitutive of individual action, independently of any claim regarding a possible causal relationship between institutions and individual action. As Abell *et al.* (2010: 387–388) note, causal and constitutive inter-level relationships can perfectly coexist, at least in principle.

9 There is one obvious way in which institutional elements can be seen to produce downward effects in a game-theoretic framework: since pre-existing institutional elements define the 'rules of the game' in which players are involved, a change in the institutional elements would change the outcome of the game. Indeed, these kinds of situational mechanisms are at the heart of modern institutional economics. However, contrary to our concept of 'belief-mediated downward effect', it does not make explicit the constitutive relationships between institutions and individual agency. This is the reason why we choose to focus on the latter.

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

Figure 3. The stop-and-go game

		Bob	
		Go	Stop
Ana	Go	0; 0	3; 1
	Stop	1; 3	1; 1

established that an institution can be defined as a correlated equilibrium, we will then be able to show that some assumptions must be satisfied for an institution to act effectively as a correlating device. These assumptions are part of the constitutive relationship between macro-structures and agency. This framework has another advantage, namely that it may be more representative of the 'true' eductive process that takes place before individuals' choices. Indeed, in many social interactions we do not act because we deliberate upon what anyone else will do. We simply take for granted that an institution exists and we rely on nothing else than this 'social object' to make a choice.

We start with a fairly simple social interaction where two drivers, Ana and Bob, are arriving simultaneously at a crossroads from perpendicular directions. It is called the 'stop-and-go game' and is defined through a payoffs matrix (Figure 3).

Both players have two pure strategies available. This game has two Nash equilibria in pure strategy where one player stops and the other goes and one mixed-strategy equilibrium where each player chooses to stop with a probability of 2/3. Though simple, this interaction is problematic both for the analyst and from the point of view of the players. This is so because the game has several Nash equilibria, making the analyst unable to predict the outcome of the game. More importantly, the players cannot confidently choose a pure strategy because they cannot know what the other player will choose. The players can still play their mixed strategies but then coordination will fail in a significant proportion of time.[10] Hence, this interaction leads to significant coordination failures. However, imagine that there is a convention C indicating that the driver coming from the right has priority. Since Ana and Bob are coming from perpendicular roads, there are two possibilities: either Ana is coming from the right (w_1), or this is Bob (w_2). In the specific case of Ana and Bob, the convention C can be formulated as: 'if w_1 Bob stops, if w_2 Ana stops'. If both players follow the convention, it is clear that they always succeed in coordinating, leading to an expected payoff of 2 for each of them. The convention C is an example of a correlated equilibrium. Since the removal of C clearly changes the payoffs of the players, C can be said to have an effect on individuals' behaviour.

10 More precisely, coordination will fail $2/3*2/3 + 1/3*1/3 = 5/9$ of the time, leading to an expected payoff of 1 for each player.

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

This statement is made stronger with a generalised account of institutions as correlated equilibria in a game. A *game* is a mathematical description of a strategic interaction involving two or more players. Formally, a game G is defined as an ordered triple (N, S, $\mathbf{u}$) where:

$N = (1, 2, \ldots n)$, the set of the players of the game;
$S = S_1 \text{ x } \ldots \text{ x } S_n$, the Cartesian product of the sets S_i of available pure strategies for each player i;
$\mathbf{u} = S \rightarrow \Re$, the payoff function mapping S into an ordered n-tuple of real numbers for each strategy combinations in S.

The triple (N, S, $\mathbf{u}$) allows us to specify any strategic interaction. A foundational assumption stipulates that players are instrumentally rational, meaning that they maximise expected utility. Formally, writing $u_i(s_j; s^*)$ for the utility of player i playing strategy j with other players playing the vector of strategies s^*, player i playing s_j is rational if for any other strategy k we have $u_i(s_j; s^*) \geq u_i(s_k; s^*)$. For any game $G = (N, S, \mathbf{u})$, define Ω as the set of all possible worlds and $\omega \varepsilon \Omega$ the state of the world that actually obtains. Each player i possesses an information partition P_i on the states of the world; P_i can be thought as i's private information.[11] Define $\boldsymbol{P} = \cap_i P_i$ as the *common information partition*, i.e. the intersection of the information partitions of all players, i intersect such that $\boldsymbol{P}\omega$ means that at ω, each player knows which cell in $\boldsymbol{P}$ obtains. An event E* is *public* when $\boldsymbol{P}\omega \subseteq$ E* since for any $\omega \in \boldsymbol{P}\omega$, everyone knows E*, everyone knows that, and so on. We define a function f: $\Omega \rightarrow S$ as a system of strategies correlated with the state of the world, with $f(\omega) = (f_1(\omega), \ldots, f_n(\omega)) \in S$. Finally, we define ($\Omega$, $\mathbf{q}$) as the probability space over all possible states of the world.

