Economic Conventions

STOCKHOLM SCHOOL OF ECONOMICS
The Economic Research Institute/EFI

Address: Sveavägen 65, Box 6501, S-113 83 Stockholm, tel 08-736 90 00

Basic Orientation

The Economic Research Institute, at the Stockholm School of Economics (EFI) is devoted to the scientific study of problems in management science and economics. It provides research facilities for scholars belonging to these disciplines, training advanced students in scientific research. The studies to be carried out by the Institute are chosen on the basis of their expected scientific value and their relevance to the research programs outlined by the different sections of the Institute.

Research Divisions:

A	Work Life Issues and Personnel
B	Accounting and Managerial Finance
C	Managerial Economics
CFR	Center for Risk Research
D	Marketing, Distribution and Industry Dynamics
ES	Economic Statistics
F	Public Administration
FI	Finance
I	Information Management
IEG	International Economics and Geography
P	Economic Psychology
S	Economics

Independent Research Programs:

PMO	Man and Organization
	The Foundation for Distribution Research
	Research Program in Policy Sciences

Additional information about research in progress and published reports is described in our project catalogue. The catalogue can be ordered directly from The Economic Research Institute, Box 6501, S-113 83 Stockholm, Sweden.

Economic Conventions

Essays in Institutional Evolution

Karl Wärneryd

STOCKHOLM SCHOOL OF ECONOMICS
THE ECONOMIC RESEARCH INSTITUTE

A Dissertation for the
Doctor's Degree in Economics
Stockholm School of Economics 1990

ISBN 91-7258-302-9

Keywords:
Game theory
Public choice
Property ownership
Firm (theory of)
Monetary theory

Gotab, Stockholm 1990

Foreword

I have been fortunate in having Karl-Göran Mäler as my thesis adviser. His encouragement and expertise has been invaluable.

I also wish to thank the other two members of my dissertation committee, Mats Persson and Ingolf Ståhl. The former, in particular, spent many hours beyond the call of duty reading and discussing my papers in detail.

The following have contributed helpful suggestions: Robert Clower, Richard Day, Jan Fornell, Lars Hörngren, Nils Karlson, Sten Nyberg, Paul Segerström, Göran Skogh, Viktor Vanberg, Anders Warne, Jörgen Weibull, Lawrence White, Hans Wijkander, and Ulrich Witt.

Financial support for this project was provided by the Center for Study of Public Choice at George Mason University, Finanspolitiska forskningsinstitutet, Humanistisk-samhällsvetenskapliga forskningsrådet, Institute for Humane Studies, and the Stockholm School of Economics.

Chapter 4 is assembled from "Legal Restrictions and the Evolution of Media of Exchange," *Journal of Institutional and Theoretical Economics* 145 (4), ©1989 J C B Mohr (Paul Siebeck), and "Legal Restrictions and Monetary Evolution," *Journal of Economic Behavior and Organization* 13 (1), ©1990 Elsevier Science Publishers B V (North-Holland).

The initiate will appreciate that the text and figures were typeset by TeX.

Contents

Chapter 1

Conventions

1.1 Introduction

This work deals with conventions. I use the term "convention" to designate a logical property held in common by, for instance, ordinary day-to-day activities such as the following.

- Driving on the right-hand side of the road.
- Speaking a language.
- Using money.

The common property of the problem situations underlying these examples is the desirability in each case for an individual to conform to the pattern of behavior adhered to by most other people. If you expect most other people to drive to the right, you would want to do so yourself in order to avoid collisions. When planning a longer stay in a foreign country, you would like to have made an investment in the local language in order to be understood. And you would like to keep stores of goods that other people will accept in exchange—such as used to be the case with Kent brand cigarettes, but apparently no other brand of cigarettes, in the Romanian black market.[1]

To put it another way, these situations are all coordination problems. Traffic rules, languages, and money are conventions that solve them.

[1] Perhaps they will now become the official currency?

Which particular behavior should be the standard is typically not the source of conflict. In the extreme case, individuals are indifferent. So a convention is a social institution that prescribes individual behavior that all involved would prefer in the situations concerned, and that everyone knows about (Lewis [38]).

Less obvious examples of conventions are compatibility standards in, e g, the personal computer market. The choice between buying an IBM-compatible computer or an Apple Macintosh is much like the choice between learning different foreign languages. The suitability of one standard depends on who, if anyone, you expect to be interacting with—and ultimately on the total number of adherents to the standard. This type of coordination problem, and specifically the solution of having conventions embodied in goods, is sometimes called a problem of *network externality*, and is the subject of a special literature (see, e g, Farrell and Saloner [19,18] and Katz and Shapiro [33]). In general, however, any symmetric coordination problem with many people involved, where the attractiveness of a particular type of behavior will be frequency-dependent, may be said to exhibit network externality. So although the present work does not deal with industry standards *per se*, the network externality concept is central to the discussion here as well.

I wish to discuss two aspects of conventions. One is how they work. The other is how they may emerge from an initially convention-less state through the decentralized decision-making of individuals. The act of showing this I will term an *evolutionary explanation.*[2] Specifically, I will be using simple game-theoretical models and concepts. Examples of recent contributions in a similar vein are Schotter [57], Sugden [61],

[2]It should be mentioned at an early stage that I use the word "evolutionary" in a broader sense, to be explained more fully below, than just genetic evolution. In fact, since my explanatory approach does not in any way depend on genetic transmission of behavior, but on the conscious imitation of successful behavior by rational individuals, a misinterpretation by the reader would lead to major confusion. It is necessary to point this out since there does exist a minor literature in economics based on the idea of genetic selection for behavioral traits in human beings, typically inspired by mistaken readings of biologist Dawkins [14]. Apart from being reminiscent of social Darwinism, this approach suffers from a major conceptual flaw. There simply has not been enough time for any significant selection of behavioral genes (if, indeed, such exist) to have taken place *within* the human species. Cultural evolution, i e, by means of learning and imitation, on the other hand, is very fast.

Ullmann-Margalit [64], and Witt [75].

The rest of this chapter is organized as follows.

In Section 1.2 I introduce a game-theoretic formalism for talking about coordination problems. I note that the idea of conventions and conventional behavior go somewhat beyond the ordinary game theory approach. Conventions which are not enforced do not affect the solution sets of games. They embody a notion of bounded rationality. I argue that this self-imposed limitation of responses in certain types of situations is necessary. Such a limitation of repertoire lowers the uncertainty of systems. I note that this may be seen as the same as saying that conventions minimize transaction costs.

This work is concerned with explaining the emergence and functioning of central conventions of the advanced market society. In particular, an evolutionary, or spontaneous order, perspective is taken. In Section 1.3, I give a brief introduction to this methodological tradition in the social sciences.

Section 1.4 discusses how evolutionary explanations of the origin of conventions and other social institutions can be given in terms of game theory. I show that for coordination problems in general, evolution does not necessarily lead to efficient conventions. There is therefore a flaw in the well-known "evolutionary" argument for the as-if optimization approach in neoclassical economics.

Finally, Section 1.5 summarizes the specific applications that make up the rest of this work.

1.2 Conventions and Game Theory

1.2.1 Coordination Games

Stylizing the common features and casting the result in the terminology of game theory, real-life conventions can be seen as solutions to recurring coordination games, many of them similar to this basic, symmetric, two-player one:

$$\begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}, \end{array} \tag{1.1}$$

where the payoffs are those of the row player, and α and β are non-negative numbers. Depending on the context and the intended application, these numbers represent levels of "utility," amounts of money, number of children, or whatever. I will take them in general to represent the players' preferences over situations, interpreted in such a way that a situation associated with a larger number is preferred to one associated with a smaller number. When necessary, this scale will be taken to be cardinal.

Assume the players have to make their strategy choices simultaneously and independently of one another. The payoff matrix is common knowledge, but the players cannot communicate. Then this game has at least two reasonable solutions, (A, A) and (B, B). Each player wishes to choose the same strategy as the other. Assume that $\alpha > \beta$. Both would prefer the (A, A) situation, but cannot exclude the possibility that the other player will play B. Given that a player expects the other player to play B, his own best choice is B. But given that he plays B, his opponent's best choice is also B. In this sense, both (A, A) and (B, B) are plausible outcomes of this game. In the terminology of game theory, they are Nash equilibria, or simply equilibria.

The Nash equilibrium is an intuitively attractive solution concept. It captures the idea that the players should act so as to secure the best result given the circumstances. The *least* we would want to demand of a solution is that no individual could be better off by changing his action from the one prescribed by the solution. Because of this property, an equilibrium can be said to be self-enforcing.

However, an equilibrium can involve strategies which are *weakly dominated* in the sense that there is some other strategy available which yields the same payoff against the equilibrium profile, but a greater payoff against any other profile. A player who plays the weakly dominated equilibrium strategy could then be said to be making a rather irrational "threat" to deliberately hurt himself should the other players not play their equilibrium strategies. The criterion of *perfection*, which suggests discounting equilibria based on such threats was a first step toward refinements of the Nash equilibrium concepts, to be discussed below in Section 1.2.2.

If the example game is taken to be a unique situation, not much more can be said than noting that it has two equilibria in pure strategies. For purposes of predicting the outcome of actual play, there is a basic indeterminacy.

In fact, there is a third equilibrium if the players can randomize over their pure strategies. Strategies that are probability distributions over pure strategies are known as mixed strategies, and it is a fundamental theorem that in mixed-strategy equilibria the expected payoffs of all pure strategies that are assigned positive probabilities are equal. So in the example game above, there is a unique completely mixed equilibrium where the players play A with probability $\beta/(\alpha+\beta)$ and B with probability $\alpha/(\alpha+\beta)$. [3]

I will not be concerned here with two-player games as such, but with a special type of many-player game based on bimatrix games. Consider a large (for practical purposes infinite) population of players who are randomly paired to play a single symmetric bimatrix game. Such games have two types of equilibria in pure strategies, those where all players choose the same strategy (*monomorphisms*, to borrow terminology from biology) and those where different strategies are represented with different population proportions (*polymorphisms*). The second kind can easily be shown to be mathematically identical to mixed-strategy equilibria in ordinary bimatrix games, i e, they occur where the population proportions are such that the expected values under random pairing are equal for all strategies that are positively represented, when more than one strategy is positively represented.

Let n_A be the share of the population that use strategy A, and $n_B = 1 - n_A$ the share of the population that use strategy B. Under random pairing, the expected value of using strategy A is then $V_A(n_A) = \alpha n_A$, and the expected value of using strategy B is $V_B(n_A) = \beta n_B = \beta - \beta n_A$. Note that the expected values of the two strategies are both increasing in their respective numbers of adherents. This is the network externality.

The game has three equilibria. In the first, all players choose to play A, because if you expect all others to play A, you get α if you go along and zero if you deviate. Similarly, for all to choose B is an equilibrium. The polymorphism (in this case, a

[3]Note that the "better" strategy is played with the lower probability in the mixed equilibrium. This result may be considered counter-intuitive and perverse at first glance. But in the example, $\beta/(\alpha+\beta)$ is the maximum probability that can be assigned to A without making pure A the unambiguously best choice for the opponent. Since expected payoffs have to be equal at the equilibrium, it follows that the better the coordination equilibrium associated with a particular strategy is, the lower the probability in the mixed equilibrium must be. This is easier to understand in a dynamic context, which I will discuss in Section 1.4.3 below.

bimorphism) is found by setting the expected values equal to one another. It occurs at $n_A^* = \beta/(\alpha+\beta), n_B^* = \alpha/(\alpha+\beta)$, i e, it corresponds to the mixed equilibrium in the underlying two-player game.

Another distinct type of coordination game in the 2×2 family may be represented by the matrix

$$\begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix} \end{array}, \tag{1.2}$$

where the players would like to make sure they use ***different*** strategies. For this reason, I would like to label them ***division of labor*** games.[4] These games have no equilibria where all players choose the same action. In the unique polymorphism, $n_A^* = \alpha/(\alpha+\beta)$.

Note that the unique equilibrium is an asymmetric one for a symmetric game. Standard game theory supplies no answers to the question of who should do what in equilibrium, at least not without introducing payoff-irrelevant asymmetries. As will be seen below, an evolutionary perspective takes care of this assignment of roles, however.

1.2.2 Equilibrium Refinements

In general, any game which has more than one Nash equilibrium may be considered a coordination game, although not necessarily a ***pure*** one such as the one above. This is because the rational choice among equilibria is not usually self-evident.

Sometimes a particular solution has a strong intuitive plausibility to it. An example would be when all players prefer one equilibrium to all others and can communicate before acting. A closer look will reveal complexities from a game-theoretical point of view. A costless verbal agreement cannot be said to commit the players to anything. So why should the game that remains after preplay communication be any less a coordination problem? Yet we intuitively expect this to be the case.

[4]In standard game theory, the two types of coordination games would not be considered distinct from each other. This is because one can be gotten from the other by changing the numbering of the strategies of one player, and the labels of strategies are considered irrelevant for the solution. In the many-player games with random pairing studied here, however, the fact that the strategy sets of all players are assumed to represent the ***same behaviors*** is of crucial importance.

Major work in game theory is therefore today devoted to the search for reasonable Nash equilibrium refinements, i e, for methods of counting out some equilibria based on stronger notions of rationality. (For example, see the survey in van Damme [68].) One of the earliest refinements is Selten's [59] notion of *perfection*. This can be seen as an attempt to formalize the idea that real players cannot be counted on to never make mistakes. So for an equilibrium to be plausible it should be an equilibrium even if there is some small probability that "wrong" strategies will be played. For bimatrix games, this can be shown to be equivalent to requiring that equilibria do not involve weakly dominated strategies.

Note that the bimorphic equilibrium of the first example game of the preceding section is not a perfect equilibrium. At the equilibrium, individuals are indifferent between the two strategies, and at any other population proportions arbitrarily close to the equilibrium, some individuals would be better off using the strategy they are not using. This imperfection turns into dynamic instability in the evolutionary setting of Section 1.4.3. The polymorphism of the division of labor game is perfect, however.

The present work could also be seen as falling in the refinement category.[5] However, the refinement literature treats multiplicity of equilibria as a problem. For purposes of making predictions about the actual outcomes of real-world games, the refinement theorist would ideally want a *single* solution for each game (see Harsanyi and Selten [24]).

My perspective will be slightly different. The very multiplicity of equilibria for certain often recurring game situations in society can be seen as the explanation for the emergence of social institutions, i e, conventions. A convention is a rule that tells you how to act in certain common situations which might otherwise be ambiguous. Many important institutions of the market society are conventions in this sense.

1.2.3 Conventional Behavior as Bounded Rationality

Many coördination problems in real life occur again and again. The set of players interacting may change, but the basic structure remains the same. Consider the driver who is approaching another car head on. There is a question as to which side of the

[5] A nice brief discussion that relates the abstract world of equilibrium refinements to the reality of social institutions and evolution is found in Dasgupta [13].

road to keep to. In this situation, I now wish to suggest, most people do not conduct a rational calculation from scratch. In fact, if they did, they would reach the conclusion that the equilibrium situation in pure strategies is indeterminate. Then they might as well flip a coin. If this was the case, we would observe collisions in about fifty percent of cases like this, which seems a bit high.

In fact no mental computation of any significance need be carried out at all. The typical driver just adheres to a rule (like driving to the right) he has learned and expects other drivers to follow. This is how an arbitrary convention solves coordination problems.

In terms of standard economics and game theory, the typical driver's behavior is slightly suspect, however. It is not strictly rational. Rational, economic man would consider each new situation on its own merits, considering all and only objective facts about the game structure at hand. That a convention is in existence is *not* such an objective fact. The existence of a convention does not make it physically impossible for strangers you meet on the road to drive on the wrong side. They still have all their strategies available. So a convention does not affect the structure of the game, and therefore not the set of equilibria for rational players. However, indulging in a bit of "bounded" rationality would seem to save your life occasionally in this case.[6]

For this reason I believe the case for a "bounded rationality" approach is strongest in this type of situation. This has not been stressed enough in the literature. Neoclassical theory deals almost exclusively with a very special family of games which have well-defined optimum behavior. That is, each player has a dominant strategy in some sense. Those who wish to argue that real people would not play the optimal strategy in such a situation, i e, that their rationality is bounded, would have to do so based on ideas about the limited capacity of the brain to perform complex computations. However, relaxing the very stringent assumptions guaranteeing a unique equilibrium, we recognize the coordination problem nature of many real-life economic and other deci-

[6]The impact of my example would seem to be diminished slightly by the fact that, in most cases, real-world conventions as to which side of the road to use are actually enforced by law. This of course *does* change the payoff structure of the game. However, I venture that most people would adhere to such conventions even without enforcement, since it is in their interest to do so.

sions. The indeterminacy of situations with multiple equilibria poses a problem even for agents whose minds are perfectly able to perform computations of arbitrary complexity. Conventional behavior is *necessary* to resolve recurrent coordination problems.

A number of authors have noted the importance of self-imposed limitations on the set of actions. Most, however, like Elster [16] and Brennan and Buchanan [5], discuss mainly situations where there is a conflict either between an individual's short-run and long-run interests, or between individual self-interest and collective interest. This type of situation is often modeled by means of the famous "Prisoners' Dilemma" and similar games. I will not be concerned here with institutions which can be said to owe their existence to PD-like problems.[7]

1.2.4 Conventions Minimize Transaction Costs

Some authors, such as Heiner [30] and Langlois [36], specifically single out uncertainty as the origin of the necessity of limiting the repertoire of actions. I would like to suggest that this is the same as saying that conventions minimize that elusive statistic known as transaction cost.

This is true in a formal sense for the models under scrutiny if one accepts the following operationalization of the transaction cost concept. Dahlman [12] defines transaction costs as "resource losses due to lack of information." This would make them equal to the expected value of perfect information.

Consider the many-player version of the game in (1.1). Under perfect information, i e, no uncertainty, each individual would know beforehand who he would be paired with, and could adjust his decision accordingly. The expected value of the game would then be equal to $n_A\alpha + n_B\beta$.

For each strategy under uncertainty we may now form the difference between this

[7]If there are any such institutions. It is not obvious how one escapes the PD in a noncooperative setting. The PD framework has also been somewhat overexploited. Even if people do not know any game theory, they are likely to be familiar with the PD. There is an abundant literature dealing with "solutions" to one-person, two-person, and multiperson PD games, one-shot and repeated. Since the classic PD is essentially a paradox, and inherently unsolvable, however, I would like to avoid abusing it further.

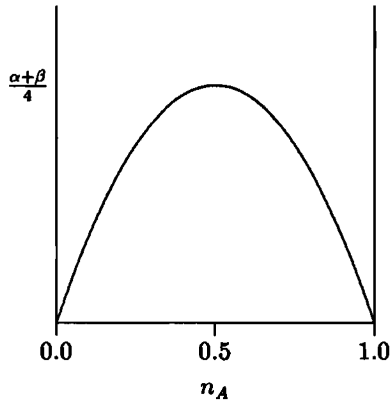

Figure 1.1: Transaction cost function for a game with two monomorphic equilibria.

optimal expected value and the actual expected value. For strategy A users, the transaction cost becomes $\tau_A = n_A\alpha + n_B\beta - n_A\alpha = n_B\beta = (1 - n_A)\beta$. For strategy B users, $\tau_B = n_A\alpha + n_B\beta - n_B\beta = n_A\alpha$. Then the average population transaction cost is

$$\tau_p = n_A\tau_A + n_B\tau_B = (\alpha + \beta)n_A(1 - n_A), \tag{1.3}$$

which is graphed in Figure 1.1.

Note that the transaction cost minima correspond to the two monomorphic equilibria, or conventions. The unstable bimorphic equilibrium is not a minimum.

It is also interesting to note that the transaction cost has a maximum, which always occurs at $n_A = .5$, regardless of the equilibrium payoffs. This point is of course the point of maximum uncertainty, measured as variance or entropy.

For the division of labor game in (1.2), the transaction cost function is

$$\tau_p = \alpha - 2\alpha n_A + (\alpha + \beta)n_A^2, \tag{1.4}$$

which is graphed in Figure 1.2 for the case where $\alpha = \beta$. This game has a transaction cost minimum at the unique bimorphic equilibrium.

The implications of this particular notion of transaction cost are discussed further in Chapter 5.

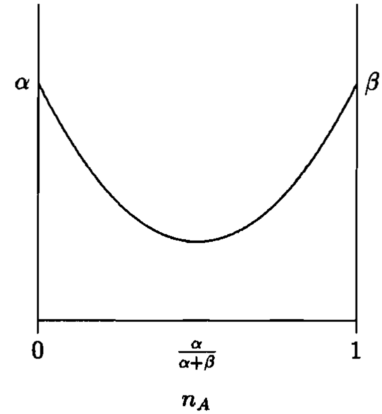

Figure 1.2: Transaction cost function for a division of labor game.

1.3 Explaining the Emergence of Conventions

1.3.1 The Spontaneous Order Tradition

This work is not be concerned with social institutions that have been deliberately introduced by some planning agency, such as government. Rather, I wish to discuss the *spontaneous* emergence and change of important conventions, where spontaneous means unintended by the actors involved.

The spontaneous order idea is arguably the oldest approach in the social sciences. It is well known to be older there than the same approach in biology. Charles Darwin explicitly mentions the population theory of Thomas Malthus as a major influence, and Hayek [28] finds evidence that he had read Adam Smith.

A nice survey of the oldest known material in this genre, starting with the medieval Spanish scholastics of the Salamanca School, is found in Barry [2]. I will content myself here with giving a brief overview of important modern contributions.

Austrian economist Carl Menger [43] expresses the evolutionary methodological view very clearly. Two types of social institutions are discerned.

Pragmatic. These are institutions explicitly designed by someone for a specific purpose.

Organic. Institutions that have evolved without being consciously directed by any single individual or committee.

Different types of explanations for the existence and functioning of the different types of institutions are necessary. Functional explanations, i e, in terms of what an institution achieves, are clearly sufficient for pragmatic institutions. Functional explanations of organic institutions are often possible, but can be misleading. (For interesting discussions of neoclassical economics and functional explanation, see Elster [15] and Langlois [37].)

Consider the case of natural language, the most prominent example of a spontaneous order. It is very easy to see what kind of human problems are solved by having a language. Explaining the origin of language as the result of design toward this end would be an obvious fallacy, however. (Though this has not stopped some people from trying.)

Menger also generously supplied the finest example yet of an evolutionary explanation of an economic institution, the theory of the origin of media of exchange.

Even for institutions known to be pragmatic, explanations of an organic kind can yield important insights (Ullmann-Margalit [65], Nozick [49]).

Perhaps most importantly, it could be argued that organic phenomena should be the main concern of the social scientist, since pragmatic social institutions really require very little explanation. Menger (p 146) speaks of

> ...a noteworthy, perhaps the most noteworthy, problem of the social sciences:
>
> How can it be that institutions which serve the common welfare and are extremely significant for its development come into being without a *common will* directed toward establishing them?

This view of the proper subject-matter of the social sciences is echoed by such important later contributors as Hayek [27], Polanyi [52,53], and Popper [55].

The life-long research project of Friedrich Hayek, exemplified by [26] and [29] is the more complicated case of these. Hayek is concerned with the spontaneous emergence of important social institutions in general, not only conventions, and argues for the central role of rules-of-thumb-following in this context. However, not all social institutions solve coordination games, and Hayek may be in trouble when seemingly arguing that PD institutions come about in a manner similar to conventions (see Vanberg [70]).

1.3.2 Desiderata for Evolutionary Explanations

An evolutionary system can be said to consist of two parts.

An innovation generator. This may be random, like genetic fluctuations, or the result of purposeful action, as when agents deliberately invent new modes of behavior or new tools.

A selection mechanism. Van Parijs [69] distinguishes between two types. The first, which could be called natural selection, works through the eventual expiration of the individual entities carrying unfit behavioral patterns or genes. An obvious example is the physical conflict for resources in nature. In economics, explanations pointing to this kind of selection mechanism may be exemplified by the work of Schumpeter [58] and Nelson and Winter [47], where firms that behave in inferior ways go bankrupt. The second kind of mechanism, and the one to be studied here, could be called reinforcement. This kind of mechanism works directly on the behavior pattern itself, without necessarily wiping out unfit carriers. The category covers diffusion of superior patterns through learning and imitation, i e, in general, processes of *cultural evolution* (See, e g, Cavalli-Sforza and Feldman [7]).

Note that in general no specific implications follow from simply invoking the concept of "evolution." To begin with, there is no guarantee that the relevant innovation generator will in fact generate our favorite candidate. Secondly, in order to explain a phenomenon, we must make plausible that the relevant selection mechanism would in fact select in favor of the phenomenon in question.

In a non-trivial evolutionary argument, the mechanism of selection should not "know" the phenomenon to be explained. Consider a number of women competing in a beauty pageant. Here the judges of the competition act as a selection mechanism. Clearly, an "explanation" of the selection of the most beautiful contestant as the winner, although an evolutionary explanation, would not be a very interesting one.

Non-trivial evolutionary explanations show how the explanandum arises as a byproduct, or epiphenomenon, of selection based on other criteria. In economics, the classic approach is to demonstrate that decentralized, self-interested decision-making by individuals can lead to socially beneficial outcomes without this having been explicitly intended by anyone.

An interesting explanation that fulfills the requirements for an evolutionary explanation, exemplifies my own approach, and is also part of the core of neoclassical economics (at least as it is taught to students) is Marshall's [40] derivation of the long-run market supply curve under perfect competition, which may be looked up in most intermediate-level microeconomics textbooks.

1.4 Evolutionary Games

1.4.1 Game Theory and Optimization

The introduction to biologist John Maynard Smith's book on evolutionary game theory [41] is interesting reading for an economist. Like neoclassical economics, evolutionary biology has been dominated by an optimization paradigm. This is justified as follows. Since Darwinian selection weeds out the less fit (by definition), what remains should be best adapted to the environment. The characteristics of surviving species may therefore be explained by showing how they are the optimal characteristics given the environment.

This argument is familiar to economists from, e g, Alchian [1] and Friedman [21]. If an individual or a firm does not maximize utility or profit, it will soon be outcompeted by those who do. Or rather, behavioral patterns other than optimization will disappear, as agents switch to the better way of doing things. So the existing institutional framework and phenomena that occur within it may be explained *as if* deliberate optimization had taken place. This is obviously a form of functional explanation.

As Maynard Smith points out, in many interesting situations in evolutionary biology there is no given environment to adapt to. Rather, the relevant environment consists of other individuals and the strategies they use. In economics, in fact, this is the *only* kind of situation studied. So what is sometimes called the "evolutionary argument" for optimization and efficiency in economics is somewhat misconceived. I will give an illustration of this in Section 1.4.3.

In this latter case, the appropriate tool of analysis is game theory, rather than optimization theory. Standard game theory, however, is based on a notion of rationality that for purposes of studying animal behavior is often too strong. Evolutionary game theory can be said to be concerned with the fitness of strategies rather than the conscious decision-making of individuals.

Strategies may reproduce in a population through being adopted by individuals imitating other individuals, or by being genetically programmed into offspring.

1.4.2 Evolutionarily Stable Strategies

Consider a symmetric two-player game G, with a strategy set $S = \{s_1, s_2, \ldots, s_n\}$ and payoff function $P : S^2 \rightarrow R$ which are the same for both players.

Definition 1.1 *A strategy $s^\star \in S$ is said to be an* evolutionarily stable strategy (ESS) *if, for all $s \in S$,*

$$P(s^\star, s^\star) \geq P(s, s^\star), \tag{1.5}$$

and, if $P(s^\star, s^\star) = P(s, s^\star)$,

$$P(s^\star, s) > P(s, s). \tag{1.6}$$

Note that condition (1.5) is just the definition of a best reply. Since we are dealing with symmetric games, this means an ESS must be part of a symmetric equilibrium. Condition (1.6) handles the weak case. It imposes a form of perfection requirement, so not all strategies which are part of symmetric equilibria are also ESS.

The motivation behind the second condition is the idea that a population using the strategy should be immune to a small invasion of "mutants." Consider a population where all individuals play a strategy $s^\star$. Suppose some individuals mutate and now play a mixture of $s^\star$ and some other strategy s, where s is played with some small probability ϵ. In order for an individual playing $s^\star$ to be better off than when playing s against a mutant, it must be the case that

$$P(s^\star, (1-\epsilon)s^\star + \epsilon s) > P(s, (1-\epsilon)s^\star + \epsilon s) \tag{1.7}$$

for small ϵ. As $\epsilon \to 0$, this condition approaches condition (1.5). For the case of a small but positive ϵ, note that (1.7) may (because of the linearity of expected payoff functions) be expanded to yield

$$(1-\epsilon)P(s^\star, s^\star) + \epsilon P(s^\star, s) > (1-\epsilon)P(s, s^\star) + \epsilon P(s, s). \tag{1.8}$$

Setting $P(s^\star, s^\star) = P(s, s^\star)$ and simplifying, we get condition (1.6).

1.4.3 Evolution Does Not Imply Efficiency

The ESS concept is a static property of certain strategies in one-shot games. There is no explicit dynamics. In many cases, however, one would like to study specific evolutionary processes.

It can be shown that ESS are fixpoints of a certain class of dynamical systems derived from bimatrix games (Zeeman [76], Taylor [62]). I will give a simple example of such a system.

Consider a large population of animals who are randomly paired in each time period. Each pair play a symmetric game G. Let the payoffs to an individual be the number of offspring (a non-negative number) it will asexually produce at the end of the period. The offspring are assumed to be genetically programmed to play the same strategy as the parent.

Then the population proportion of adherents to strategy i at time (generation) $t+1$ is given by

$$n_{it+1} = \theta(n_{it}) := \frac{\sum_j n_{it} n_{jt} P(s_i, s_j)}{\sum_i \sum_j n_{it} n_{jt} P(s_i, s_j)}. \tag{1.9}$$

Note that this is just the ratio of the average payoff to a strategy i user and the population average, times the population proportion of i-users. It is sufficiently general to be used also as an approximation for processes of imitation rather than genetic reproduction, with which the present work is concerned. The specification should then be taken as saying mainly that a strategy will attract more followers when it yields a locally better payoff.

Substituting the game in (1.1) into (1.9) yields the transition function for the pop-

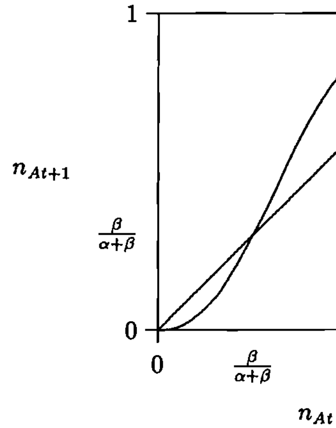

Figure 1.3: Phase diagram for a game with two monomorphic equilibria.

ulation proportion of users of strategy A for this specific game,

$$\theta_G(n_{At}) = \frac{n_{At}(n_{At}\alpha + n_{Bt}0)}{n_{At}(n_{At}\alpha + n_{Bt}0) + n_{Bt}(n_{At}0 + n_{Bt}\beta)} = \frac{\alpha n_{At}^2}{(\alpha+\beta)n_{At}^2 - 2\beta n_{At} + \beta}, \quad (1.10)$$

(since $n_{Bt} = 1 - n_{At}$).

Note that not all Nash equilibria are also fixpoints of the mapping θ_G. Let $\alpha = 0$. Then (A, A) is still an equilibrium of the two-player game, and $n_A = 1$ is an equilibrium population proportion for the game with random pairing. However, it is not a fixpoint, since (1.10) collapses to $\theta_G(n_{At}) = 0, \forall n_{At}$.

Letting α and β be strictly positive, then in general $\theta_G(n_{At})$ has three fixpoints. These are $n_{At} = 0$, $n_{At} = 1$, and $n_{At} = \beta/(\alpha+\beta)$. The situation is graphed in Figure 1.3.

Differentiating, we find that $\theta_G'(\beta/(\alpha+\beta)) = 2$, so the polymorphism is dynamically unstable. It is a critical mass, the minimum initial population proportion of A-users necessary for convergence to the A-convention.[8] Conversely, $\alpha/(\alpha+\beta)$ is the critical mass for the B-convention.

Now this model can be used for an evolutionary explanation of the kind outlined earlier. Assume that $\alpha > \beta$, i e, that the convention where all players choose A is the only efficient convention. For an innovation generator, consider randomly selected

[8]See Schelling [56] and Granovetter and Soong [23] for discussions of critical masses and "bandwagon effects" in the context of slightly different models.

starting-points of the dynamical system. That is, let there be a uniform distribution over the initial population proportions. The limit set, i e, the set of states that can occur as time approaches infinity from a randomly selected starting-point, is very simple for this system. The probability that $\beta/(\alpha+\beta)$ will be hit upon as the initial state is obviously zero, since any particular point has zero density. So it is not in the limit set. For starting points below $\beta/(\alpha+\beta)$, the system will converge to the B-convention, and for starting-points above to the A-convention. The limit set of values for n_{At} is therefore $\{0,1\}$.

The inefficient B-convention clearly cannot be counted out. Since this, if certainly not the only plausible model, is at least an evolutionary model, we may conclude there is nothing about the idea of evolution *per se* that guarantees efficiency in the way envisaged by the neoclassical "evolutionary" methodological argument.

However, it should be noted that evolution in this sense does lead to coordination. A convention will always be established. Furthermore, the better the efficient convention is, the smaller is its critical mass and the larger the probability that it will be converged upon from a random initial population proportion.

Now consider the division of labor game in (1.2). Its transition function is

$$n_{At+1} = \frac{\alpha}{\alpha+\beta}, \tag{1.11}$$

which is graphed in Figure 1.4. This means that the system converges to the bimorphic equilibrium in a single period.

The dynamic view also solves the problem of who should do what in equilibrium. If the initial proportions are incorrect, individuals will migrate from the less successful strategy until balance is achieved. So any starting distribution will do.

1.5 A Summary of the Rest of the Thesis

The rest of this work consists of essentially self-contained applications of the general theory of conventions sketched above to specific economic problem areas. Having read the present chapter will probably make these treatments easier to follow, especially for the reader unfamiliar with game theory.

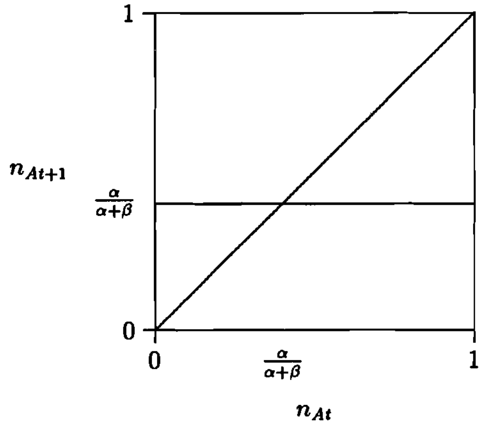

Figure 1.4: Phase diagram for a division of labor game.

In each case I have tried not only to convey a sense of how the phenomena in question can be seen as conventions, but also to use this insight to provide some new theoretical results for each problem area.

1.5.1 Language

It is intuitively obvious that the possibility of communication solves coordination problems. Explaining why this is, in terms of standard game theory, is not so easy, however. Assume a pre-play communication stage is added to a coordination game. It seems reasonable to assume that verbal communication is practically costless, i e, that it is "cheap talk." But if the payoffs in the ensuing game are unaffected by the messages sent at the communication stage, the set of equilibria is also unaffected.

As an example, consider two people who discuss a dinner date, to which each party must arrive independently. Assume there are two restaurants, A and B, and that restaurant A is also the preferred alternative for both individuals. Now suppose restaurant A is in fact agreed upon. In real life, we would predict that both players would then go to A, and not consider the alternative. To the game theorist, however, the situation where both go to B, regardless of what has been said, is still an equilibrium.

In order to understand what happens in reality, it must be realized that in situations like this, the coordination has already taken place at a higher level. For a world in which an individual who has stated his intention of doing A would be considered equally likely to do B cannot exist. Because then whatever string of sounds the individual uttered could not be said to mean "I intend to do A."

Languages are mappings that associate conventional meanings to symbolic actions. A symbolic action σ can be said to be a statement of the intention to do some action x in the future only if it is the case that people can realiably be expected to do x upon having done σ.

The assignment of meaning to symbols is the solution to a higher-order coordination game. Given that this game is in equilibrium, people who have signaled σ will tend to do x in equilibrium. However, as noted above, the existence of a convention does not affect the payoffs of a game. So when people act according to the conventional meanings of words in natural language, they are deliberately limiting their sets of responses. They are being boundedly rational.

1.5.2 Property

The principle that the first individual to claim a previously ownerless resource is to be considered its owner, sometimes called the *homestead* principle, is frequently found in liberal ethical systems and actual legal systems. It is easily seen as the solution to a coordination game.