A strategy profile $f^*(\omega) = (s_1^*, \ldots, s_n^*)$ forms a correlated equilibrium if, for each player i, for each state of the world ω and given the probability space (Ω, $\mathbf{q}$) we have:

$$\mathrm{E}_i[u_i(f^*(\omega)) \,|\, P_i\,](\omega) > \mathrm{E}_i[u_i(s_i, f^*_{-i}(\omega)) \,|\, P_i\,](\omega) \qquad (1)$$

Given each player's conjectures conditional on their information partition and the state of the world ω, nobody should have interest to deviate from the strategy prescribed by $f^*(\omega)$. We can usefully interpret any institutional elements (convention, norms, rules) as a correlated equilibrium by linking explicitly an institutional element with the occurrence of a public event E*. As indicated above, a public event is an event such that it refers to a world ω that belongs to the common information partition of the players. Hence, E* is common knowledge: everyone knows that E* obtains, everyone knows that everyone

11 P_i is called a possibility operator. It indicates for any player i and for any state of the world ω what other states $\omega' \in \Omega$ a player thinks may obtain. The information partition is made of cells containing one or several possible states. When one state ω obtains, all the states contained in the same cell are thought equally possible by the agent.

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

knows that E* obtains, and so on *ad infinitum*. We say that a public event E* is governed by an institution I (or one of its elements) when, whenever E* occurs, it is also common knowledge that everyone follows the correlated equilibrium $I(E^*) = (s_1^*, s_2^*, \ldots s_n^*)$, with s_k the strategies of the k player and $k = (1, \ldots, n)$.

One interesting property of correlated equilibria is that sometimes they are not Nash equilibria.[12] Therefore, the existence of a correlating device such as a convention or any other institutional element can 'make a difference': this correlating device can lead to an outcome that otherwise would be unattainable. The nature of this downward effect needs, however, to be clarified. Whatever the nature of the correlated device, we can consider that the downward effect captured by the correlated equilibrium operates through the formation of the players' beliefs. Hence, we propose to call this kind of downward effect triggered by institutions a *belief-mediated downward effect*.

First, note that the definition of an institution as a correlated equilibrium makes explicit the fact that an institution supervenes upon individuals' actions. Indeed, the *definition* of an institution I(E*) is the vector of strategies distribution $s_1^*, s_2^*, \ldots s_n^*$. Therefore, a change in this vector logically entails a change in the institution I(E*). This is exactly what supervenience means. Second, the traditional concept of Nash equilibrium figures that players form their beliefs independently.[13] The allowance of correlated equilibria departs from this assumption by allowing the possibility that players' beliefs do not form independently. This can be made clearer by writing ϕ_i^ω as the conjecture of player i concerning the strategies of the other players at ω. A (Bayesian) rational player i will choose the strategy $s_i^*(\phi_i^\omega)$ that maximises his expected utility given his conjecture ϕ_i^ω. The set of conjectures $\phi^\omega = (\phi_1^\omega, \ldots, \phi_k^\omega)$ forms an equilibrium (i.e. players' beliefs are mutually compatible) at ω if the set of strategies $s^*(\phi^\omega) = (s_1^*(\phi_1^\omega), \ldots, s_n^*(\phi_n^\omega))$ is itself an equilibrium. At the equilibrium, each player knows the others' beliefs and the corresponding strategies chosen by rational individuals form a correlated equilibrium.[14] The formulation $s^*(\phi^\omega)$ is just another way to describe the institution I(E*) and makes explicit its belief-mediated effect. The downward effect of a correlating device resides in its capacity to coordinate individuals' expectations by associating a distribution of conjectures to a specific

12 This statement needs some qualifications: for any underlying game where we allow for correlated strategies, we can write a new game where correlated strategies appear explicitly in the set of pure strategies. A correlated equilibrium is always a Nash equilibrium in such a game. However, the correlated equilibrium is not always a Nash equilibrium in the underlying game. In our example above, the correlated equilibrium does not correspond to any of the Nash equilibria of the basic game.

13 For a Nash equilibrium to be played, players must have common knowledge of the structure of the game, common knowledge of their instrumental rationality and of their mutual expectations, and their expectations must be formed independently (Vanderschraaf, 2010). If only the first three conditions are true, then the equilibrium is correlated.

14 Indeed, a fundamental theorem in epistemic game theory demonstrates that in a game where Bayesian rational players have common priors, the resulting distribution of strategies corresponds to a correlated equilibrium (Aumann, 1987; Gintis, 2009).

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

signal. Formally, an institution correlates a set of conjectures with the occurrence of a specific event.