Chapter 3 presents a very simple model to explain the emergence (or non-emergence) of spontaneous respect for possessor rights in a non-cooperative setting. Of particular interest is the effect of a "veil of uncertainty" surrounding one's own future possessions on individuals' commitment to respecting property rights. Contrary to a social-contract-theoretic hypothesis, a set of cases is identified where the stable degree of such respect falls as uncertainty rises. This is due to the fact that increased uncertainty, or income inequality, also means that, for some, the expected value of not respecting rights is correspondingly greater.

1.5.3 Money

One of the fundamental coordination institutions of the market society is that of the medium of exchange. Money solves the problem of the "double coincidence of wants" necessary for trade to take place in a pure barter economy, making specialization in production possible. When money is on one side of most transactions, and the money prices of other goods are known, the profitability of various projects can be compared according to a single standard. This is an essential prerequisite for the existence of firms.

Neoclassical monetary theory is not a good place to search for understanding of the medium of exchange function of money. In neoclassical general equilibrium, transactions are mediated by an imaginary auctioneer, which obviates the need for a medium of exchange. The great popularity of this approach in areas other than monetary economics, together with the idea that every theoretical construct must be amenable of incorporation into the GE framework, has led to some curious propositions. One of these is the so-called "legal restrictions" hypothesis: The only reason individuals hold currency, which yields no interest, rather than interest-bearing instruments such as government bonds, is that they are forced to by legal tender laws. (See Wallace [71].)

The function of a medium of exchange can be understood only in a setting where individuals have to actively search for someone to trade with, and where there is uncertainty. An individual would like to hold goods that other people will find attractive. In the case of a medium of exchange, a large part of the potential attractiveness of a money candidate is the likelihood that it will later be accepted in exchange by others. So the coordination problem in monetary anarchy consists of choosing a good to acquire and offer in exchange.

On a very abstract level, the problem with only two medium of exchange candidates could be modeled as a game of the type discussed previously. Let the payoff matrix of each agent be

$$G = \begin{array}{c c} & \begin{array}{cc} M & B \end{array} \\ \begin{array}{c} M \\ B \end{array} & \begin{pmatrix} \mu & 0 \\ 0 & \beta \end{pmatrix} \end{array}, \tag{1.12}$$

where $\mu, \beta > 0$. There are two goods, M and B. Individuals from a large population are paired randomly. An individual receives a positive payoff only when he has chosen the same good as the agent he is paired with to hold and accept in exchange.

Note that the discussion in Section 1.2.4 gives a precise content to the often-mentioned notion that a medium of exchange exists to lower transaction costs.

In light of the discussion of the previous section, the "legal restrictions" hypothesis may now be addressed. Let $\beta > \mu$. Clearly, the efficiency of a situation where everyone holds good B is not enough to make such a choice unilaterally rational. Furthermore, if the medium of exchange is deregulated, i e, legal tender laws are repealed, in a situation where M is the convention, there is no immediate evolutionary tendency to abandon M. Such a switch can only occur when a critical mass of B-users has been assembled. This could be the result of entrepreneurial or collective action. An analysis extended by the possibility of such action would of course also have to take into account its costs.

1.5.4 The Firm

The classic Coase [9] is cited in most modern "Theory of the Firm" literature. However, very little actual use is made of the notion of "costs of using the market mechanism" as the *raison d'être* of the firm. Obviously this is because no accepted coherent theory of what transaction costs are exists.

In terms of the division of labor game in Section 1.2.4 above, however, Coase's theory has a definite meaning. The random-pairing game may be seen as a simple model of a market where individuals have to search for trading partners under uncertainty. If α and β are positive, the equilibrium transaction cost as defined can never be zero. This means that it would be profitable for a group to get together and accept the allocation of actions by some central authority, thus reducing uncertainty.

In these terms, Coase's theory of the firm is seen to be an information-oriented one. It is then easier to relate it to various later theories which are informational in nature.

Chapter 2

Language

2.1 Introduction

Intuition tells us that verbal communication solves coordination problems. This could very well be the reason communication is seldom studied explicitly in non-cooperative game models. Another reason may be that, when actually modeled, it turns out to be far from obvious how communication solves anything.

We know from everyday life that the opportunity to communicate about, say, a dinner date, will usually make people end up in the same restaurant. A round of "cheap talk," i e, communication that does not affect payoffs, added to a game model, however, should not change the set of ultimate equilibria. You could have equilibria where the parties verbally agree on playing a particular equilibrium, but then actually play a different one. The only thing that should matter is that the actions finally taken constitute an equilibrium. But then communication has not solved the coordination problem.

Beyond the simple coordination of actions, it seems plausible that real communication would select Pareto-undominated equilibrium outcomes. This idea is an axiom in the literature on renegotiations-proofness (see, e g, Bernheim, Peleg, and Whinston [4]). No more fundamental game-theoretic basis seems to be available, however, as noted by van Damme [67], who suggests efficiency as a reasonable criterion but then finds that "unfortunately, one cannot turn to game theory for formal support of this argument as existing concepts fail to incorporate the meaning of ordinary language" (p 221).

Here a simple remedy will be attempted. Some aspects of the idea of conventional meaning of signals can be incorporated into equilibrium under verbal communication, as a Nash refinement. I will first argue that the seemingly most natural way of modeling communication is misleading.

Consider two people discussing whether to visit restaurant A or restaurant B, both of which are equally appealing to both players. This presumably brief communication session may be assumed to go on until both parties have stated intentions of going to, for example, restaurant A. The choice of restaurant might be decided on by the flipping of a coin. Now, regardless of this agreement, they will both be equally well off if they instead go to B. But it could be even worse. Suppose both players prefer restaurant A. Then there seems to be an equilibrium in which the players agree to go to A, but actually go to B.

In real life, we would not expect the latter case to be a plausible outcome. I would like to suggest that the paradox arises because of the failure in the above story to distinguish between a verbal signal and a statement of intention. Objectively, a "statement of the intention to go to restaurant A" is just a signal that could as well be called x. It does not have any meaning unless there is an established convention to the effect that the player who says x can thereupon reasonably be either expected to go to A, or to go to B. There is no such thing as a statement of intention to do A if you are not then also expected to actually do A.[1]

Communication presupposes existing conventions of language, which are themselves higher-level coordination equilibria. The reason the intuitively bizarre equilibrium described above may be ruled out is that if it could occur, there would not exist a well-defined language. We would then not be able meaningfully to call what goes on "communication" at all. Coordination is not a consequence of communication. It is the definition of it.

In Section 2.2 below, I spell out the problem with the intuitive way of modeling

[1]A statement of the type "Tomorrow, I intend to have three legs" has the form of a statement of intention. However, even if one could actually intend to have three legs, being mistaken about its feasibility, the statement would not generate the corresponding expectations in the audience. I will therefore not be referring to this kind of statement as a statement of intention.

communication in the context of a simple bimatrix game. I note that the statements to be made in communication are ultimately just arbitrary signals. Section 2.3 presents the notion of a ***signaling system*** used by a communicator and an audience to convey information about the state of the world from the former to the latter. This concept is an axiomatic attempt to determine when signals should be said to have meaning, i e, when there is communication going on.

In Section 2.4, I try to generalize the signaling system concept to situations where the communicator (or Sender) can choose the state of the world. This is done by studying extended games where the Sender in addition to playing a strategy from the original game can send a signal. In this extended game, the Sender chooses an encoding function, i e, a relation between his strategies and the signals he sends. The Receiver chooses a decoding function, which specifies which action should be taken when a particular signal is received. The equilibria of this new game are studied. Most importantly, I find that in games of pure coordination, only efficient equilibria are also equilibria of the extended game.

Section 2.5 gives examples for 2×2 games, and discusses some problems. Finally, in Section 2.6, I relate the axiomatic notion of communication of intentions to the way natural language seems to work.

2.2 The Problematic Intuition

Consider the following 2-player, 2-strategy game:

$$G = \begin{array}{c c} & \begin{array}{cc} a & b \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} 2,2 & 0,0 \\ 0,0 & 1,1 \end{pmatrix} \end{array}. \tag{2.1}$$

This pure coordination game has two equilibria, (A, a) and (B, b), in pure strategies. Both players would prefer (A, a). However, the possibility of the other player choosing the wrong action prevents a player from simply playing his Pareto-dominant equilibrium strategy. The problem is for a player to know beforehand which will be played by the other.

Until rather recently, when the standard literature at all addressed games of this kind, it was usually assumed that no communication is possible. Sometimes explicitly stated, the intuition is that if the players could talk before making their strategy choices, then they would verbally agree on one of the equilibria and subsequently play that equilibrium. This seems completely obvious. Indeed, it is verified by our normal experience of everyday life.

A popular textbook suggests that a round of communication should be explicitly added to the model. Interestingly enough, this is then not actually carried out. There is now an equally obvious intuition that says that if communication is "cheap talk," i e, does not change payoffs, then it should not change anything in the relevant part of the game.

To see this for the example game (2.1), introduce a restricted form of potential communication. Assume there is a first stage where the row player (denoted the Sender) may make a costless message of the form "I intend to play i," or μ_i for short, with $i \in \{A, B\}$. The column player, or Receiver, observes the Sender's statement. Thereupon both players take actions simultaneously.

The strategy set of the Sender for this extended game encompasses the four possible combinations of statements and conditional actions. The Receiver's strategy set consists of the four different possible contingency rules that associate an action with each of the Sender's two possible statements. Let the Receiver's strategies be written with the action if Sender stated μ_A in the first position and the action if Sender stated μ_B in the second position. Then the new game is

$$G' = \begin{array}{c c} & \begin{array}{cccc} aa & ab & ba & bb \end{array} \\ \begin{array}{c} \mu_A A \\ \mu_A B \\ \mu_B A \\ \mu_B B \end{array} & \begin{pmatrix} 2,2 & 2,2 & 0,0 & 0,0 \\ 0,0 & 0,0 & 1,1 & 1,1 \\ 2,2 & 0,0 & 2,2 & 0,0 \\ 0,0 & 1,1 & 0,0 & 1,1 \end{pmatrix} \end{array}. \tag{2.2}$$

There are now six subgame perfect equilibria, instead of the two in the original game. Any statement in the "communication" round is compatible with equilibrium, since all that matters is that an equilibrium is played in the final game. So the intuition seems to be wrong. This could hardly be termed a solution to the problem of coordination.

More importantly, what goes on here cannot meaningfully be termed communication.

2.3 Signaling Systems

The intuition is incorrect only as far as the above interpretation of messages goes. In fact, there is a straightforward way to exclude the counter-intuitive equilibrium where the Sender says he will play B and then plays A, and where the Receiver interprets the messages the other way around. This is done by first noting that it is incorrect to call μ_B a statement of the intention to play B if you will not play B on having made it and you are not expected to. A verbal signal can have a particular meaning only if a certain state of the world (in this case, action) can be reliably associated with it. If this is the case, we have a signaling system. One of the functions of natural language is to be an established signaling convention.

The seminal study of conventions of signaling is Lewis [38]. The situation studied is the following. There is an agent called the communicator, and one or more agents called the audience. The communicator knows which one of a set of states of the world holds. The audience can do one of a number of actions. For each possible state of the world, there is a unique action on the part of the audience which is preferred by both audience and communicator. The communicator can do one of a set of actions called signals, which number at least as many as the possible states of the world.

What Lewis calls the communicator's contingency plan, and I will call encoding function, is any function from the set of states of the world to the set of signals. If it is an injective (one-to-one) function, it is called admissible.

The audience's contingency plan (which I will call decoding function) is any function from a subset of the set of signals to the set of actions. If it is injective, it is called admissible.

These two functions are said to make up a signaling system if the range of the encoding function and the domain of the decoding function coincide, and, furthermore, their composition always maps a state of the world to the unique action preferred by all parties in that state of the world. Lewis proves that all and only admissible encoding and decoding functions belong to signaling systems.

More importantly, though, Lewis observes that signaling systems are noncooperative equilibria of a coordination game where the communicator and the audience choose their respective functions independently of one another (and act according to the chosen functions). This property makes this particular notion of "communication" something markedly more robust than just an arbitrary set of axioms.

I now wish to generalize the signaling system concept slightly to encompass two-person normal form games where the communicator can, in effect, choose the state of nature. In the following section, this is done by explicitly extending games so that they come to involve choosing contingency plans, which are acted upon.

2.4 Signaling as an Equilibrium Refinement

Consider a general two-person, normal form (i e, bimatrix) game G. One player is called the Sender, the other the Receiver. Let $S_i, i \in \{s,r\}$, be the players' sets of pure strategies in G. A typical element of S_i is s_i.

Let Σ be a set of actions, which will be called signals, which may be undertaken independently of the game G and which do not affect the outcomes of G, where $|\Sigma| = \max_{i\in\{s,r\}} |S_i|$. A typical element of Σ is σ. Define the set of functions Φ_s from the Sender's strategy set to the set of signals, and the set of functions Φ_r from the set of signals to the Receiver's strategy set. Φ_s is a subset of the set of all subsets of $S_s \times \Sigma$. Φ_r is a subset of the set of all subsets of $\Sigma \times S_r$. Further let Φ'_s be the subset of injective functions of Φ_s, and Φ'_r the subset of onto functions of Φ_r.

Now consider the following signaling-extended version G' of G. The Sender's strategy set in G' is $\Phi'_s \times S_s$, i e, the Sender chooses an action of G and an injective function $f_s : S_s \to \Sigma$. The set of signals is assumed unique to the game G'. The function f_s will be called the encoding function. It may be seen as a machine which automatically sends the signal associated by it with the chosen action.

The Receiver simultaneously, and independently of the Sender, chooses a decoding function f_r from Φ'_r. This automatically performs for the Receiver the action associated with the signal sent by the Sender.[2]

[2]This specification of G' could make it the normal form of a variety of extensive form games. For

In G', the payoff associated with the profile $(f_s s_s, f_r)$ is that associated with $(s_s, f_r(f_s(s_s)))$ in G.

Note that $f_r \circ f_s : S_s \to S_r$ is defined. If Sender and Receiver have the same number of strategies, $f_r \circ f_s(s_s)$ is bijective.

Let $E(G')$ be the set of pure-strategy Nash equilibria of G'.

Definition 2.1 *A game G will be said to* admit a signaling system *if $E(G')$ is non-empty.*

The justification for studying G' is as follows. The criterion that f_s should be injective captures the notion that signals in order to communicate something must distinguish between different actions. In the terminology of the literature on signaling games with costly signals, only "separating" equilibria can be said to involve communication.

That f_r should be onto rules out, among other things, the Receiver ignoring all signals.[3]

Why should only Nash equilibria be considered as potentially involving communication? The reason is that the Receiver must "understand" what the Sender "meant." The Receiver, if rational, may be said to understand a signal if he chooses a best reply to the action the signal stands for. In other words, for us to say of an $s_s \in S_s$ that it is understood, f_r must be chosen so that $f_r \circ f_s(s_s)$ is a best reply to s_s. But for the signal $f_s(s_s)$ to mean s_s, the Sender must do s_s. In particular, the Sender must want to do s_s when the Receiver expects him to. So f_s and s_s must be chosen so that the combination constitutes a best reply to f_r. Then $(f_s s_s, f_r)$ is a Nash equilibrium.

simplicity, no distinctions as to the actual sequence of the Sender's choices are made.

[3]The assumption also goes considerably beyond this, of course. One might intuitively find the weaker restriction of excluding just the constant functions, i e, those where the range is a single action, more appealing. Under this assumption, some much weaker results than those below can be proved (although they coincide for the 2×2 case). However, the "onto" assumption does not imply that the Sender may force the Receiver to do anything, since only equilibria are studied. We are interested in the potential coordinative properties of communication. A signaling system must then allow the possibility of selecting any avaliable equilibrium. In case a game has non-equilibrium actions available to the Receiver, which may be a source of trouble, the restricted game constructed by eliminating these actions may still admit signaling.

Let $x = (f_s s_s, f_r)$ be a strategy profile of G'. Then $(s_s, f_r \circ f_s(s_s))$ is said to be the profile induced in G by x.

The set $E(G')$, if it exists, will have at least 2 elements, since which signal is assigned to which action does not matter.

I first prove that the notion of equilibrium in G' is a Nash refinement for G.

Proposition 2.1 *Every element of $E(G')$, if there are any, induces an equilibrium in G.*

Assume there is some $e \in E(G')$ which is not in $E(G)$. Then for at least one player, the action in G implied by e is not a best reply. Given the choice of the other player, there is some other action that is a best reply. But this latter action can also be implemented by a strategy available in G'. Then e cannot be an equilibrium of G'.

Definition 2.2 *If a game G admits a signaling system, then we may say that the equilibria of G induced by the elements of $E(G')$ are equilibria of G under verbal communication, or properly conceived "cheap talk."*

The equilibria have the following properties.[4]

Proposition 2.2 *Let $u_s^\star$ be the greatest outcome that the Sender can get in G. Then every element of $E(G')$, if there are any, gives the Sender $u_s^\star$.*

To prove this, assume there is a strategy profile e in G' which does not give the Sender $u_s^\star$. Note that given the function f_r^e chosen by the Receiver in e, the Sender can attain any feasible outcome in G by varying f_s and his strategy choice. So if e does not give the Sender $u_s^\star$, it cannot be a best reply.

Proposition 2.3 *If there is an equilibrium of G that gives the Sender $u_s^\star$, then G' has an equilibrium.*

Let $(s_s^\star, s_r^\star)$ be the equilibrium of G that gives the Sender $u_s^\star$. There are now $f_s^\star$ and $f_r^\star$ such that $f_r^\star(f_s^\star(s_s^\star)) = s_r^\star$. Clearly, $f_s^\star s_s^\star$ is a best reply for the Sender against $f_r^\star$ since

[4]These are properties similar to those possessed by the rationalizability refinement in Farrell [17]. The latter, however, is based on the unsatisfying idea that following your own equilibrium suggestions is "focal," a term that is not defined. No distinction is made between signals and suggestions.

it gives him u_s^*. The Receiver's strategy f_r^* is a best reply for G' since s_r^* is a best reply against s_s^* in G. Therefore $(f_s^* s_s^*, f_r^*)$ is an equilibrium of G'.

The following propositions follow immediately.

Corollary 2.1 *A game G admits a signaling system if and only if there is an equilibrium of G that gives the Sender u_s^*.*

Examples of games which do not admit signaling systems in this sense are, trivially, games which do not have Nash equilibria, i e, games of pure conflict. In general, games where there is an incentive to bluff, i e, where u_s^* can only be had in non-equilibrium profiles, do not admit signaling systems.

Corollary 2.2 *If G is a game of pure coordination, the set of equilibria of G induced by $E(G')$ is equal to the set of Pareto efficient equilibria of G.*

Note that G' is a bimatrix game which may itself be extended into G'', and so on. Let $E^n(G)$ be the set of equilibria of G induced by the nth iterated extension.