4. Indeterminacy, institutions and constitutive relationships

The argument of the preceding section demonstrates that institutional elements (and more generally social structures) can have beliefs-mediated effects on agency. However, it undermines only partially the idea that explanations of institutions can be entirely based on individualistic grounds. In fact, one can still argue that the institutional elements serving as correlated devices are ultimately created by individual action. For pragmatic reasons, it is impossible or even unnecessary to provide in a rigorous way the micro-foundations of all institutional elements in a specific study but the theoretical possibility of a total reduction of institutions to micro-elements is not contradicted. However, the above argument can be expanded to demonstrate that a full micro-explanation of institutions is not only pragmatically difficult; it is also impossible *in principle* in a game-theoretic framework because institutions are constitutive of individual action. Our argument in this section is related to the 'determination process' associated to an institution and discussed in section 2: an institution selects one of the multiple outcomes amenable to self-reproduction through an institutional mechanism.

The definition of institutions as correlated equilibria can be used to demonstrate that the relationships between institutions and agency are *constitutive relationships*. Constitutive relationships are linked to the fundamental problem of indeterminacy in game theory: many games do not have a unique solution. This makes game theorists unable to predict the outcome of these games. More importantly, indeterminacy makes it impossible for rational players to settle on a choice of a strategy since what a player should do depends on what others will do. In other words, the optimal strategies for the players are non-computable.[15] The only way for the analyst to account for the players' actual choice is then to postulate the existence of some exogenous features not included explicitly in the mathematical description of the game. For example, following Thomas Schelling (1981), one can consider the existence of some 'focal points', making a strategy or a specific outcome particularly salient.[16] Alternatively,

15 See Binmore (1987) for a sophisticated statement of this argument. Of course, the problem of indeterminacy occurs because players have common knowledge of their rationality. Players with bounded rationality escape the problem of indeterminacy.

16 In this paper, we take salience (as any other kind of macro-structure) to be exogenous. However, see the recent attempt of Alberti et *al.* (forthcoming) to model salience as an emergent property. The authors provide an evolutionary model demonstrating how a shared conception of salience can develop in a population where individuals are using a replication heuristic. As long as a player's attempt to coordinate succeeds, he chooses an action (described by a label) that he perceives as similar to the one he used in a previous and similar (but not identical) game. The authors demonstrate that such a heuristic favours label

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

following Robert Aumann (1987), it is possible to assume that players have common priors, i.e. that they share the same beliefs distribution on how others will play.[17] We can also figure that, through a framing process, players have the same alternative description of a game where a unique equilibrium exists (Bacharach, 2006). Finally, we can simply assume that an exogenous institution (a convention or a social norm) already solves the game, as we have done in section 3. Focal points, common priors, frames and conventions are all examples of 'macro-structures', that is, elements irreducible to the characteristics of a sole individual.

Macro-structures are needed to overcome indeterminacy. They are *constitutive* of the action of rational agents because they are needed to make sense of agents' choices in a game-theoretic framework. The idea of constitutive relationships, of which supervenience is a special case, applies to part–whole relationships. As a general definition,[18] the relation between a higher-level entity S which has the property to do ψ (i.e. S's ψ-ing) and a lower-level entity P endowed with property φ is constitutive if the instantiation of ψ by S logically implies that P has φ. Since constitutive relationships (unlike causal relationships) are symmetric, the reverse also applies: the instantiation of φ by P logically implies that S has ψ. To see that supervenience is a form of constitutive relationship, take the institution I(E*) to be the higher-level entity. The property ψ of I(E*) is to instantiate a set of strategies $s_1^*, \ldots s_n^*$. The lower-level entities are the players of a game G. Each player i has the property φ_i defined by its strategy s_i^*. Hence, for ψ to be instantiated, each player i must have φ_i by logical necessity. By symmetry, the reverse also applies. This is the core of the idea of downward effect: when an agent follows strategy s_i^* as part of an institution I(E*), his action is *constituted* by this institution.

However, the definition of the concept of constitutive relationship in a game-theoretic framework goes farther. A rule (no matter if it is a convention, a norm, a legal rule or a contractual agreement) is never self-transparent to the individuals who have to follow it. The French sociologist Emile Durkheim (2004) was among the first social scientists to argue convincingly that any contract is necessarily incomplete: even when an agreement has been achieved among parties through an explicit contractual bargain, one can never be completely

types that have a higher intrinsic replicability. This model is a typical bottom-up evolutionary model but it also suggests that salience (measured by label types with high intrinsic replicability) matters in agents' decision problems. Note, however, that salience is ultimately exogenous in the model (measured through a parameter of intrinsic replicability), which makes the title of the paper slightly misleading.

17 As noted in footnote 14, this is equivalent to assuming that players are playing a correlated equilibrium. The justification of this assumption is known as the Harsanyi–Aumann doctrine according to which it is impossible to 'agree to disagree' (Aumann, 1976): if players are Bayesian rational and if their posterior beliefs are common knowledge, then they must update their beliefs until they agree. The relevance of this claim is clearly disputable but the assumption of common priors can be made more plausible when players share a common culture (Gintis, 2009).

18 See Craver and Bechtel (2007) and Vromen (2010) for similar characterisations.

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

sure that it will be respected since it is impossible to stipulate every condition in advance. Moreover, any contract (or legal rules) must be *interpreted* to have a meaning. This interpretation is necessarily 'outside' the contract or the rule and lies somewhere else. It seems sensible to suggest that the constitutive relations between institution*s* (note the plural) and agency is responsible for the possibility to go beyond the incompleteness of rules or contracts. For an institution to produce a downward effect on agency, there must be a *common understanding* among a population of what an institution means. The constitutive nature of the relationship between institutions and individual action is implied by the necessity of common understanding.

Our definition of institutions as correlated equilibria helps to account for this fundamental characteristic. Recall the definition in the preceding section: an institution I is a correlated device that specifies a set of mutually consistent beliefs and strategies $I(E^*) \in S$ such as, given a public event E^*, it is common knowledge that everyone will conform to $I(E^*)$.[19] This definition makes the institution I an external device making players' beliefs probabilistically dependent. This definition also makes explicit the fact that an institution is contingent on the realisation of specific states of the world through the occurrence of a (public) event E^*. This implies that an institution relies on the link between an event E^* and a set of strategies $I(E^*)$ by making common knowledge in a population that everyone will conform to $I(E^*)$. How E^* generates this common knowledge of expectations is a non-trivial question and indeed refers to what Durkheim and others call the 'incompleteness of contracts'. Following David Lewis's (1969: 52–59) theory of common knowledge, we say that an event E^* *indicates* an institution I when the common knowledge of E^* makes I common knowledge. This involves four conditions (Lewis, 1969; Vanderschraaf, 1998; Gintis, 2009):

(1) Everyone knows that E^* holds;
(2) The fact that E^* holds gives each individual good reasons to believe that everyone knows that E^* holds (E^* is common knowledge);
(3) Individuals are symmetric reasoners with respect to E^*;
(4) Because of (3), everyone infers from E^* that I holds.

Obviously, and as Lewis has first demonstrated, if E^* is common knowledge and if everyone infers I from E^*, then there is in the population an infinite set of iterated mutual knowledge of I, which implies that I is also common

19 Binmore (2008) disputes the claim that a convention (or any other institutional element) must be common knowledge to be followed. He emphasises that a convention can emerge from an evolutionary process without sophisticated epistemic assumptions. Moreover, he argues that common knowledge of conjectures is rarely attainable. Binmore's argument is insightful but there are several reasons to not agree with it: first, Chwe (2003) has demonstrated that common knowledge can be obtained quite easily in a population. Second, several authors (Cubitt and Sugden, 2003; Paternotte, 2011) have shown that common knowledge in Lewis's sense does not involve heroic epistemic assumptions. Finally, Binmore's evolutionary alternative has also its own problems and is unhelpful to account for the players' reasoning process *once* the convention has emerged.

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

knowledge. The third assumption is crucial: individuals are symmetric reasoners means that they share the same modes of reasoning and hence make the same inference from a given event.[20] In other words, individuals that are symmetric reasoners interpret the same situation in the same way. If a situation indicates that a convention or any other institutional element exists, then this convention will be common knowledge among the population of symmetric reasoners. It seems pretty clear that it is impossible to explain how an institutional element becomes common knowledge without making an appeal to other institutional elements and, hence, to macro-structures. For example, Gintis (2009) emphasises that the assumption of symmetric reasoning is deeply linked to the fact that individuals share a common culture (see also Chwe, 2003). The implication is that the working of any institutional element or institution is dependent on a set of other macro-structures that are part of the individuals' institutional environment. Two individuals will reason symmetrically from a given event to form a set of mutually consistent conjectures (and hence generate an institutional element such as a norm) only if cultural and social factors lead them to do so intuitively or consciously.

More generally, we contend that individuals must have a *common understanding* of the situation: for an event E* to trigger an institution I, E* must indicate I to everyone. Each person must be able to infer reflexively I from E* but must also understand reflexively that the same is true for everyone else in the population. This 'shared reflexivity' is analogous to the assumption of symmetric reasoning. Common understanding of a situation in the population appears when the mutuality of shared reflexivity reaches an arbitrary high level: *j* infers I from E*, *j* understands that *k* infers I from E*, *j* understands that *k* understands that *j* infers I from E*, and so on *ad infinitum*. Common understanding can be thought as a tacit joint agreement under which members of a group have *fictitiously* agreed on the definition of some salient features helping them to interpret the situation and, crucially, to have the insurance that everyone interprets the situation in the same way. It is sensible to suggest that shared reflexivity and common understanding are possible only when people share a system of symbolic cues and meanings. This implies that most of the time the working of a sole institution depends on the cultural and institutional environment in which it is embedded.