Corollary 2.3 $E^n(G) = E^1(G)$, *for all* $n > 1$.

That is, allowing "communication about communication," up to the desired meta-level, does not further refine equilibrium.

2.5 Examples and Anomalies

I will give examples using 2×2 games. For such games, the set of admissible encoding functions for the Sender consists of the following two.

$$\begin{aligned} f_s^1(s_s) &:= \begin{cases} \sigma_1, & s_s = s_s^1; \\ \sigma_2, & s_s = s_s^2; \end{cases} \\ f_s^2(s_s) &:= \begin{cases} \sigma_2, & s_s = s_s^1; \\ \sigma_1, & s_s = s_s^2. \end{cases} \end{aligned} \tag{2.3}$$

The set of admissible decoding functions for the Receiver are these.

$$\begin{aligned} f_r^1(\sigma) &:= \begin{cases} s_r^1, & \sigma = \sigma_1; \\ s_r^2, & \sigma = \sigma_2; \end{cases} \\ f_r^2(\sigma) &:= \begin{cases} s_r^2, & \sigma = \sigma_1; \\ s_r^1, & \sigma = \sigma_2. \end{cases} \end{aligned} \tag{2.4}$$

Let

$$G = \begin{array}{c c} & \begin{array}{cc} s_r^1 & s_r^2 \end{array} \\ \begin{array}{c} s_s^1 \\ s_s^2 \end{array} & \begin{pmatrix} \alpha_s, \alpha_r & \beta_s, \beta_r \\ \gamma_s, \gamma_r & \delta_s, \delta_r \end{pmatrix} \end{array} \tag{2.5}$$

be a general 2×2 game. Then in general we have that

$$G' = \begin{array}{c c} & \begin{array}{cc} f_r^1 & f_r^2 \end{array} \\ \begin{array}{c} f_s^1 s_s^1 \\ f_s^1 s_s^2 \\ f_s^2 s_s^1 \\ f_s^2 s_s^2 \end{array} & \begin{pmatrix} \alpha_s, \alpha_r & \beta_s, \beta_r \\ \delta_s, \delta_r & \gamma_s, \gamma_r \\ \beta_s, \beta_r & \alpha_s, \alpha_r \\ \gamma_s, \gamma_r & \delta_s, \delta_r \end{pmatrix} \end{array}. \tag{2.6}$$

As an illustration of Corollary 2.2, let G_1 be the pure coordination game from (2.1). Then

$$G_1' = \begin{array}{c c} & \begin{array}{cc} f_r^1 & f_r^2 \end{array} \\ \begin{array}{c} f_s^1 s_s^1 \\ f_s^1 s_s^2 \\ f_s^2 s_s^1 \\ f_s^2 s_s^2 \end{array} & \begin{pmatrix} 2,2 & 0,0 \\ 1,1 & 0,0 \\ 0,0 & 2,2 \\ 0,0 & 1,1 \end{pmatrix} \end{array}. \tag{2.7}$$

The only equilibria of this game are $(f_s^1 s_s^1, f_r^1)$ and $(f_s^2 s_s^1, f_r^2)$, those which induce the efficient equilibrium of G_1.

Some cases of what might be seen as anomaly arise because the communication is one-sided. If there is conflict of interest, as in the "Battle of the Sexes," this conflict is resolved in favor of the Sender, who gets to choose his preferred equilibrium. Let

$$G_2 = \begin{array}{c c} & \begin{array}{cc} s_r^1 & s_r^2 \end{array} \\ \begin{array}{c} s_s^1 \\ s_s^2 \end{array} & \begin{pmatrix} 0,0 & 2,1 \\ 1,2 & 0,0 \end{pmatrix} \end{array}. \tag{2.8}$$

Then

$$G_2' = \begin{array}{c c} & \begin{array}{cc} f_r^1 & f_r^2 \end{array} \\ \begin{array}{c} f_s^1 s_s^1 \\ f_s^1 s_s^2 \\ f_s^2 s_s^1 \\ f_s^2 s_s^2 \end{array} & \begin{pmatrix} 0,0 & 2,1 \\ 0,0 & 1,2 \\ 2,1 & 0,0 \\ 1,2 & 0,0 \end{pmatrix} \end{array}, \tag{2.9}$$

where the only equilibria are those which give the Sender his greatest outcome. It can safely be said, however, that no symmetric model of communication will ever resolve the indeterminacy of this kind of game. In the case at hand, the Receiver might as well choose to ignore communication altogether.

Another reason to restrict the scope of implications of one-sided "cheap talk" to pure coordination games is that the refinement selects some equilibria which are not perfect. As an example, let

$$G_3 = \begin{array}{c c} & \begin{array}{cc} s_r^1 & s_r^2 \end{array} \\ \begin{array}{c} s_s^1 \\ s_s^2 \end{array} & \begin{pmatrix} 2,0 & 0,0 \\ 0,0 & 1,1 \end{pmatrix} \end{array}. \tag{2.10}$$

Here, only (s_s^2, s_r^2) is perfect, i e, involves only undominated strategies. Under communication, however, the only equilibria are those which induce (s_s^1, s_r^1).

The present notion of communication is extremely sensitive to a certain kind of perturbation which might be called "transmission error." Assume a signaling convention is in existence when G_1 above is to be played, e g, the (f_s^1, f_r^1) one. Now assume there is a positive probability ϵ that given his action choice, the Sender will send the wrong signal. If the Receiver is allowed the opportunity to ignore the signal, we get the game

$$\hat{G}_1 = \begin{array}{c c} & \begin{array}{ccc} s_r^1 & f_r^1 & s_r^2 \end{array} \\ \begin{array}{c} f_s^1 s_s^1 \\ f_s^1 s_s^2 \end{array} & \begin{pmatrix} 2,2 & 2-2\epsilon, 2-2\epsilon & 0,0 \\ 0,0 & 1-\epsilon, 1-\epsilon & 1,1 \end{pmatrix} \end{array} \tag{2.11}$$

in expected payoffs. For $\epsilon = 0$, $(f_s^1 s_s^1, f_r^1)$ is the unique perfect equilibrium. For ϵ arbitrarily small but positive, however, the same strategy profile is not even an equilibrium at all. This would seem to mean the Receiver has no reason to pay attention when there is even the slightest probability of transmission error. This is intuitively unappealing, since we would expect some communication, even if error-prone, to be better than none at all.

2.6 Natural Language

In real life, no extended games of the G' sort are actually played. This game is merely a device for deriving reasonable conditions that have to be fulfilled for us to speak of something as "communication" in any meaningful sense.

In actual situations there is seldom any need to contemplate which signaling system to use. There is typically an already established one that suggests itself. As an example, you tend instinctively to use the language spoken in the geographical area where you currently are, if you know it.

One might therefore say that it is not so much particular equilibria which are "focal," but rather particular signaling conventions.

Above, the set of signals was assumed unique to the particular game. This is in fact an important difference from the way natural language functions. In the latter domain, the same signals can be used in many different situations, a property Barwise and Perry [3] call the "efficiency" of language. This raises the possibility of lying, a meaningless concept in the artificial game G'.

In order to be statements of intention in a language, specific signals need to be associated with specific action patterns in others' expectations. These behavior patterns may have a very abstract, high-level description, such as "delivering on Thursday," so they do not fit precisely the extremely rigorous game-theoretical definition of a strategy (which refers only to the situation at hand). Since signals in natural language can be used in different types of game situations, it may be the case that the conventional expectation generated by a particular signal can be profitably misused in some situations.

Consider, however, a particular expression σ in natural language. For σ to mean s, the speaker must be expected to follow the action of uttering σ by the action s. For such an expectation to arise generally, it must usually be the case that it is rational to actually do s when expected to do s. That is, s, which is an action that can be undertaken in a large number of different games, must in general be capable of being credibly promised. This leads to an intriguing *a priori* observation about the kinds of situations in reality to which statements of intention in language are applicable. These must be predominantly of a non-competitive character, since otherwise no linguistic convention could arise or be stable. There can be no statements of intention attached to actions which no one would ever perform.

Chapter 3

Property

3.1 Introduction

The property rights principle that the first individual to claim a previously unowned resource is to be considered its owner, sometimes referred to as the *homestead* principle, is a basis for many classical liberal ethical systems (notably that of Locke [39]) and is recognized in common law. The reasons for the appearance of the principle in the two different instances must be thought to be different, however. In the case of natural law ethics, the concept is typically derived from metaphysical and epistemological axioms concerning the nature of reality, human nature, etc, and is an imperative independently of whether any individual ever chooses to respect it. Common, or judge-made, law, on the other hand, is an evolutionary process of pragmatic human social problem solving.[1] In this chapter, I will be concerned with explanations of when and why we in the latter case (and similar situations of decentralized decision-making) would expect the homestead principle to emerge spontaneously.

Curiously few economists have devoted attention to state-of-nature theory, i e, the study of situations where there are no institutional constraints on individuals' use of violence to get what they want. Two exceptions are Bush [6] and Umbeck [66]. The latter tests his theory of rights creation in anarchy using empirical data from the 19th

[1]As an aside, at least one author (Sugden [61]) seems to think that no useful distinction between natural and common law exists. This appears to me an unwarranted blurring of the dividing line between normative moral philosophy and positive economics.

century goldrush in California. Although the model presented here differs radically from Umbeck's, I have retained the suggestive gold-digger framework.

In evolutionary biology, the theoretical situation is different. Respect for first-claimer "rights" when a number of individuals are in potential conflict over a resource has been empirically observed in a variety of species. The seminal game-theoretic analysis of this appearance of the homestead principle is due to Maynard Smith and Price [42], in an article which is probably also the first example of the *evolutionary game theory* approach (see Maynard Smith [41] or van Damme [68] for comprehensive introductions). Assuming that an individual is genetically programmed to be either aggressive or passive, that resource-possessors are randomly paired with non-possessors for interaction, and that a given individual is as likely to be in possession of something as not, Maynard Smith and Price find conditions for non-violence to be viable equilibrium behavior.

One motivation for the present chapter has been, in applying this kind of thinking to human societies, to investigate the results of allowing the probability of being in the possessor role to possibly be something else than 1/2.

That the model is discussed in terms of gold-digging is not intended to reflect a limitation of the scope of conclusions to such scenarios only. From the Lockean construct a variety of comprehensive rights could possibly be derived (see, e g, Nozick [49]). Most notably the homestead principle seems to imply the invidual's right to the product of his own talents. Income redistribution schemes interfere with such a notion.

One question of interest in this context is whether the presence of a higher relative degree of uncertainty about one's own future position would, as argued by Brennan and Buchanan ([5], p 28 ff), make individuals more inclined to cooperative solutions, i e, lead them to desire a stronger commitment to well-defined property rights. Since my aim is a positive theory, however, the social-contract-theoretic approach of Brennan and Buchanan is replaced by a non-cooperative game situation.

The chapter is organized as follows. Section 3.2 discusses the symmetric biological model in some detail. Section 3.3 introduces a simple model of a non-cooperative binary choice situation, related to the biological model, but with a built-in asymmetry of possession. Section 3.4 discusses equilibria of the model. Uncertainty in the contractarian sense is crucial, but a set of cases can be identified where it works contrary to the contractarian hypothesis.

In Section 3.5 I argue that the equilibrium is a stable fixpoint of an evolutionary process of imitation of successful decisions in the repeated game, analogous to the "evolutionary stability" discussed by Maynard Smith in biological models.

Finally, in Section 3.6, I discuss possible implications of the model for real-world societies.

3.2 Baboons

3.2.1 Symmetric "Hawk and Dove"

The seminal game-theoretic analysis of the homestead principle in animal behavior is due to Maynard Smith [41]. The version presented here is, apart from some superficial differences, essentially identical to Maynard Smith's treatment.

Respect for first-claimer rights when a number of individuals are in potential conflict over a resource has been empirically observed in a variety of species. I will consider here a hypothetical case involving two male baboons, who are the players of the game, and a female baboon, who is the prize. Being able to mate with the female has a value $V > 0$ to each of the parties. Assume that one of the males, labelled the *possessor*, has come upon the female first, and is challenged by the *intruder*. Each agent may choose to be either aggressive ("Hawk") or passive ("Dove")[2]. In case both choose to aggress, they are assumed to have an equal chance of winning the fight, but both suffer injuries $C > 0$ as a result. In case both choose to be passive, they are assumed to share the female, each getting enjoyment $V/2$ out of this *ménage à trois*. Finally, in the other possible cases, whoever plays "Hawk" when the other is a "Dove" gets the female for himself.

The symmetric "Hawk and Dove" game may be described completely by the follow-

[2]Although I am supposedly discussing baboons here, these designations of the strategies are conventional.

ing matrix of expected possessor payoffs:

$$\Pi = \begin{array}{c} \\ \text{Hawk} \\ \text{Dove} \end{array} \begin{array}{c} \begin{array}{cc} \text{Hawk} & \text{Dove} \end{array} \\ \begin{pmatrix} V/2 - C & V \\ 0 & V/2 \end{pmatrix} \end{array}, \tag{3.1}$$

where the top row and left column denote "Hawk" choices by possessor and intruder, respectively.

Assume that $\bar{C} := C/V > 1/2$. This means that a "rational" actor would want to be a "Hawk" when the other one is a "Dove," and *vice versa.* Another way of putting this is to say that there are two non-cooperative equilibria in which the players choose different strategies. Since the "Dove" player receives a lower payoff than the "Hawk" player in these equilibria, there is a conflict in the assignment of roles.

Now assume that there is a very large population of players genetically conditioned to be either "Hawks" or "Doves," and who are randomly paired for interaction. There is an equal chance of being possessor and intruder in an interaction. The payoffs in (3.1) are to be interpreted as number of offspring. Over time, if the game is repeated for many generations, there would be a tendency for a genotype programmed for a dominant strategy to proliferate while the others would die out. Under the assumptions given above, however, there is of course no domination. Instead, there may be an equilibrium such that both genotypes are represented in the population. This occurs when the population proportions are such that the average offspring for both genotypes are equal, i e, when

$$n_H(V/2 - C) + n_D V = n_D V/2, \tag{3.2}$$

where n_H and n_D are the population proportions of "Hawks" and "Doves," respectively.

Condition (3.2) together with the constraint that the population proportions sum to 1 may be written as the equation system

$$\begin{pmatrix} V/2 - C & V/2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} n_H \\ n_D \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{3.3}$$

which has a solution under the assumptions given above. This solution is

$$n_H^* = \frac{1}{2\bar{C}}. \tag{3.4}$$

The solution is also non-trivial by virtue of the assumptions, i e, $n_H^\star \in (0,1)$. Furthermore, it is stable in the sense that a fluctuation in the population proportions at the equilibrium would be corrected over time, since, as can easily be checked, the genotype that is underrepresented in relation to the equilibrium proportions will have relatively more offspring.

The equilibrium proportion of "Hawks" will be "small," in the sense that $n_H^\star < 1/2$, if $\bar{C} > 1$, i e, if injuries are relatively serious.

3.2.2 "Hawk and Dove" with Recognizable Role Asymmetry

It is important to observe that, in the game as described, there is no asymmetry of payoffs even though the players are labelled differently. An interesting contribution of Maynard Smith's is the realization that a role asymmetry recognized by the participants may still, in fact, give rise to new possible solutions for the game.

We may now consider strategies which possibly prescribe different behavior dependent on whether the individual is possessor or intruder in a particular pairing. The original strategy set for each individual was $S = \{\mathrm{H(awk)}, \mathrm{D(ove)}\}$, so the new strategy set is $S \times S$, where I will consider the first element of a pair to be the conditional strategy on being possessor, and the second the strategy on being intruder. There are of course four different conditional strategies, of which two correspond to the role-independent "Hawk" and "Dove" strategies of the previous discussion.

Since, as assumed, a player is equally likely to be possessor or intruder, his expected payoffs conditional on the opponent's strategy may be summarized in the following matrix:

$$\Pi' = \begin{array}{c} \\ \mathrm{HH} \\ \mathrm{HD} \\ \mathrm{DH} \\ \mathrm{DD} \end{array} \begin{array}{c} \begin{array}{cccc} \mathrm{HH} & \mathrm{HD} & \mathrm{DH} & \mathrm{DD} \end{array} \\ \begin{pmatrix} \frac{1}{2}V - C & \frac{3}{4}V - \frac{1}{2}C & \frac{3}{4}V - \frac{1}{2}C & V \\ \frac{1}{4}V - \frac{1}{2}C & \frac{1}{2}V & \frac{1}{2}V - \frac{1}{2}C & \frac{3}{4}V \\ \frac{1}{4}V - \frac{1}{2}C & \frac{3}{4}V - \frac{1}{2}C & \frac{1}{2}V & \frac{3}{4}V \\ 0 & \frac{1}{4}V & \frac{1}{4}V & \frac{1}{2}V \end{pmatrix} \end{array}. \tag{3.5}$$

There are now equilibria in pure strategies, notably when everyone is of the HD type, i e, defends his first-claimer rights when possessor and respects first-claimer rights when intruder. This pattern is termed the *bourgeois* strategy by Maynard Smith. As Hirshleifer

[31] notes, however, the inverse principle, that the late-comer always gets the female, will do just as well in the game as specified, i e, all DH is also an equilibrium. This is of course because the underlying game is symmetric. There is no built-in advantage to being possessor. Yet, at least in human affairs, the rights of the possessor traditionally carry legal and perhaps moral significance, while the late-comer's do not. So perhaps something is missing here. In Section **3.3**, the possessor's first-comer advantage is reflected in the game structure.

3.3 Gold-Diggers

3.3.1 Players and Strategies

There is a continuum of individuals on the interval $[0, 1]$. All individuals are assumed to have the same von Neumann-Morgenstern utility function $U(x)$, where x is gold consumption. Furthermore, $U' > 0$ in the relevant interval.

At the beginning of the game period, each individual decides whether to make an investment in a mechanism of aggression, incurring a cost $c \geq 0$. This is the only decision to be made in the game. The set of pure strategies available to each individual is thus $S = \{I, N\}$, where I denotes a decision to make the investment and N a decision not to. The investment is in the nature of a commitment. It can be thought of as similar to a decision (perhaps made in a depressed frame of mind) to pay a professional "hit man" to kill you (or, in this case, someone else), regardless of whether you later change your mind, under certain objective circumstances at a future date. Or perhaps it is a time-bomb that cannot be disarmed once activated. This assumption has the same function here that the notion of genetically programmed behavioral commitment has in the biological game discussed above.