These factors refer to the macro-structures that are thus *constitutive* of any individual action. Hence, it is not only that an institution (singular) is constitutive of a particular action of an individual; institutions (plural) are always constitutive

20 Formally, for every person *i*, *j*, for any event E and any proposition P, and denoting **K**i(E) for '*i* knows that E holds', we must have:

$$\boldsymbol{K}_{\mathrm{i}}(\mathrm{E}) \wedge \boldsymbol{K}_{\mathrm{i}}[\boldsymbol{K}_{\mathrm{j}}(\mathrm{E})] \Rightarrow \boldsymbol{K}_{\mathrm{i}}[\boldsymbol{K}_{\mathrm{j}}(\mathrm{P})]$$

In plain language, if *i* knows that E holds and knows that *j* knows that E holds, then *i* is a symmetric reasoners with respect to *j* if *i* knows that *j* knows P.

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

of a specific individual action because, without them, a sole institution would be powerless and indeterminacy would be reigning.[21] Another way to put this idea is that while the relationship between a given institution I and a set of individual actions A can be described by the four logical possibilities of absence and presence (I and A, I and not-A, not-I and A, not-I and not-A), thus permitting us to claim that it is a causal relationship (but the supervenience thesis is still an objection), the effect of I on A *necessarily* implies a larger set of institutions or macro-structures generating the required common understanding.

5. Application: Avner Greif's comparative study of the role of cultural beliefs

This section will illustrate the above argument by taking Greif's (1994, 2006) insightful analysis of the role of cultural beliefs in economic organisation as a case study. It shows that what we have called macro-structures indeed play a decisive role in institutional analysis. This example is even more interesting since Greif *does not use* the concept of a correlated equilibrium. Thus, it will permit us to show that our argument is not tied to the use of a particular solution concept.[22]

In a series of work (Greif, 1989, 1993, 1994, 2006), Greif investigates the institutional foundations of several medieval economies in Europe and North Africa between the 10th and the 14th century. Greif has studied particularly the organisation of the Maghribi traders, a group of Jewish traders in the Mediterranean in the 11th century that organised into coalitions to enforce contracts over overseas exchanges. He has also examined the economic and political institutions of several Italian cities during medieval times. The basic claim of each of Greif's studies is that institutions, along with technology and geography, have played a decisive role in influencing the pace and the quality of the economic development of different regions. Hence, the link between institutions and economic performance is at the heart of Greif's work.[23] Each case study ingeniously combined a piece of formal analysis through a game-

21 The constitutive relationship between the institutional (or cultural) environment and agency seems to be an example of what the French philosopher Edgar Morin (2005) calls a 'hologrammatic relationship'. In a hologrammatic relationship, parts constitute a whole while in the mean time each part embodies the totality of the whole. Morin gives the example of cells: cells are parts of a whole (a living organism) but each cell contains the integrality of the genetic information constitutive of the whole. Our account indicates that in a game-theoretic framework, the relationship between agency and social structures is similar.

22 This section should not be read as a critique of Greif's analysis. Quite the contrary; Greif is completely lucid about the necessity to go beyond a narrow version of methodological individualism and of rational choice theory.

23 In a long and somewhat harsh review of Greif's book, Greg Clark (2007) notes that the economic efficiency of the various institutions scrutinised by Greif is never explicitly treated, even less measured. Indeed, since Greif does not use quantitative analysis (partly because economic data on the Medieval period are rare and subject to caution), no statistical link between a specific institution and some economic aggregates (growth or volume of trades) is demonstrated. Accordingly, the link between institutions and

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

theoretic model with a substantive historical narrative supported by several references to historical archives. The historical narratives are more than simple ornamentations since they provide the motivation for the construction of the formal models: the narrative justifies the assumptions of the model and the function of the model is to test a conjecture made thanks to the historical analysis (Greif, 2006: chapter 11). This methodology already indicates that the formal model builds on several exogenous features provided by the historical narrative.

In a methodological chapter called 'Endogenous Institutions and Game-Theoretic Analysis' (Greif, 2006: chapter 5), Greif makes several observations that indicate that the historical narrative provides the necessary elements to take into account the constitutive relationships between the macro-structures and agency. Notably, he states that the game-theoretic analysis 'corresponds to a situation in which institutionalised rules that aggregate private knowledge and information provide shared cognition, information and coordination' (Greif, 2006: 126). Further, he argues that any social interaction that a game is formalising takes place in a social framework and through a cognitive system that provide the 'behavioral and social rules' that indicate to each agent what he has to do.[24] These rules provide the necessary clues for each individual to form his beliefs: '*Social rules help individuals form beliefs – represented by probabilistic estimates – about the situation and what others will do*' (Greif, 2006: 133, emphasis in original). Finally, the author argues that it is crucial that individuals have a common understanding of the same objective situation or action and attribute to it the same meaning. The assumption of common knowledge in game-theoretic models is justified only in those cases where the social and behavioural rules (what we call 'macro-structures') provide the cognitive and informational resources necessary for this common understanding (Greif, 2006: 137–140).

In spite of the fact that he does not take the same analytical roads as ours, Greif reaches a strikingly similar conclusion: to work, a game-theoretic analysis of various institutions has to assume the existence of some macro-structures such as focal points or social rules. In the case of Greif's studies, this is the historical narrative that provides this necessary component. The comparative study between the institutional foundations of the economic organisation of the Maghribi traders and the merchants of the Italian city of Genoa (the 'Genoese') during the 11th and 12th centuries is a great example of the role of macro-

economic performance is only suggested. However, since Greif characterises some institutions as Pareto suboptimal, that indicates that some notion of economic efficiency is nevertheless part of the analysis.

24 '[U]sing the typification, categorization, and cognition provided by the cognitive system, the "behavioral rules" component of social rules specify what is expected of individuals with particular social positions in various circumstances: members of the two basketball teams have to stand in particular positions during a free throw, a driver must stop at what is cognitively defined as a stoplight, and a Maghribi merchant is expected to hire only honest Maghribi agent and to reciprocate in sharing information' (Greif, 2006: 131).

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

structures in the production of a 'micro-explanation' (Greif, 1994, 2006: chapter 9). In his work on the Maghribi traders (Greif, 1989, 1993, 2006: chapter 3), Greif suggests that the traders have to enforce contracts over overseas trades without the recourse to any kind of law system or formal jurisdiction.[25] He documents that overseas trades often led a Maghribi merchant to engage in a contractual relationship with an agent. The agent received a wage from the merchant to ship the merchandises to their destination. This involves a commitment problem since, during the travel, the agent is in possession of the merchant's capital and can behave opportunistically to take advantage of this fact, hurting the merchant's welfare in the process. Greif builds on historical archives to suggest that the Maghribis solved the commitment problem by organising a coalition of merchants building on a system of multilateral punishments: once a merchant estimates that he has been cheated by an agent, he alerts the rest of the coalition of the misbehaviour of the agent. The coalition was organised through an informal rule indicating that once an agent has been tagged as a cheater, he should not be hired in the future. Hence, the Maghribis were able to enforce commercial contracts thanks to a reputation-based institution. Greif develops a simple game-theoretic model showing that, under some assumptions, to not cheat a merchant is an optimal strategy for an agent given the expectation that he will not be hired in the future if he is tagged as a cheater.

Maghribis' reputation-based institution was effective in enforcing commercial contracts. However, the enforcement of contracts through a system of multilateral punishments also has some defects. Notably, since contracts are always incomplete, agents could behave strategically to take advantage of the absence of complete agreement on what is considered to be cheating.[26] More importantly, since reputation was an essential informational asset in the working of the coalition, the Maghribis were severely dissuaded to engage in commercial activities with traders not belonging to the coalition. In fact, the threat of the loss of reputation was not operating on outsiders and hence the punishment mechanism was not credible. The immediate consequence was the development of an economic segregation restricting the growth of market size and limiting the commercial expansion and activity of the region.

Some decades later in Europe and particularly in Italian cities, commercial activities were developing through radically different institutional mechanisms. Contrary to the Maghribis, the merchants of the city of Genoa, the Genoese, developed a set of formal rules enforced by proto-state institutions to organise commercial trades. These rules enabled the agents to commit *ex ante* to be

25 Greif's argument has been disputed by Edwards and Ogilvie (2008). They argue that economic cooperation between the Maghribis was enforced by a formal code and a public organization. See Greif (2009) for a response.

26 This problem was partially solved by the fact that the Maghribis shared a set of cultural rules fostering the common interpretation necessary to make common knowledge the traders' expectations.

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

honest *ex post*. Though not necessarily more efficient than the reputation-based institution of the Maghribis, the formal institutions of the Genoese did not generate any kind of economic segregation since agents' reputation was not taken into account by the merchants in the formation of their expectations. This led to the formation of an integrated economy.

Greif argues that these differences in the institutional foundations of the economic organisation of these two communities of traders originate in the respective cultural beliefs entertained by both communities. Cultural beliefs are defined as probability distributions over the different paths of play (in- and out-of the equilibrium path). By affecting players' expectations about the play out-of the equilibrium path, cultural beliefs change the threshold value of the parameters of the models required for a given set of strategies to form an equilibrium. Specifically, Greif contrasts the collectivist beliefs of the Maghribis with the individualistic beliefs of the Genoese and defines them as the amount of information the merchants have over the agents' past play and how this information is taken into account: 'Intuitively, in a collectivist society everyone is expected to respond to whatever transpires between any merchant and agent; in an individualistic society this may well not be the case' (Greif, 2006: 276). These two types of cultural beliefs lead to two different subgame perfect equilibria in the same game: in the individualistic equilibrium, a merchant randomly hires an unemployed agent, while in the collectivist equilibrium a merchant only hires an agent who has never cheated. In an individualistic society, merchants are not expected to take into account the agent's past behaviour; as a consequence, each merchant is indifferent between hiring a cheater and an honest agent. Therefore, to not cheat will be on the equilibrium path for an agent only if he is paid a higher wage than in a collectivist society.