3.3.2 The "Veil of Uncertainty"

Having made the investment decision, each individual is randomly allocated one *site*, which may or may not turn out to contain a unit of gold. The probability density of gold at a particular site is α, with $\alpha \in [0, 1]$. Therefore, when the allocation is complete, a share α of the population will be *possessors* of gold, and a share $1-\alpha$ *non-possessors.*

Now $\sigma := \alpha(1-\alpha)$ is the variance of the single-sample distribution of sites containing a unit of gold. The variance σ is a measure of the uncertainty facing agents as to which role, possessor or non-possessor, they will occupy in the future game. The thickness of this "veil" behind which decisions must be made will turn out to influence what is to be considered a good strategy choice. However, this will also depend on the expected value α. Because of its quadratic nature, each value of σ is associated with two values of α, one of which is "high" in the sense of being larger than 1/2, and the other one "low." This turns out to complicate the total effect.

3.3.3 Interaction and Payoffs

Finally, individuals are paired randomly. In case a possessor is paired with a possessor, which will occur in a share α^2 of all pairs, both are assumed to exit the encounter with their original allocations, regardless of aggression investments. I am assuming that you cannot carry more than one unit of gold. The possible outcomes for each party in this particular pairing, viewed from the *ex ante* position, may thus be summarized as follows:

$$\Pi_{\mathrm{P,P}} = \begin{array}{c} \\ I \\ N \end{array} \begin{array}{c} \quad I \qquad\qquad N \\ \begin{pmatrix} U(1-c) & U(1-c) \\ U(1) & U(1) \end{pmatrix} \end{array}, \tag{3.6}$$

where the rows represent the arms situation of the individual to whom the payoffs accrue, and the columns that of the opponent he has been paired with.

Similarly, in case a non-possessor is paired with a non-possessor, which will be the case in a share $(1-\alpha)^2$ of all pairs, nothing happens. That is, we have that

$$\Pi_{\mathrm{Non-P,Non-P}} = \begin{array}{c} \\ I \\ N \end{array} \begin{array}{c} \quad I \qquad\quad N \\ \begin{pmatrix} U(-c) & U(-c) \\ U(0) & U(0) \end{pmatrix} \end{array}. \tag{3.7}$$

Finally, if a possessor is paired with a non-possessor, the non-possessor exits with the possessor's gold if the former has invested in the aggression mechanism while the latter has not. In case both have made the investment, let there be a probability β, not necessarily equal to .5, that the possessor gets to keep his unit of gold. The outcomes for a possessor in such a pair are therefore

$$\Pi_{\mathrm{P,Non-P}} = \begin{array}{c c} & \begin{array}{cc} I & N \end{array} \\ \begin{array}{c} I \\ N \end{array} & \begin{pmatrix} \beta U(1-c) + (1-\beta)U(-c) & U(1-c) \\ U(0) & U(1) \end{pmatrix} \end{array}, \tag{3.8}$$

while those of the non-possessor are

$$\Pi_{\mathrm{Non-P,P}} = \begin{array}{c c} & \begin{array}{cc} I & N \end{array} \\ \begin{array}{c} I \\ N \end{array} & \begin{pmatrix} (1-\beta)U(1-c) + \beta U(-c) & U(1-c) \\ U(0) & U(0) \end{pmatrix} \end{array}. \tag{3.9}$$

From the viewpoint of the *ex ante* position, i e, when the investment decisions are to be made, an individual will be the possessor in such a pair with probability $\alpha(1-\alpha)$, and the non-possessor with the same probability.

The game situation may now be given a description in terms of the payoff structure for a single individual, conditional on the choices of all others.

- The individual has made the investment.
  - All others have also made the investment. The expected payoff for each individual is then $\alpha^2 U(1-c) + (1-\alpha)^2 U(-c) + \alpha(1-\alpha)(\beta U(1-c) + (1-\beta)U(-c)) + \alpha(1-\alpha)((1-\beta)U(1-c) + \beta U(-c)) = \alpha U(1-c) + (1-\alpha)U(-c)$. Note that the (possibly) conditional probabilities of winning the fight cancel out, since you are equally likely to be in either role.
  - No one else has made the investment. There are three possible cases in which the individual exits with a unit of gold: When he is a possessor and is paired with another possessor (with probability α^2), when he is possessor and is paired with a non-possessor (with probability $\alpha(1-\alpha)$), and when he is non-possessor and is paired with a possessor (with probability $\alpha(1-\alpha)$). The expected payoff for the individual is then $(2\alpha - \alpha^2)U(1-c) + (1-\alpha)^2 U(-c)$.
- The individual has not made the investment.
  - All others have made the investment. The expected payoff for the individual is then $\alpha^2 U(1) + (1-\alpha^2)U(0)$.

– No one else has made the investment either. The expected payoff for each individual is then $\alpha U(1) + (1-\alpha)U(0)$.

Equivalently, the *ex ante* payoff structure may be summarized as follows:

$$\Pi = \alpha^2 \Pi_{\text{P,P}} + (1-\alpha)^2 \Pi_{\text{Non-P,Non-P}} + \alpha(1-\alpha)\Pi_{\text{P,Non-P}} + \alpha(1-\alpha)\Pi_{\text{Non-P,P}}. \quad (3.10)$$

The matrix columns now represent the choice of the opponent one happens to be paired with:

$$\Pi = \begin{array}{c} \\ I \\ N \end{array} \begin{array}{c} \begin{array}{cc} \qquad\qquad I \qquad\qquad & \qquad\qquad N \qquad\qquad \end{array} \\ \begin{pmatrix} \alpha U(1-c) + (1-\alpha)U(-c) & (2\alpha - \alpha^2)U(1-c) + (1-\alpha)^2 U(-c) \\ \alpha^2 U(1) + (1-\alpha^2)U(0) & \alpha U(1) + (1-\alpha)U(0) \end{pmatrix} \end{array}. \quad (3.11)$$

The elements of this matrix will be referred to as π_{ij}, with i, j the names of strategies in S.

3.4 Equilibrium

3.4.1 A Classification of Social States

Now denote by n_I the Lebesgue measure of the subset of individuals that have made the aggression investment, and by $n_N = 1 - n_I$ the corresponding measure for those who have not. Various values of n_I may now be given interpretations in terms of implied property rights systems.

The Hobbesian jungle (or "state of nature"). $n_I = 1$. All individuals invest in aggression. There is no respect for the rights of anyone.

"Leviathan." $n_I \in (0,1)$. Some individuals invest in aggression. Some individuals (a share $n_N \alpha(1-\alpha)$) will *ex post* turn out to have "respected" homestead rights as non-possessors paired with unarmed possessors. There is some transfer of resources from unarmed possessors to armed non-possessors. There is some enforcement by possessors of their own claims. There is of course no useful distinction here between *private* and *organized* enforcement of rights, such as in Hobbes' original discussion of Leviathan; the point is rather that there is resource loss because of the need for defense. For these purposes, n_I may be seen as a measure of the "size of government," since it measures the share of individuals who are willing to pay for defense of their rights of possession or to have resources transferred to them should they be unlucky allocation-wise.

Anarchic cooperation. $n_I = 0$. No one invests in aggression. There can be said to be voluntary respect for homestead rights. There is no social resource loss due to a need for defense.

3.4.2 An Equilibrium Concept

The predominant solution concept in noncooperative game theory is that developed by Nash [46]. This is also the notion of equilibrium referred to in some treatments of evolutionary games. However, this seems to involve a slightly incorrect use of the Nash idea.

Although the subject of much controversy and confusion, the Nash equilibrium is normally thought to have predictive significance when the players are rational, there is no communication, and the game structure is common knowledge among the players. Clearly, this is the kind of situation that Nash was addressing. The equilibrium is thought to be interesting because, when unique, it specifies what *rational* players would do under the described circumstances.

Evolutionary game theory, however, dispenses with the idea of rationality—entirely, in the case of explanations of animal behavior, and to a large degree, in applications to processes of bounded rationality and imitation in human societies. For this reason, the use of an equilibrium concept which is based on rationality seems somewhat unsatisfactory.

I will therefore instead propose a notion of equilibrium which is a description of objective circumstances, entirely devoid of references to the behavior of the players, or indeed the information they possess about the game. The reader who now expects something looking radically different from the Nash concept will be disappointed, however, since all Nash equilibria are also equilibria in the sense to be defined below—and the converse is also true. Only the *rationale*, which will become clearer in Section 3.5, is different.

Definition 3.1 *A set of strategy choices such that unilateral deviation on the part of a single agent would cause his expected utility to be the same or lower will be termed an* equilibrium.

Furthermore:

Definition 3.2 *An equilibrium such that all agents make the same strategy choice will be termed a* monomorphic equilibrium, *or* monomorphism, *while an equilibrium such that two different strategy choices are represented in the population will be termed a* bimorphic equilibrium, *or* bimorphism.

3.4.3 Equilibria

I will consider only equilibria in pure strategies. There can, of course, depending on parameter values, be at most two different monomorphic equilibria in pure strategies simultaneously in existence in the game.

The "Hobbesian jungle" (HJ) equilibrium exists when making the aggression investment is a best response to everyone else making the investment, i e, when

$$\pi_{II} \geq \pi_{NI}, \tag{3.12}$$

which is equivalent to

$$\alpha U(1-c) + (1-\alpha)U(-c) \geq \alpha^2 U(1) + (1-\alpha^2)U(0). \tag{3.13}$$

The "anarchic cooperation" (AC) equilibrium exists when

$$\pi_{NN} \geq \pi_{IN}, \tag{3.14}$$

i e, when

$$\alpha U(1) + (1-\alpha)U(0) \geq (2\alpha - \alpha^2)U(1-c) + (1-\alpha)^2 U(-c). \tag{3.15}$$

Proposition 3.1 *The situation with $n_I = 0$ (anarchic cooperation) Pareto- dominates $n_I = 1$ (the Hobbesian jungle).*

Forming the difference between the expected individual payoffs in the two situations, we find that

$$\pi_{NN} - \pi_{II} = \alpha(U(1) - U(1-c)) + (1-\alpha)(U(0) - U(-c)) > 0, \forall\alpha \in [0,1], \quad (3.16)$$

by monotonicity. This means that the "thickness" of the veil of uncertainty (i e, the value of σ) has no bearing on the relative equilibrium status of the AC and HJ situations in the hypothetical cooperative game suggested by Brennan and Buchanan.

There are now four different conceivable cases: Both, none, or only one of the equilibria may exist.

The two monomorphic equilibria exist simultaneously when conditions (3.12) and (3.14) both hold. This implies, however, that the individuals are risk lovers or risk neutral.

Proposition 3.2 *If individuals are risk averse, i e, have strictly concave expected utility functions, then if any monomorphic equilibrium in pure strategies exists, it is unique.*

To prove this, assume that both the HJ and AC equilibria exist simultaneously. This means that $\pi_{II} \geq \pi_{NI}$ and $\pi_{NN} \geq \pi_{IN}$. This implies, by summing the inequalities and rearranging, that

$$\pi_{II} - \pi_{NI} + \pi_{NN} - \pi_{IN} \geq 0 \quad (3.17)$$

or

$$\sigma(U(1) - U(0) + U(-c) - U(1-c)) \geq 0. \quad (3.18)$$

Since $\sigma \geq 0$, this would imply that $U(1) - U(0) \geq U(1-c) - U(-c)$. But then U cannot be a strictly concave function, since in general if $f : R \rightarrow R$ is a strictly concave function, then, for any x, x' such that $x' > x$, and for any $\delta > 0$, we have that $f(x' + \delta) - f(x') < f(x + \delta) - f(x)$. Therefore any monomorphic equilibrium in pure strategies is unique.

When only one of (3.12) and (3.14) holds, however, one strategy strictly dominates the other, regardless of risk attitudes. In case (3.12) holds and (3.14) does not, the game has the familiar "Prisoners' Dilemma" (PD) structure, where the dominant strategy leads to an inefficient equilibrium.

In particular, it is enlightening to consider for a moment the special case that arises when individuals are risk neutral, i e, have linear expected utility functions. We then have that (3.12) reduces to

$$c \leq \sigma. \tag{3.19}$$

Conversely, the AC equilibrium exists when

$$c \geq \sigma. \tag{3.20}$$

When (3.20) holds strictly, what we have is that happy (but seldom discussed—unless one counts extremely naïve readings of Adam Smith) thing, a game situation where adopting the efficient strategy is a dominant way of behaving. I hereby christen it the "Inverse Prisoners' Dilemma" (or DP).

When we have that $c = \sigma$, all possible outcomes are equal and any population proportions constitute an equilibrium.

Now assume a uniform distribution over the parameter space, and let $c < 1$. For a given value of α, the set of parameter values for which "anarchic cooperation" is an equilibrium is $\{c \in (0,1) : c \geq \alpha(1-\alpha)\}$. The measure of this set is simply $1-\sigma$. When uncertainty is highest, i e, when σ is at its maximum at 0.25, the "anarchic cooperation" set is smallest. In other words, the lesser is uncertainty, and the greater the cost of using violence, the larger is the scope for spontaneous cooperation.

The intuition for this result is very straightforward. Having made the investment will be useful only in possessor versus non-possessor conflicts. The probability of being one of the parties in such a pairing is 2σ. The larger σ, and therefore this probability, and the smaller c is, the greater is the likelihood of everyone deciding that the investment is a good idea.

The uncertainty σ is, of course, at a maximum when the states of being possessor and non-possessor are equally likely. When α is small, a given player is most likely to be going to be "poor," and most of the other players are also going to be poor, so mixed conflicts, where something can be gained, will be rare. Similarly, when α is large, an individual will most likely be "rich," and this goes for his potential opponents as well, so that the investment in aggression will be unlikely to seem a worthwhile project.

Returning to the more realistic case of risk aversion, when neither condition (3.12) nor (3.14) holds, there is no monomorphism. However, there may be a *bimorphism*, or "Leviathan" equilibrium, that is, population proportions $(n_I^\star, n_N^\star)$ such that expected payoffs are equalized over the strategies[3], so that

$$n_I^\star \pi_{II} + n_N^\star \pi_{IN} = n_I^\star \pi_{NI} + n_N^\star \pi_{NN}. \tag{3.21}$$

We may now prove the following general existence result.

Proposition 3.3 *A monomorphic or bimorphic equilibrium for the game exists.*

We already know that a monomorphic equilibrium may exist. Equation (3.21) together with the condition that the equilibrium population proportions sum to 1 may be written as the equation system

$$\begin{pmatrix} \pi_{II} - \pi_{NI} & \pi_{IN} - \pi_{NN} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} n_I^\star \\ n_N^\star \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.22}$$

which has the solution

$$n_I^\star = \frac{\pi_{NN} - \pi_{IN}}{(\pi_{II} - \pi_{NI}) + (\pi_{NN} - \pi_{IN})}. \tag{3.23}$$

Assuming that neither monomorphism exists, it can easily be checked that $n_I^\star \in (0,1)$, which guarantees the existence of at least one equilibrium.

Correspondingly, we have that

$$n_N^\star = 1 - n_I^\star = \frac{\pi_{II} - \pi_{NI}}{(\pi_{II} - \pi_{NI}) + (\pi_{NN} - \pi_{IN})}. \tag{3.24}$$

Proposition 3.4 *The bimorphism is inefficient for* $\alpha \in (0,1)$.

This can be proved by noting that the expected payoff to each individual is equal to $n_I^\star \pi_{NI} + n_N^\star \pi_{NN}$. We have that $\pi_{NN} \geq \pi_{NI}$ since

$$\pi_{NN} - \pi_{NI} = \sigma(U(1) - U(0)) \geq 0 \tag{3.25}$$

[3] It should be noted that this definition of an equilibrium in fact has three different interpretations. The one used here is that of a bimorphism in pure strategies. It could also be a monomorphism in mixed strategies, with $(n_I^\star, n_N^\star)$ the probabilities assigned by all players to strategies I and N, respectively. Finally, it could define a truly polymorphic situation where individuals in fact use different mixed strategies, but where the aggregate probabilities are $(n_I^\star, n_N^\star)$.

by monotonicity. Since the expected payoff at the bimorphism is a convex combination of π_{NN} and π_{NI}, it is equal to or lower than π_{NN}. Each individual would be as well off if all chose strategy N for $\alpha \in \{0,1\}$, and better off for $\alpha \in (0,1)$.

If the cases where one of (3.12) and (3.14) holds are included, we get the domination equilibria as limit cases of the bimorphism, so that $n_I^\star, n_N^\star \in [0,1]$. Therefore, only the behavior of the bimorphism (3.23) needs to be studied.

Fairly unsurprisingly, the larger the cost of the investment as a share of what can be gained, the smaller is the equilibrium arsenal.

Proposition 3.5

$$\frac{\partial n_I^\star}{\partial c} < 0$$

This is found by differentiating (3.23). Let $N = \pi_{NN} - \pi_{IN}$ (not confusing this N for the strategy of the same name) and $D = \pi_{II} - \pi_{NI} + \pi_{NN} - \pi_{IN}$. Then

$$\frac{\partial n_I^\star}{\partial c} = \frac{N_c D - D_c N}{D^2}, \tag{3.26}$$

where the subscripts denote derivatives. We have that $N, D < 0$ by the non-existence of monomorphisms, $N_c = (2\alpha - \alpha^2)U'(1-c) + (1-\alpha)^2 U'(-c) > 0$, and $D_c = \sigma(U'(1-c) - U'(-c)) < 0$ by strict concavity, which makes the expression negative.

The role of uncertainty in this general context is partly indeterminate, i e, it will depend on the specific shape of the utility function. We might want to ask, for instance, what the effect of a small change in σ due to a continuous adjustment of α would be on $n_I^\star$. That is, by excluding sudden jumps in α, we could consider it locally a function of σ and differentiate.

Proposition 3.6

$$\frac{\partial n_I^\star}{\partial \sigma}\Big|_{\alpha<1/2} \geq 0$$

We have that

$$\frac{\partial n_I^\star}{\partial \sigma} = \frac{(N_\alpha - D_\alpha)(\pi_{NN} - \pi_{IN}) + N_\alpha(\pi_{II} - \pi_{NI})}{D^2}\frac{\partial \alpha}{\partial \sigma}. \tag{3.27}$$

For $\alpha < 1/2$, we have that $N_\alpha = U(1) - U(0) - 2(1-\alpha)(U(1-c) - U(-c)) < 0$, $N_\alpha - D_\alpha = 2\alpha(U(1) - U(0)) - (U(1-c) - U(-c)) < 0$, and $\partial\alpha/\partial\sigma \geq 0$. That is, for a situation with relatively poor prospects, an increase in uncertainty (due to an increase in the expected value α) would make the equilibrium size of "Leviathan" larger. For $\alpha \geq 1/2$, nothing can be said without further assumptions about the functional form of U.