The formal institutions in the city of Genoa developed in response to these cultural factors. Indeed, individualistic beliefs prompted the Genoese to adopt formal mechanisms to prevent opportunistic behaviours. More importantly, individualistic beliefs have been a favourable factor in the commercial expansion of European economies. Greif's concept of cultural beliefs is an exemplary case of macro-structures: these beliefs determine the actual path of play and, more importantly, the distinction between the two equilibria cannot exist without them. However, they are irreducible to the individuals' characteristics and can be linked to various structural factors such as the influence of religion on the evolution of kinship and tribes. Hence, Greif's comparative institutional analysis is a great illustration of a micro-explanation building on 'macro-structural foundations'.

6. Conclusion

Contemporary institutional economics is interested in providing micro-explanations to economic phenomena, particularly to the link between

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

institutions and economic efficiency. Since it starts from agents' actions and beliefs, game theory is the main tool to build such micro-foundations. In this respect, institutional economics can be seen as an attempt to uncover the social mechanisms responsible for both institutions' reproduction and performance. However, this paper has shown that the use of game theory to study institutions makes it necessary to assume the existence of macro-structures such as prior beliefs, frames or focal points. As a result, institutions in a game-theoretic framework are more than causal factors of individual action: they are constitutive of it. The constitutive relationship between institutions and individual action notably emphasises the fact that to understand the downward effect of an institution, the analyst cannot ignore the social structures surrounding it.

Our analysis relies on the somewhat underrated concept of a correlated equilibrium. Following Gintis (2009), we consider that the concept of a correlated equilibrium is the proper one to study norms and other institutional elements. However, our argument does not depend on this particular concept; it still holds even when institutions are not defined as correlated equilibria. The example of Avner Greif's comparative institutional analysis of the economic organisation of the Maghribis and the Genoese provides an interesting illustration. This example also indicates that micro-explanations building on formal models in institutional analysis cannot be complete both in practice and in principle: macro-structures are always necessary as an input in micro-explanations. Finally, our argument reinforces the claim that strict methodological individualism is either wrong or impossible (Hodgson, 2007). It is possibly extendable to looser version of methodological individualism such as Udehn's structural individualism, the latter being compatible with Coleman's diagram. Interestingly, this point can be made even clearer with the help of game theory (Gintis, 2009).

However, our argument should not be read as a critique of game-theoretic analyses of institutions. It underlines that game-theoretic models are not self-sufficient and need to be complemented with other tools and methods, such as historical case studies. Moreover, it shows that the rhetoric of methodological individualism which is still alive is a wrong-headed justification for the use of game theory in institutional analysis. Still, once the necessity to take macro-structures into account is recognised, game theory is a powerful tool to produce theoretical and empirical insights.

References

Abell, P., T. Felin, and N. J. Foss (2010), 'Causal and Constitutive Relations and the Squaring of Coleman's Diagram: Reply to Vromen', *Erkenntnis*, **73**(3): 385–391.

Acemoglu, D. and J. A. Robinson (2005), *Economic Origins of Dictatorship and Democracy*, Cambridge: Cambridge University Press.

Alberti, F., R. Sugden, and K. Tsutsui (forthcoming), 'Salience as an Emergent Property', *Journal of Economic Behavior and Organization*.

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

Amable, B. (2003), The Diversity of Modern Capitalism, Oxford: Oxford University Press.
Aoki, M. (2001), *Toward a Comparative Institutional Analysis*, Cambridge, MA: The MIT Press.
Arrow, K. J. (1994), 'Methodological Individualism and Social Knowledge', *AEA Papers and Proceedings*, **84**(2): 1–9.
Aumann, R. J. (1976), 'Agreeing to Disagree', *Annals of Statistics*, **4**(6): 1236–1239.
Aumann, R. J. (1987), 'Correlated Equilibrium and an Expression of Bayesian Rationality', *Econometrica*, **55**(1): 1–18.
Bacharach, M. (2006), *Beyond Individual Choice: Teams and Frames in Game Theory*, Princeton, NJ: Princeton University Press.
Bates, R. H., A. Greif, M. Levi, and J.-L. Rosenthal (1998), *Analytic Narratives*, Princeton, NJ: Princeton University Press.
Binmore, K. (1987), 'Modeling Rational Players. Part I', *Economics and Philosophy*, **3**: 179–214.
Binmore, K. (2008), 'Do Conventions Need to Be Common Knowledge?', *Topoi*, **27**: 17–27.
Chwe, M. S.-Y. (2003), *Rational Ritual: Culture, Coordination and Common Knowledge*, Princeton, NJ: Princeton University Press.
Clark, G. (2007), 'Review of Avner Greif's *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*', *Journal of Economic Literature*, **45**(3): 725–741.
Coleman, J. (1990), *Foundations of Social Theory*, Cambridge, MA: Harvard University Press.
Craver, C. F. and W. Bechtel (2007), 'Top-Down Causation Without Top-Down Causes', *Biology and Philosophy*, **22**(4): 547–563.
Cubitt, R. and R. Sugden (2003), 'Common Knowledge, Salience and Convention: A Reconstruction of David Lewis' Game Theory', *Economics and Philosophy*, **19**(2): 175–210.
Durkheim, E. (2004) [1895]), *De la division du travail social*, Paris: PUF.
Edwards, J. and S. Ogilvie (2008), 'Contract Enforcement, Institutions and Social Capital: The Maghribi Traders Reappraised', *CESifo Working Papers*, no. 2254, Munich: CESifo.
Elster, J. (1982), 'Marxism, Functionalism and Game Theory', *Theory and Society*, **11**(4): 453–482.
Field, A. J. (1981), 'The Problem with Neoclassical Institutional Economics', *Explorations in Economic History*, **18**: 174–198.
Field, A. J. (1984), 'Microeconomics, Norms, and Rationality', *Economic Development and Cultural Change*, **32**(4): 683–711.
Gintis, H. (2009), *The Bounds of Reason*, Princeton, NJ: Princeton University Press.
Glaeser, E., R. La Porta, F. Lopez-De-Silanes, and A. Shleifer (2004), 'Do Institutions Cause Growth?', *Journal of Economic Growth*, **9**(3): 271–303.
Granovetter, M. (1985), 'Economic Action and Social Structure: The Problem of Embeddedness', *American Journal of Sociology*, **91**(3): 418–510.
Greif, A. (1989), 'Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders', *Journal of Economic History*, **49**(4): 857–882.
Greif, A. (1993), 'Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition', *American Economic Review*, **83**(3): 525–548.
Greif, A. (1994), 'Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies', *Journal of Political Economy*, **102**(5): 912–950.

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press

Greif, A. (2006), *Institutions and the Path to the Modern Economy. Lessons from Medieval Trade*, Cambridge: Cambridge University Press.
Greif, A. (2009), 'Contract Enforcement and Institutions Among the Maghribi Traders: Refuting Edwards and Ogilvie', *Working Paper*, Stanford, CA: Stanford University.
Hedström, P. (2005), *Dissecting the Social. On the Principles of Analytical Sociology*, Cambridge: Cambridge University Press.
Hedström, P. and P. Bearman (eds.) (2009a), *The Oxford Handbook of Analytical Sociology*, Oxford: Oxford University Press.
Hedström, P. and P. Bearman (2009b), 'What is Analytical Sociology All About? An Introductory Essay', in P. Hedström and P. Bearman (eds.), *The Oxford Handbook of Analytical Sociology*, Oxford: Oxford University Press, pp. 3–24.
Hedström, P. and R. Swedberg (eds.) (1998a), *Social Mechanisms. An Analytical Approach to Social Theory*, Cambridge: Cambridge University Press.
Hedström, P. and R. Swedberg (1998b), 'Social Mechanisms: An Introductory Essay', in P. Hedström and R. Swedberg (eds.), *Social Mechanisms. An Analytical Approach to Social Theory*, Cambridge: Cambridge University Press, pp. 1–31.
Hodgson, G. M. (2006), 'What are Institutions?', *Journal of Economic Issues*, **40**(1): 2–25.
Hodgson, G. M. (2007), 'Meanings of Methodological Individualism', *Journal of Economic Methodology*, **14**(2): 211–226.
Jepperson, R. and J. W. Meyer (2011), 'Multiple Levels of Analysis and the Limitations of Methodological Individualisms', *Sociological Theory*, **29**(1): 54–73.
Kim, J. (2005), *Physicalism, or Something Near Enough*, Princeton, NJ: Princeton University Press.
Lewis, D. (1969), *Conventions. A Philosophical Study*, Oxford: Wiley-Blackwell.
Morin, E. (2005), *Introduction à la pensée complexe*, Paris: Seuil.
North, D. C. (1990), *Institutions, Institutional Change and Economic Performance*, Cambridge: Cambridge University Press.
Paternotte, C. (2010), 'Being Realistic about Common Knowledge: A Lewisian Approach', *Synthese*, **183**(2): 249–276.
Schelling, T. (1981) [1960]), *The Strategy of Conflict*, Harvard, MA: Harvard University Press.
Udehn, L. (2001), *Methodological Individualism: Background, History and Meaning*, London and New York: Routledge.
Vanderschraaf, P. (1998), 'Knowledge, Equilibrium and Convention', *Erkenntnis*, **49**(3): 337–369.
Vanderschraaf, P. (2010), 'Review of *The Bounds of Reason*, Herbert Gintis, Princeton University Press, 2009', *Economics and Philosophy*, **27**(1): 88–96.
Vromen, J. (2010), 'MICRO-Foundations in Strategic Management: Squaring Coleman's Diagram', *Erkenntnis*, **73**(3): 365–383.

https://doi.org/10.1017/S1744137412000045 Published online by Cambridge University Press