3.5 Evolutionary Dynamics

Now imagine this interaction structure is repeated over many periods—that is, the same agents are in each period confronted with a new field of sites, all having the same frequency of gold sites, and an identical investment decision to be made. In keeping with the evolutionary analysis of Maynard Smith's, one might want to know the likelihood of the static equilibria occurring as steady states in such a dynamical system.

Explanations based on genetic selection, i e, by means of differential reproductive capacity in agents, have a dubious status in the social sciences. Assuming, however, that agents are ***boundedly rational***, and update their strategy choices from period to period based on observed results, we get a process of ***cultural evolution*** similar to the genetic evolution primarily studied by evolutionary biologists.

When the expected values of investing and not investing are equal, no agent who has made one type of decision could observe agents who have made the other do better on average. A reasonable updating process therefore has the bimorphism defined above as a fixpoint. Without assuming any specific functional form for this process, one necessary condition for stability of its fixed points is evident already: A small deviation from the equilibrium population proportions (such as might occur due to, for example, an influx of new players having different proportions of aggressors and non-aggressors) must not make the expected value of the strategy that is then over-represented in relation to the equilibrium larger than that of the other.

Note that the bimorphic equilibrium in pure strategies is mathematically identical to a symmetric equilibrium in mixed strategies for a 2×2 symmetric bimatrix game. Since this is the case, the stability condition described above is equivalent to both

Selten's [59] notion of *perfection* (or *"trembling-hand"* perfection) and Maynard Smith's of *evolutionary stability.*

Let $E_I(n_I) = \pi_{II}n_I + \pi_{IN}n_N$ and $E_N(n_I) = \pi_{NI}n_I + \pi_{NN}n_N$ be the respective expected values of the two strategies. We have that $E_I(n_I^*) = E_N(n_I^*)$. For stability in the sense described above, it must be the case that for a small perturbation ϵ, $E_I(n_I^* + \epsilon) < E_N(n_I^* + \epsilon)$ for $\epsilon > 0$, and $E_I(n_I^* + \epsilon) > E_N(n_I^* + \epsilon)$ for $\epsilon < 0$. Since

$$E_I(n_I^* + \epsilon) - E_N(n_I^* + \epsilon) = (\pi_{II} - \pi_{NI})\epsilon + (\pi_{NN} - \pi_{IN})\epsilon \qquad (3.28)$$

is, if (3.12) and (3.14) do not hold, negative for $\epsilon > 0$ and positive for $\epsilon < 0$, this is in fact the case.

3.6 Concluding Remarks

Most of actual economic and social life does not have the character of "gold-digging" in the sense discussed above. For instance, there is production activity. Although the model might at first glance seem highly contrived, it could perhaps nonetheless say something about property rights and redistribution in actual societies. To begin with, the model's variables are related to observable measures. Note, for instance, that the variance σ may also be seen as an "inequality" measure for the income distribution, directly related to the entropy measure discussed by Theil [63]. (The use of the information-theoretical term "entropy," which is equivalent to "uncertainty," by Theil should be especially intriguing to contractarians concerned with decisions made behind "veils.")

Although it is not immediately obvious how one would best measure the degree of respect for individual property rights in talent and its product (although transfer payments as a share of GNP might be one example), the above discussion would lead us to expect poor societies with a large degree of pre-tax income inequality to also have a lesser degree of such respect, assuming that the costs of using violence are equal across societies.

Chapter 4

Money

4.1 Introduction

What is known as "money" in present-day societies has several distinct functions. It is

- a store of value,
- a unit of account, and
- a commonly accepted medium of exchange (CAMOE).

A recent discussion (with important forerunners, see Cowen and Kroszner [11]) questions the theoretical necessity and efficiency of these functions all being performed by the same good. In any case, care should be taken to point out which particular functions are intended to be covered when discussing the good "money." Different definitions of what is to be considered "moneyness" may have radically different implications for modeling results.

In particular, the origins and peculiar functioning of a *medium of exchange* have often been neglected or misunderstood. What a CAMOE accomplishes has, of course, always been recognized. Money solves the problem of the "double coincidence of wants" necessary for trade to take place in a pure barter economy. Until very recently (see Jones [32]), Kiyotaki and Wright [35], [34]), and Oh [51])), however, the modeling of this important function has eluded formal neoclassical theory. One reason for this lack is noted by Niehans [48]. This is the more fundamental problem that we do not even have a formal understanding of the workings of an actual auctioneer-less barter economy.

An example of confusing the CAMOE function with the store-of-value function is found in one of the claims of the "legal restrictions" school of monetary theorists. According to Wallace [71] (p 1):

> An obvious instance of a paradoxical pattern of returns among assets is the coexistence of, on the one hand, U S Federal Reserve Notes (U S currency) and, on the other hand, interest-bearing securities that are default-free. .... Examples of such securities are U S savings bonds and Treasury bills. Our first task is to identify the features of these securities that prevent them from playing the same role in transactions as Federal Reserve notes. For if they could play that role, then it is hard to see why anyone would hold non-interest-bearing currency instead of the interest-bearing securities.

Wallace goes on to suggest that the use of non-interest-bearing money, i e, ordinary currency, to carry out transactions, would disappear in favor of interest-bearing instruments—such as government bonds—without legal tender laws. It is also apparent from the context that it is meant seriously as a statement about reality—not just about Walrasian general equilibrium models.

The Wallace hypothesis could be restated in the following form. It is in the interest of individuals to hold assets which yield positive rather than zero interest. Therefore, legal restrictions as to what may be used in exchanges must be what creates a demand for ordinary currency.

This chapter is an attempt to bring a better understanding of the problem of medium-of-exchange choice in decentralized monetary anarchy to bear on this claim. This leads to a theoretical rejection of the claim, for the following reasons.

- A CAMOE is an institutional solution to a coordination problem. The crucial feature of such problems is that no action is individually dominant independently of what other people do. Therefore, the interest yields on various instruments are in an important sense irrelevant for their CAMOE potential.

- Indirect exchange in general, however, individually dominates direct exchange. Government intervention is not necessary to convey the CAMOE property on *something.*

In Section 4.2, support is provided for the latter point by reference to 19th century Austrian economist Carl Menger's [44] theory of the spontaneous origin of money. Section 4.3 then goes on to argue that for coordination conventions in general, and money in particular, switches to efficiency are not automatic.

Section 4.4 presents a simple, static, two-instrument model of the problem of media of exchange choice in a situation without central enforcement. It is shown to be unlikely that the strategy of holding the "interest-bearing" asset will dominate in this coordination game, since there will typically be no dominance at all. The game also has an equilibrium where the uses of the two instruments coexist.

In Section 4.5, the problem of embedding the static game in a dynamic context is discussed. The familiar "supergame" approach is rejected, since it seems particularly unsuited to coordination problems, in favor of an adaptive, or evolutionary, approach.

The dynamized game studied in Section 4.6 is related to the kind of evolutionary processes originally proposed by Maynard Smith and Price [42] for the study of animal behavior, though the mechanism of selection is here the imitation of successful practices rather than the differential reproductive capacities of individuals. It is demonstrated that even if the use of the non-interest-bearing medium of exchange is assumed to be inefficient, it may still persist in the face of feasible alternatives, and that the "critical mass" of adherents to the efficient strategy necessary for a switch to occur is not necessarily "small," or even dependent on the relative efficiencies of the two conventions.

Finally, the general problem of media of exchange in macro-models is discussed from the strategic interaction point of view.

This chapter addresses only the problem of the emergence of media of exchange conventions in monetary anarchy. Other important issues central to a more broadly defined "legal restrictions" school of monetary economics are not dealt with here. An example would be the macro-stability of *laissez-faire* money.

4.2 The Emergence of Media of Exchange

Early attempts at explaining the existence of CAMOE were typically "legal restrictions" theories. They assumed that both the institution itself and the stability of a currency's value could only emerge and persist because of governmental enforcement. "Social contract" theories provide a slightly more sophisticated, although obviously factually untrue, argument. Witness, for instance, Locke's [39] praise for "the *invention* of money, and the tacit *agreement of men* to put a value on it" (emphasis added). This naïve view is still prominent today.

Carl Menger pointed out that the social institution of a CAMOE can be explained by reference only to the decentralized actions of self-interested individuals. Invoking the *deus ex machina* of state power is not necessary. From a purely methodological point of view, as noted by, among others, Nozick [49] and Ullmann-Margalit [65] such *invisible hand* explanations are more satisfying than others. They are more parsimonious, particularly as regards assumptions about the information available to a central planner.

Menger considers a pure barter economy where there is some degree of division of labor. Agents do not themselves produce all the goods they ultimately wish to consume. Every week or so they venture out in search of someone to trade with. Assume agents initially do not trade until they have found someone who has something they want and who is also willing to accept what they offer in return.

Now some enterprising individual E may realize that his search time can be reduced. This is done by trading his original basket for one which he does not wish to consume, but contains goods for which consumption demand is more wide-spread. Let there be a good M which many people desire for ultimate comsumption. Then the probability of finding someone who has M and wants what E has and later finding another agent who wants M and has what E wants may be greater than the probability of finding a "first-best" exchange opportunity. Let the *saleability* of a good or basket of goods mean the probability of finding someone who wishes to acquire it. Then we may say that M is more saleable than E's original basket. Assume this to be the case.

E in accepting M has discovered indirect exchange, and is using M as a medium of exchange. The story of the emergence of money now proceeds as follows.

1. E's use of M as a medium of exchange lowers his total search time. This leaves him more time each week to pursue other interests.

2. Other agents in E's immediate environment may be assumed to notice this improvement in E's conditions.

3. Assume they can correctly identify the source of E's improved want-satisfaction as the use of a medium of exchange. They will then want to imitate this practice.

4. Assuming that M was a highly saleable good when E started accepting it in indirect exchange, his doing so will increase its saleability even more. A transactions demand is added to final consumption demand for the good. Therefore, imitating agents will find M even more attractive.

5. The choice of M by the second generation further increases its saleability.

This cumulative process may converge to M being the single good with supreme saleability. It would then be the generally accepted medium of exchange. Starting from less severe premises, several goods might acquire this property. In the real world, of course, there are no generally accepted media of exchange, but local CAMOE.

Two features of this story that are of particular interest to the present discussion are the following.

- Acquiring highly marketable goods in the absence of legal restrictions not for their direct consumption-value but to use them in indirect exchange is individually rational. This amounts to saying that a CAMOE is an equilibrium in a non-cooperative game, rather than being a cooperative institution. A non-cooperative equilibrium is self-enforcing. By definition nothing can be gained by unilaterally deviating from an equilibrium strategy. An external enforcement agency, e g, government, is not necessary. Observe, however, that this is a theory of commodity money. A pure token money is one backed by nothing but expectations as to its common acceptance. Whether token money could ever emerge spontaneously in the decentralized manner suggested is a tricky issue. Once such an institution is in existence, though, we would expect it to be able to persist even if legal enforcement of it is abolished (see White [73]).

- The Mengerian account suggests that global knowledge of the game structure and unbounded rationality in agents is not necessary for the emergence of money. The self-interested practices of a few innovative individuals are diffused through an evolutionary process of imitation of successful behavior.

The Mengerian theory of "money" is a theory of a good that solves the double-coincidence-of-wants problem in a barter economy. An alternative version of this statement could serve as a definition of money. A money is whatever good an individual agent can reasonably expect will be commonly accepted in indirect exchange, i e, a CAMOE. This subjectivist definition of money (see White [74] for an in-depth discussion) radically departs from the still common textbook alternative. The latter locates moneyness in, on the one hand, the property of being legal tender, and, on the other, certain physical attributes of a good, such as its divisibility, storability, transportability etc. While these attributes are often present, what serves as a CAMOE in a particular society at a particular time cannot be identified by looking for the good that satisfies a list of physical criteria best. Again, a CAMOE is *whatever* is commonly used for indirect exchange. Still, there is a long tradition in monetary theory of discussing at length the "optimal" physical properties of a money. This is a curious feature of a discipline that has adopted a subjectivist viewpoint in most other areas. Money is not really all that different.

4.3 Media of Exchange as Conventions

The "legal restrictions" idea has been subjected to various criticisms (see, e g, White [72] and O'Driscoll [50]). These more or less stay within the same general sort of contemporary macro-model framework that Wallace discusses. This framework may not be adequate, however. The Wallace claim addresses issues of monetary anarchy. This is done without a real theory of a situation where agents are (in principle) at liberty to use whatever takes their fancy as a medium of exchange.

I would like to suggest that a convincing theory of a medium of exchange must first be a theory of an actual exchange process. This is a setting where no central coordinator is present. Individuals engage in specialized production and then have to search for someone to trade with. In fact, indirect exchange has no meaning except in such a world. Many standard neoclassical texts start out by defining money as the most saleable good. But saleability is a meaningless concept in a general equilibrium world.

What is meant is the fact that all individuals hold a special good "money." That this is not the same thing is easily checked by asking oneself which is the second most saleable, third, etc, in general equilibrium. Such a distinction makes no sense in that context.

Consider the decision problem facing a single individual when there is no enforced rule as to what can be used as a means of payment. An interest-bearing instrument can for our purposes be defined as something that always yields a payoff, regardless of whether it can be traded or not. A first implicit assumption of Wallace's now seems to be that interest-bearing instruments in some sense individually dominate non-interest-bearing money as as media of exchange. This does not necessarily follow simply from the definition.

Note that the value of any good intended to be used as an instrument for indirect exchange is dependent on expectations concerning the commonness of its acceptance. There is a network externality aspect involved. The larger the "installed base," or population proportion that accepts a particular good X in trade, the more attractive it is as a medium of indirect exchange. The probability of finding somebody who has something that you value directly and is also prepared to accept X is greater. Conversely, if nobody wishes to accept X in exchange, it is of no use to you *as a medium of exchange.*

Whether the medium-of-exchange candidate X is interest-bearing or not is irrelevant in this case. The unlikely exception would be when the expected yield from holding X is larger than any payoff you might expect to gain from participating in trade. Switching to speaking Esperanto, which supposedly is a simpler and more logical language, does not pay off if the person you are speaking to does not know Esperanto. In exactly the same way it is not necessarily true that interest-bearing assets are better than others as media of exchange to the individual agent. Like natural language, the institution of a CAMOE is a convention in the sense of Lewis [38], i e, an equilibrium solution to a coordination game.

This immediately points to a problem with a second implicit assumption of the Wallace hypothesis. This is the (admittedly common) notion that efficient practices always become prevalent in situations where there is no outside interference. Even if we assume that everybody would be better off if everybody used bonds as a medium of exchange, without legal enforcement agents face a coordination problem. In a decentralized environment, where individuals are unable to all get together and explicitly agree on a CAMOE, an agent is instead forced to base his decision on the experience of himself and those with whom he interacts as to what practices happen to work out better than others. It is never the case that you choose yourself what should conventionally be considered a medium of exchange. Furthermore, whether the evolutionary selection process of imitation of successful behavior converges to "efficiency" will depend on where it started from. If the means of transaction is deregulated when an "inefficient" instrument is commonly used, the inefficient convention may very well persist because of wide-spread expectations of its persistence.

As a simple example, let $V_M(n_M)$ be the expected utility of holding ordinary currency for transactions purposes when the population proportion doing so is n_M. V_M is, of course, a non-decreasing function of n_M. If the only alternative is holding bonds, then the population proportion holding bonds is $n_B = 1-n_M$, so the expected utility of bonds may be written $V_B(n_M)$, a non-increasing function. Agents are for simplicity assumed to hold either currency or bonds, but not both. Incidentally, to hold something you must acquire it from somewhere. A decision to hold bonds is therefore equivalent to a decision to accept them in exchange.

Holding bonds will be a dominant strategy only when $V_B(n_M) > V_M(n_M), \forall n_M$. This condition would imply that $V_B(1) > V_M(1)$, or that holding bonds only for the interest paid on them, when nobody accepts them in exchange, is more profitable than taking part in exchanges. This is clearly an unreasonable assumption.

Ruling this case out, both $n_M = 1$ and $n_M = 0$ are non-cooperative equilibria. In both cases an agent would like to conform to whatever everyone else is doing. Now assume that universal use of bonds would in fact be efficient, i e, that $V_B(0) > V_M(1)$. The payoff functions summarize the expected outcome of a search process, where an individual faces an entire economy of potential traders. Most of these he has never met before. It seems outrageous to assume that individuals would trust in the rationality of strangers to such a degree that they would automatically coordinate themselves to the efficient solution. Instead, individuals would switch to bonds only when a group of bond-holders sufficiently large to make the expected value of holding bonds greater

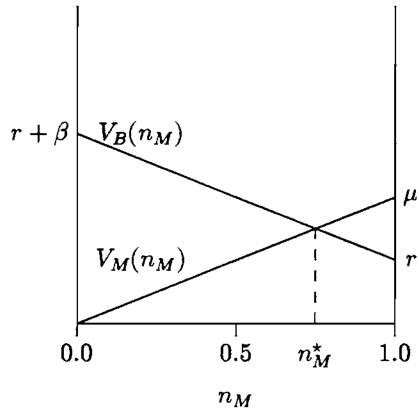

Figure 4.1: Two instruments with linear payoffs.

than that of holding currency has somehow been assembled.

Assuming the V_M and V_B functions to be continuous, then in general there will be some $n_M^\star$ such that $V_M(n_M^\star) = V_B(n_M^\star)$. If the functions are strictly monotonic, it is trivial to prove that $n_M^\star$ is unique. It has the property that $V_M(n_M) > V_B(n_M)$ if $n_M > n_M^\star$, and $V_M(n_M) < V_B(n_M)$ if $n_M < n_M^\star$. Then $n_B^\star = 1 - n_M^\star$ is the "critical mass" necessary for a switch from currency to bonds.

The relative efficiency of the two equilibria does not in general determine the size of $n_B^\star$. It could be the case that the critical mass is large, for instance in the sense of being greater than 1/2, even though using bonds is the efficient equilibrium.

In the special case where the payoff functions are linear, $n_B^\star$ will be less than 1/2 if using bonds is efficient. This case is illustrated in Figure 4.1, with $V_M(n_M) = \mu n_M$ and $V_B(n_M) = r + \beta n_B = r + \beta - \beta n_M$, where μ, β are parameters and r is the interest paid on bonds. Efficiency of bonds implies $r + \beta > \mu$.

In the following, a more detailed model is presented.

4.4 The Static Model

Consider an economy consisting of a continuum of individuals who are ordered on the interval $[0, 1]$ and have von Neumann-Morgenstern expected utility functions. At the beginning of the game period, each agent decides whether to hold a unit of good M

("gold") or a unit of good B (a "fruit tree"), two goods that are assumed to have historically been used as media of exchange. Holding one of these goods is assumed to be equivalent to a decision of offering and accepting the same good in indirect exchange. Since in order to hold a good, it must be accepted in exchange, these decisions are essentially the same. A number of other goods are produced in the economy. At the end of the game period, holders of a unit of B are paid an amount r.

Let n_M be the population proportion (or Lebesgue measure) of holders of good M, and n_B that of holders of B. Since there is an infinite number of agents, the decision of a single agent *given* the decisions of all the others will not measurably affect the population proportions.[1]

During the game period, the agents are paired randomly for exchange transactions where the economy's other goods, which in different ways enter directly into agents' utility functions, are traded. This exchange process will be subsumed in the following. The intention here is not to construct a complete, detailed model of an exchange economy, but to study specifically the coordination problem aspects of media of exchange choice.

A prerequisite for a successful transaction is assumed to be that both parties have chosen the same good to hold and accept in exchange. The payoff to an agent holding good M who is paired with an agent also holding good M is assumed to be a continuous, differentiable, non-decreasing function $f(n_M)$. In addition to the market value of good M, which, again without going into the details of the exchange process, can reasonably be taken to be at least not decreasing in the number of users, the f function reflects intrinsic properties, such as its degree of divisibility.

An agent holding M and accepting only M in exchange who is paired with an agent holding B receives a zero payoff. This should be considered a matter of convenient scaling of the payoffs only. Assuming that good M has some value apart from its use in transactions would not change the results, while, indeed, such an assumption clearly

[1]This will be a crucial, though hopefully not unreasonable, assumption for the following discussion. The game with a finite number of agents may have very different properties from those studied here. Observe, however, that the assumption may be thought of as only being about the *expectations* of a single individual as to the influence of his decision.

increases the plausibility of the Mengerian commodity money story.

Holders of good B receive r at the end of the game under any circumstances, but in addition receive $g(n_M)$, a continuous, differentiable, non-increasing function, in the case of having been paired with another B-holder. Finally, let $f(0), f(1), g(0), g(1), r > 0$.

The expected payoff of an agent holding good M is now $n_M f(n_M)$, while that of an agent holding good B equals $n_M r + n_B(r + g(n_M))$. It should be obvious that the M strategy can never dominate the B strategy, in the sense that holding M will always yield a higher payoff no matter what the population proportions are. There will therefore be no dominant strategy in this problem unless the expected payoff from holding good B is always greater than that from holding good M, or

$$r > n_M f(n_M) - n_B g(n_M), \forall n_M. \tag{4.1}$$

In particular, this would imply that $r > f(1)$, or that holding B just for the interest paid on it, when nobody accepts it in exchange, is more profitable than taking part in exchanges. If this extreme case is ruled out, the problem is one of coordination.

In the absence of government-imposed restrictions on the medium of exchange, and assuming that $f(0) > r$, the static game has three non-cooperative equilibria when allowing only pure strategies. Two of these have all agents choosing to hold the same good, each receiving a payoff of $f(1)$ in the case of good M being the unique CAMOE, or of $g(0) + r$ in the case of good B being the standard.

The third possible equilibrium has the uses of both instruments coexisting in the population. It will therefore be referred to as the *bimorphic equilibrium*, or *bimorphism*. It occurs where the expected payoff from holding a unit of good M equals that of holding a unit of good B, i e, where $n_M f(n_M) = n_M r + n_B(r + g(n_M))$. Equivalently, the bimorphic equilibrium is a fixed point $n_M^\star \in (0, 1)$ of the mapping

$$\phi(n_M) = \frac{r + g(n_M)}{f(n_M) + g(n_M)} \tag{4.2}$$

Proposition 4.1 *There exists a bimorphic equilibrium $n_M^\star \in (0, 1)$.*

Since $f(n_M)$ and $g(n_M)$ are continuous functions, and since the other assumptions imply that $\phi(n_M) \in [0, 1]$, $\phi(n_M)$ is a continuous function from the unit interval to itself. Then Brouwer's well-known fixed point theorem applies. Inspection of (4.2) also reveals that 0 and 1 cannot be fixpoints.

Proposition 4.2 *The bimorphic equilibrium $n_M^\star$ is unique.*

To prove this, assume another fixpoint $n_M^{\star\star} \neq n_M^\star$ exists. Suppose we have that $n_M^{\star\star} > n_M^\star$. Then $n_M^{\star\star} f(n_M^{\star\star}) > n_M^\star f(n_M^\star)$. Since expected payoffs are equal at the equilibrium, it must then also be the case that $n_B^{\star\star} g(n_M^{\star\star}) > n_B^\star g(n_M^\star)$. But this is impossible. Similarly, we cannot have that $n_M^{\star\star} < n_M^\star$. Therefore, $n_M^\star$ is unique.

Proposition 4.3 *The bimorphism $n_M^\star$ is inefficient.*

We have that $n_M^\star f(n_M^\star) < f(n_M^\star) \leq f(1)$ for any $n_M^\star$, because of monotonicity.

In addition to this inefficiency, a version of the well-known *perfection* concept suggests that the bimorphism is not particularly interesting in the static game. Consider an arbitrarily small perturbation of the population proportions at the bimorphism, such that $n_M = n_M^\star + \epsilon$. It follows immediately from the logic used in the uniqueness proof that one of the instruments will now perform better than the other, so that some individuals would want to switch. Only the two "pure" equilibria, or conventions, are stable in the face of such a test. This criterion is similar to Maynard Smith's *evolutionary stability*. The latter is not directly applicable since it is only defined for symmetric equilibria. Also, the term "evolutionary" should perhaps be saved for the explicitly dynamic version of the game studied in Section 4.6 below. The unstable bimorphism will be of interest there, since it can be given a "critical mass" interpretation.

4.5 A Positive Approach to Dynamics

We may now wish to study the behavior of the model as the static situation is repeated over many time periods. A familiar approach would be to consider the infinitely repeated game from an *ex ante* viewpoint, and, assuming agents have infinite planning horizons, ask which sequences of actions constitute equilibria for the "supergame." That is, we would now consider expected payoff functions of the form

$$\sum_{t=0}^{\infty} \beta_i^t U_i(s_{it}, n_{Mt}) \tag{4.3}$$

where $s_{it} \in \{M, B\}$ is the action of individual i at time t, n_{Mt} is the population proportion of M-holders at t, β_i is a discount factor, and

$$U_i = \begin{cases} n_{Mt} f(n_{Mt}), & s_{it} = M; \\ n_{Bt} g(n_{Mt}) + r, & s_{it} = B. \end{cases} \tag{4.4}$$

Such an approach has several deficiences. To begin with, we may trivially identify a very large family of equilibria for this "supergame": All sequences of n_{Mt} such that $n_{Mt} \in \{0, n_M^*, 1\}, \forall t$ (with n_M^* defined as in Section 4.4 above), are non-cooperative equilibria. That is, since we are dealing with a coordination problem, any sequence of actions is an equilibrium for the "supergame" as long as all agents choose the same action, or different actions in the expectation-equalizing combination, in a given time period.

We also have not assumed anything about the relative size of payoffs at the two "pure" equilibria, so there is no ground for making a distinction on an efficiency basis. The predictive power of this approach is therefore somewhat limited. This is, of course, an instance of a general problem with the repeated game concept. As is well known, any individually rational outcomes of the one-shot game can be supported by a perfect equilibrium of the supergame (see, e g, Fudenberg and Maskin [22]).

Secondly, as stressed by proponents of the "bounded rationality" paradigm (see, e g, Simon [60]), there are reasons to believe that the fact that real-world decision-making is to a greater or lesser extent myopic rather than infinitely "rational" has important implications for modeling. This rings particularly true if the game situation is thought to be recurrent over a very long interval of time, i e, for a succession of generations.

For these reasons, an adaptive, evolutionary approach will be used here instead. This involves the assumption that, instead of being instantaneous optimizers, individuals adhere to behavioral rules (in this case, as to which good to use for indirect exchange) which change only slowly over time. Specifically, an individual will be assumed to change his strategy choice only when he observes adherents to another strategy doing better than himself. This is also clearly in the spirit of Menger's money theory.

4.6 The Evolutionary Model

Now define $V_{Mt} = n_{Mt}f(n_{Mt})$ and $V_{Bt} = n_{Mt}r + n_{Bt}(r + g(n_{Mt}))$, and let $\Pi_{Mt} = V_{Mt}/(n_{Mt}V_{Mt} + n_{Bt}V_{Bt})$ and $\Pi_{Bt} = V_{Bt}/(n_{Mt}V_{Mt} + n_{Bt}V_{Bt})$ be measures of the performance (relative to the population average) at t of the M and B strategies, respectively. (Observe that the *ex ante* expectation of the single agent and the actual *ex post* subpopulation averages are equal.)

We now wish to specify an adaptive process that relates the inflow and outflow of adherents to the two strategies to Π_{Mt} and Π_{Bt}, i e, transition functions of the general form

$$\begin{aligned} n_{Mt+1} &= n_{Mt} + \psi^i_M(\Pi_{Mt})n_{Bt} - \psi^o_M(\Pi_{Bt})n_{Mt} \\ n_{Bt+1} &= n_{Bt} + \psi^i_B(\Pi_{Bt})n_{Mt} - \psi^o_B(\Pi_{Mt})n_{Bt} \end{aligned} \tag{4.5}$$

where ψ^i and ψ^o represent newcomers and and quitters, respectively (in proportional terms).

The aggregate specification below has the reasonable properties that the inflow to a strategy subpopulation responds positively to the difference in relative performance, and that the adjustment rate is faster the larger the subpopulation is already. The latter property could be thought of as reflecting the increased probability of the individual's detecting a positive payoff differential when successful users are relatively numerous. Thus let

$$\begin{aligned} \psi^i_M(\Pi_{Mt}) &= \psi^o_B(\Pi_{Mt}) &= \begin{cases} n_{Mt}(\Pi_{Mt} - \Pi_{Bt}), & \Pi_{Mt} \geq \Pi_{Bt}; \\ 0, & \text{otherwise}; \end{cases} \\ \psi^o_M(\Pi_{Bt}) &= \psi^i_B(\Pi_{Bt}) &= \begin{cases} n_{Bt}(\Pi_{Bt} - \Pi_{Mt}), & \Pi_{Bt} \geq \Pi_{Mt}; \\ 0, & \text{otherwise}. \end{cases} \end{aligned} \tag{4.6}$$

Observe that this specification does not allow for stationary states where a positive inflow and outflow exactly balance each other. While this could be considered as rendering the dynamic version somewhat less interesting than might have been the case otherwise, it has the positive feature of retaining a very clear connection to the underlying static game.

Substitution of (4.6) into (4.5) yields the specific transition functions for the population proportions

$$\begin{aligned} n_{Mt+1} &= n_{Mt}\Pi_{Mt}, \\ n_{Bt+1} &= n_{Bt}\Pi_{Bt}. \end{aligned} \tag{4.7}$$

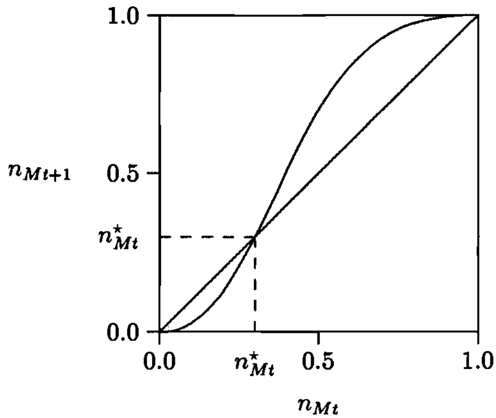

Figure 4.2: The phase diagram for n_{Mt}.

A cursory inspection reveals that the (n_M, n_B) pairs $(0,1)$ and $(1,0)$, which were equilibria in the static game, are stationary states of this system. The system also reproduces the static bimorphism as a dynamic fixpoint $n^\star_{Mt}$, where

$$V_{Mt}(n^\star_{Mt}) = V_{Bt}(n^\star_{Mt}). \tag{4.8}$$

The dynamic stability of the bimorphism $n^\star_{Mt}$ may now be investigated by differentiating (4.7) and utilizing condition (4.8):

$$\left.\frac{\partial n_{Mt+1}}{\partial n_{Mt}}\right|_{n_{Mt}=n^\star_{Mt}} = 1 + \frac{n^\star_{Mt} n^\star_{Bt}(V'_{Mt}(n^\star_{Mt}) - V'_{Bt}(n^\star_{Mt}))}{V_{Mt}(n^\star_{Mt})} > 1, \tag{4.9}$$

(since $V'_{Mt} = f(n_{Mt}) + n_{Mt}f'(n_{Mt}) > 0, \forall n_{Mt}$, and $V'_{Bt} = -g(n_{Mt}) + n_{Bt}g'(n_{Mt}) < 0, \forall n_{Mt}$), i e, $n^\star_{Mt}$ is a repellor. Figure 4.2 shows the general shape of the phase diagram for n_{Mt}.

Being a repellor, dynamically $n^\star_{Mt}$ will never actually occur. (I e, the set of initial states such that $n^\star_{Mt}$ occurs has measure zero.) The interpretation should be of $n^\star_{Mt}$ as a *critical mass* of M-users necessary for convergence to the pure M-equilibrium. Conversely, starting from a situation where M is the CAMOE, $n^\star_{Bt} = 1 - n^\star_{Mt}$ is the relative size of the subpopulation of B-users that must be assembled for a convention switch to occur.

Now assume that the B convention is indeed strongly efficient, i e, that $r + g(0) > f(1)$. Does this imply that the critical mass $n^\star_{Bt}$ is small in the sense of being less than .5?

Using (4.2) to form the difference

$$n_{Bt}^{\star} - .5 = \frac{.5(f(n_{Mt}^{\star}) - g(n_{Mt}^{\star})) - r}{f(n_{Mt}^{\star}) + g(n_{Mt}^{\star})}, \tag{4.10}$$

we find its sign indeterminate in the absence of assumptions about the relative sizes of f, g and r at the equilibrium. In particular, the values of $f(1)$ and $g(0)$ do not enter explicitly into this expression, which means that the critical mass is not necessarily dependent on the relative efficiency of one convention over the other.

The independence result is of course in a way an artefact of the particular adaptive process specified. In adaptive models such as those studied in Friedman and Rosenthal [20], individuals are assumed to evaluate their outcomes not only in relation to current population payoffs, but relative to the absolute feasible maximum (which would be equal to $g(0)+r$ in the present model). Such a specification obviously makes the critical mass directly dependent on efficiency. However, such a construction places more stringent informational demands on agents. Here I only assume that individuals are able to observe their own payoffs and the population average, as well as identify the strategies in use in the population. They need have no global knowledge of the game.

We may, however, state some necessary and sufficient conditions for the expression (4.10) to be positive, i e, for $n_{Bt}^{\star}$ to be large.

Proposition 4.4 *The critical mass for a switch from the M-convention is large, i e, $n_{Bt}^{\star} > .5$, only if $r \leq .5(f(1) - g(1))$.*

For (4.10) to be positive, we must have that $r < .5(f(n_{Mt}^{\star}) - g(n_{Mt}^{\star})) \leq .5(f(1) - g(1))$, because of monotonicity.

Proposition 4.5 *We have that $n_{Bt}^{\star} > .5$ if $f(n_{Mt}) > g(n_{Mt}), \forall n_{Mt}$, and $r < .5(f(n_{Mt}) - g(n_{Mt})), \forall n_{Mt}$.*

This is obviously true.

While the nature of the game as specified does thus not necessarily imply that $n_{Bt}^{\star}$ will be small, it comes as no surprise that it is relatively smaller the larger is r. This may be checked by considering (4.8) as implying the existence of a function $n_{Mt}^{\star}(r)$ and differentiating:

$$\frac{\partial n_{Bt}^{\star}}{\partial r} = -\frac{1}{V_{Mt}'(n_{Mt}^{\star}) - V_{Bt}'(n_{Mt}^{\star})} < 0. \tag{4.11}$$

It could be noted that since the model always has three stationary states there is no possibility of "catastrophic" behavior. That is, continuous changes in r will not lead to sudden convention switches.

Finally, the special case where the f and g functions are constants may be studied. Letting $f(n_{Mt}) = \bar{f}, \forall n_{Mt}$, and $g(n_{Mt}) = \bar{g}, \forall n_{Mt}$, we may solve explicitly for $n^{\star}_{Bt}$:

$$n^{\star}_{Bt} = \frac{\bar{f} - r}{\bar{f} + \bar{g}}, \tag{4.12}$$

where the smallness of $n^{\star}_{Bt}$ is now directly dependent on the relative efficiency of one convention over the other. Interestingly enough, if we have that $\bar{g} + r < \bar{f} < \bar{g} + 2r$, $n^{\star}_{Bt}$ will be small even though the convention where all agents use M and receive $\bar{f}$ is efficient. If we have that $\bar{f} = \bar{g}$, (4.12) reduces to

$$n^{\star}_{Bt} = .5 - \frac{r}{2\bar{f}} < .5. \tag{4.13}$$

To summarize and relate these results to the Wallace proposition, we cannot conclude that, starting from a situation where M (the "non-interest-bearing" good) is the unique CAMOE, a switch to using B (the "interest-bearing" good) would be either spontaneous or easy, unless the interest payment is very much larger than the payoff to be had from engaging in exchange transactions, or B is generally much superior to M as a means for carrying out transactions.

4.7 Concluding Remarks

The "legal restrictions" theory of money seems to have its origin in the following problem: How to explain why anyone would want to hold ordinary currency in a typical, general-equilibrium-based macro-model world, unless they are forced to. In such a model, it is normally simply postulated that transactions must be carried out in a specific good which has no other uses. The Wallace argument serves the important purpose of focusing attention on the function of a medium of exchange.

In the real world, money is an important instrument for achieving what the auctioneer takes care of in general equilibrium models, i e, the coordination of the different economic plans of a myriad of different individuals. While this may of course be assumed away in any given model, care must be taken when actually discussing what would happen in monetary anarchy, where no central coordinator is at hand. The "legal restrictions" confusion consists in continuing to treat "money" as an asset among others after restrictions on the means of transaction have been removed. From that point of view, it is hardly surprising that holding ordinary currency appears highly undesirable.

Another typical macro-model approach is also part of the problem: The assumption that the economy may be treated as if it effectively consisted of a single agent. Obviously, models which study only the portfolio choices of a "representative" individual will not be able to take into consideration any strategic aspects, and are thus not able to deal with the peculiarities of an actual medium of exchange.

Chapter 5

The Firm

5.1 Introduction

According to the author himself, Ronald Coase's 1937 article "The Nature of the Firm" [9] (henceforth, NF) has been "much cited and little used" (Coase [8]). This finding is corroborated by recent surveys of the field such as Hart [25] and Milgrom and Roberts [45]. In the paper, the existence of firms is suggested to be explainable by the fact that there are "costs of using the market mechanism," or transaction costs, which can be avoided through explicit organization. But planning of course has its own costs, so the size of the firm is determined by the relationship of market transaction costs and internal transaction costs.

The lack of subsequent applications of Coase's view of the firm probably has its origin both in the non-neoclassical framework of the paper and the vagueness of the concept of transaction costs. Although the verbal argument is readily understood, putting its insights into a model is not so easily done.

In Coase's later "The Problem of Social Cost" [10] (henceforth, PSC), the transaction cost notion surfaces again. This time, in a perhaps seemingly unrelated context, transaction costs may stop people from bargaining about certain effects of the actions of others. This potential problem of "externality" prompts Dahlman in "The Problem of Externality" [12] (henceforth, PE) to try to get a firmer grip on the slippery concept of transaction costs. He finds that the only "costs" worthy of being considered different in kind from the ordinary production costs that neoclassical competitive equilibrium handles well are resource losses due to imperfect information, or uncertainty.

One important source of such uncertainty may be lack of information about trading opportunities in markets where agents have to search for trading partners. Here lies one clue to the strange status of NF in the theory of the firm. In a neoclassical general equilibrium world, there is no uncertainty about trading opportunities. Nobody has to search for someone else to trade with. The imaginary auctioneer conveniently takes care of the coordination of individual plans. There are no transaction costs.

The purpose of the present discussion is to suggest how the notion of transaction costs arises naturally in a different type of model, where the individual is faced with uncertainty about trading opportunities. The ultimate question is whether the Coasian firm may be better understood in such a context.

The chapter is organized as follows. Section 5.2 develops a measure of transaction cost in the sense of expected value of perfect information for models where agents are randomly paired for interaction. The assumption of random pairing is a simple approach to modeling situations where agents have to search for trading partners, and there is uncertainty about who will be encountered. I find that at least for binary choice situations, certain market equilibria have the property of minimizing transaction cost.

A particular type of game will have positive transaction costs in equilibrium. This leaves a profit opportunity for an entrepreneur who knows how to organize group production. Equilibrium with a firm in this sense is studied in Section 5.3. The equilibrium size of the firm will depend negatively on the cost of organization, and positively on the complementarity of the goods produced.

Finally, in Section 5.4, I note that this account of the emergence of the firm has a serious flaw. Since the marginal cost of a market transaction is zero, i e, the expected value of market participation is unaffected by the size of firms, the theory cannot explain why all production is not organized in firms.

5.2 The Cost of Market Transactions

5.2.1 The Coase-Dahlman Transaction Cost Concept

Dahlman's discussion of the role of externality in economic theory has two major parts.

To begin with, he notes that at least since the publication of PSC, externalities are thought to exist in a welfare-relevant sense only when the presence of transaction costs makes full individual contracting about some good impossible.

But what are transaction costs? They must be a category fundamentally different from ordinary production and transportation costs. Otherwise, how could they, among other problems, be overcome by government intervention? Dahlman finds that the literature on general equilibrium with "transaction costs" as a share of resources that gets lost in exchanges, used for instance to explain the use of money, fails to satisfy the criterion of differentness.

> These, then, represent the first approximation to a workable concept of transaction costs: search and information costs, bargaining and decision costs, policing and enforcement costs.
>
> Yet, this functional taxonomy of different transaction costs is unnecessarily elaborate: fundamentally, the three classes reduce to a single one—for they all have in common that they represent resource losses due to lack of information. (PE, p 148)

The second part of Dahlman's argument is to note that since only resource losses due to imperfect information can be a category apart from ordinary production costs, the argument for intervention must be based on the value judgment that the government is omniscient.

For the present purposes, transaction costs in the Coase-Dahlman sense can be seen to equal the expected value of perfect information.

5.2.2 Coordination Problems and Transaction Costs

I will study models where each individual faces uncertainty regarding who he will in fact be interacting with. Consider a large population of agents who are paired randomly.

Each pair then plays a symmetric bimatrix game. For simplicity, I will consider 2×2 examples. Let the game in general be

$$G = \begin{array}{cc} & \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{array}{c} s_1 \\ s_2 \end{array} & \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \end{array}, \tag{5.1}$$

where, because of symmetry, the payoffs are those of either player when he is taken to be the row player. We can assume with no loss of generality that $p_{ij} \geq 0, \forall i, j$.

Each player has to decide on his action *before* the pairing takes place. Further assume only pure strategies can be used. Let n_i be the population proportion that uses strategy s_i. It is also the probability of encountering the strategy under random pairing.

If there was perfect information, every individual would know beforehand the action choice of the opponent he was to be paired with. Assuming he used this information optimally, he could guarantee himself an expected payoff equal to $n_1\Pi_1 + n_2\Pi_2$, where $\Pi_j = \max_i p_{ij}$. Under uncertainty, however, the expected payoff for pure strategy i is equal to $n_1 p_{i1} + n_2 p_{i2}$. The expected value of perfect information to a strategy i user, or the *transaction cost* of strategy i, is therefore

$$\tau_i := \sum_{j=1}^{2} n_j \Pi_j - \sum_{j=1}^{2} n_j p_{ij}, \tag{5.2}$$

and the average transaction cost for the population as a whole is

$$\tau_p := \sum_{i=1}^{2} n_i \tau_i = \sum_{j=1}^{2} n_j \Pi_j - \sum_{i=1}^{2}\sum_{j=1}^{2} n_i n_j p_{ij}. \tag{5.3}$$

When speaking of transaction cost in the following, I will be referring to τ_p if nothing else is indicated.

These games have two types of equilibria in pure strategies, *monomorphisms*, where $n_i = 1$ for some i, and *polymorphisms*, which have more than one strategy positively represented in the population.[1] Some of these equilibria have the interesting property of minimizing transaction costs. In fact,

Proposition 5.1 *Let G be a symmetric 2×2 game. Then every pair (n_1, n_2) such that $n_i \geq 0$ and $n_1 + n_2 = 1$ that minimizes τ_p is also an equilibrium of G.*

[1] The terminology is borrowed from biology and evolutionary game theory.

A proof of this is given in the Appendix to this chapter.

Consider the example game

$$\begin{array}{c} \\ s_1 \\ s_2 \end{array} \begin{array}{c} \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \end{array}, \tag{5.4}$$

where $\alpha, \beta > 0$. The transaction cost function for this type of coordination game is graphed in Figure 1.1 on page 10. The set of monomorphic equilibrium values for n_1 is $\{0, 1\}$, which corresponds to the set of transaction cost minima.

Another distinct type of coordination problem can be represented by the matrix

$$\begin{array}{c} \\ s_1 \\ s_2 \end{array} \begin{array}{c} \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix} \end{array}, \tag{5.5}$$

where $\alpha, \beta > 0$. This game has no monomorphic equilibria, because individuals would want to coordinate on using different strategies. This type of game could therefore be called a ***division of labor*** game.

The division of labor game has a unique polymorphism, which is also a transaction cost minimum as seen in Figure 1.2 on page 11. The equilibrium transaction cost can never be zero, however, which provides the rationale for using this type of game to explain the emergence of the firm in the following Section.

5.2.3 A Division of Labor Problem

A modest fleshing out of the abstract division of labor coordination problem gives rise to possibly the simplest "general equilibrium" model that has something to say about the *raison d'être* of the firm. Assume there is a continuum of individuals and two goods, good A and good B. Each individual can produce either two units of good A or two units of good B during the relevant time period. All individuals have the same preferences. The most preferred situation is to consume one unit of A and one unit of B, the next most preferred situation is to consume two units of A, with the consumption of two units of B the lowest ranked alternative. This is just an ordinary assumption of convex preferences. The rankings will be taken to be cardinal measures with the expected utility property. Note that the larger u_{AB} is relative to u_{AA} and u_{BB}, the greater is the complementarity in consumption of the two goods.

Now assume the individuals make their production decisions and are then paired randomly. Each pair may trade in their produced goods if desired. Obviously trade will take place only if the paired individuals have produced different goods. That is, each individual will *ex ante* face a situation of the following structure, where the rows represent his own production decision and the columns that of the person he is paired with, and $u_{AB} > u_{AA}, u_{AB} > u_{BB}$.

$$\begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} u_{AA} & u_{AB} \\ u_{AB} & u_{BB} \end{pmatrix} \end{array}. \tag{5.6}$$

Let n_A be the population proportion of individuals who choose to produce good A, and $n_B = 1 - n_A$ the proportion of individuals who choose to produce good B. The expected utility of producing A is then $n_A u_{AA} + n_B u_{AB}$, and that of producing B is $n_A u_{AB} + n_B u_{BB}$.

The transaction cost concept of the preceding discussion may now be applied to this situation. Under certainty, an individual would know who he would later be paired with when making his production decision. Optimal adjustment to this knowledge would mean that regardless of the proportions of strategy representation in the population, the individual would have an expected utility of u_{AB}. The reduction in utility that is a result of these ideal circumstances not being the case is then the difference between the utility under certainty and the expected utility of each strategy under uncertainty. For strategy A, this transaction cost is then $\tau_A := u_{AB} - (n_A u_{AA} + n_B u_{AB}) = n_A(u_{AB} - u_{AA})$, and for strategy B, $\tau_B := u_{AB} - (n_A u_{AB} + n_B u_{BB}) = n_B(u_{AB} - u_{BB})$. The average transaction cost for the market as a whole is $\tau_m = n_A \tau_A + n_B \tau_B$.

The game has a unique non-cooperative equilibrium where the expected values of the production strategies are equal, i e, when

$$n_A = n_A^\star := \frac{u_{AB} - u_{BB}}{2u_{AB} - u_{AA} - u_{BB}}. \tag{5.7}$$

We have that $n_A^\star \in (0, 1)$ by virtue of the assumptions. As expected, (5.7) is also the first-order condition for an average market transaction cost minimum, i e,

$$n_A^\star = \arg\min_{n_A} n_A \tau_A + n_B \tau_B. \tag{5.8}$$

Since the function is strictly convex, this does indeed define a minimum.

Let $V_m^\star$ be the equilibrium expected payoff for market participation. We have that

$$V_m^\star = \frac{u_{AB}^2 - u_{AA}u_{BB}}{2u_{AB} - u_{AA} - u_{BB}}. \tag{5.9}$$

Market equilibrium transaction cost is

$$\tau_m^\star = u_{AB} - \frac{u_{AB}^2 - u_{AA}u_{BB}}{2u_{AB} - u_{AA} - u_{BB}}, \tag{5.10}$$

which is positive. This means organization could potentially improve on the market situation.

5.3 The Firm

5.3.1 The Internal Workings of Organizations

In the two-good economy with random pairing of trading partners studied above, any alternative production arrangement that is less uncertain than market equilibrium uncertainty represents a profit opportunity. If a group of individuals could get together and agree to follow a single plan, there is an obvious optimal way of organizing production. Half the group produces good A and the other half good B.

If planning was costless, this scheme could guarantee each group member an expected outcome of u_{AB}. However, as noted most convincingly by Polanyi [54] (almost never credited in this context), central planning of a certain task has a natural disadvantage against the decentralized performance of the same task. The argument goes as follows. Assume a certain collective task can be performed during a certain time period by a number of individuals acting independently of one another. The action of each individual involves making one or more decisions, or adjustments, during the time period. By definition, central planning means substituting a single decision-maker, or a hierarchy of decision-makers, for the individual decision-makers in the independent-action situation. In any case, a smaller number of people now have to make the same total number of adjustments that make up the performance of the task. The number of adjustments per decision-maker increases, which would tend to lessen the efficiency of performance.

In terms of the present model, I will characterize the firm by assuming some utility is lost during the time period due to time spent centrally coordinating the production of the individuals in the firm. This loss is increasing in the number of participants in the firm. Letting x_f denote the population proportion that participates in the firm, I will for simplicity assume the utility loss to be cx_f, where c is a constant greater than u_{AB}. This means that, at some point before the firm has grown to encompass all individuals in the society, the entire period would have to be spent coordinating the task.

5.3.2 Equilibrium with a Firm

Each individual is now faced with a choice between independent production and marketing of good A, independent production and marketing of good B, and participation in the firm, where the production decision will be centrally coordinated. In equilibrium, the expected values of the three strategies are equal, i e, the following is true of the equilibrium population proportions $x_A^\star$, $x_B^\star$ and $x_f^\star$:

$$\frac{x_A^\star}{x_A^\star + x_B^\star}u_{AA} + \frac{x_B^\star}{x_A^\star + x_B^\star}u_{AB} = \frac{x_A^\star}{x_A^\star + x_B^\star}u_{AB} + \frac{x_B^\star}{x_A^\star + x_B^\star}u_{BB} = u_{AB} - cx_f^\star \qquad (5.11)$$

and

$$x_A^\star + x_B^\star + x_f^\star = 1. \qquad (5.12)$$

The system may be solved by first noting that it is recursive. For market equilibrium we must have that

$$\frac{x_A^\star}{x_A^\star + x_B^\star} = n_A^\star \qquad (5.13)$$

and

$$\frac{x_B^\star}{x_A^\star + x_B^\star} = n_B^\star, \qquad (5.14)$$

where the $n_i^\star$ are those of the preceding Section.

The equilibrium firm size is then given by

$$u_{AB} - cx_f^\star = V_m^\star. \qquad (5.15)$$

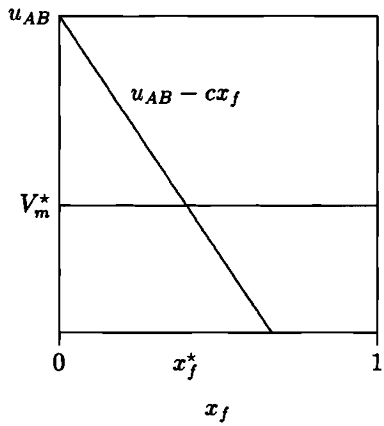

Figure 5.1: Equilibrium with a firm.

The situation is illustrated in Figure 5.1. The explicit solution is

$$x_f^\star = \frac{(u_{AB} - u_{AA})(u_{AB} - u_{BB})}{c(2u_{AB} - u_{AA} - u_{BB})}. \tag{5.16}$$

A greater cost of organization corresponds to a lower equilibrium firm size, as one would expect, since

$$\frac{\partial x_f^\star}{\partial c} = -\frac{(u_{AB} - u_{AA})(u_{AB} - u_{BB})}{c^2(2u_{AB} - u_{AA} - u_{BB})} \tag{5.17}$$

is a negative number.

An increase in the complementarity of good A and good B affects both the expected value of the market equilibrium and the "profit function" of the firm, so the net effect on equilibrium firm size is not as intuitive. It is always positive, however, since

$$\frac{\partial x_f^\star}{\partial u_{AB}} = \frac{(u_{AB} - u_{AA})^2 + (u_{AB} - u_{BB})^2}{c(2u_{AB} - u_{AA} - u_{BB})^2} \tag{5.18}$$

is positive.

5.4 Remarks

The previous discussion depends crucially on the implicit assumption of the impossibility of more than one firm. Such a situation could be the case if the organizational coordination scheme is the monopoly property of a single entrepreneur through, e g, a patent.

It is immediately obvious that if more than one firm can be in existence at the same time, and if each faces the profit function $u_{AB} - cx_f$, then the optimal firm size is infinitely small, there would be infinitely many firms, and no one would want to participate in the market.

Observe that adding set-up costs, in the form of a minimum size of the firm, does not affect the result that all individuals would be organized in firms. As long as the single-firm profit function lies above $V_m^\star$ at some point, there is an optimal way of arranging the entire membership of society into firms.

The reason for this lies in the fact that $V_m^\star$ (and market transaction cost) is unaffected by the size of the proportion of individuals who are in firms. That is, the *marginal* cost of a market transaction is equal to zero.

One way of introducing an upward-sloping $V_m^\star$ curve, which could cure the problem, would be to let the number of market participants influence the probability of coordination in the market. Pairing would no longer be random, and the probability of being paired with a matching trading partner would be greater the smaller the market size. In the simple formulation given above, the "transaction cost" theory of the firm paradoxically enough has problems with explaining the existence of *markets*.

Appendix: Proof of Proposition 5.1

To prove this, construct a new matrix R from G by subtracting the best reply payoff of each column from the elements in that column. That is,

$$R = \begin{pmatrix} p_{11} - \Pi_1 & p_{12} - \Pi_2 \\ p_{21} - \Pi_1 & p_{22} - \Pi_2 \end{pmatrix}. \tag{5.19}$$

Let r_{ij} be the elements of R. We have that $r_{ij} \leq 0, \forall i, j$. If R is considered as a game, then if (n_1, n_2) is an equilibrium of R,

$$\sum_i \sum_j n_i n_j r_{ij} \geq \sum_i \sum_j q_i n_j r_{ij}, \forall q_1, q_2 \geq 0, q_1 + q_2 = 1 \tag{5.20}$$

(since the population game is analogous to a symmetric two-player game with mixed strategies allowed and only symmetric equilibria considered). This can be rewritten

$$\sum_i \sum_j n_i n_j p_{ij} - \sum_j n_j \Pi_j \geq \sum_i \sum_j q_i n_j p_{ij} - \sum_j n_j \Pi_j, \forall q_1, q_2 \geq 0, q_1 + q_2 = 1, \tag{5.21}$$

which implies that (n_1, n_2) is also an equilibrium of G. So G and R have the same equilibria. Furthermore, a maximum of $\sum_i \sum_j n_i n_j r_{ij}$ is a transaction cost minimum, since one is the negative of the other. So it suffices to show that if (n_1, n_2) is a maximum of $\sum_i \sum_j n_i n_j r_{ij}$, then (n_1, n_2) is an equilibrium of R.

Consider the case when G has monomorphic equilibria. Then one or both of the diagonal elements of R are equal to zero. If $r_{ii} = 0$, then setting $n_i = 1$ will give $\sum_i \sum_j n_i n_j r_{ij} = 0$ and be an equilibrium. Unless all elements of R are equal to zero, in which case the proposition holds trivially since any pairs (n_1, n_2) are also equilibria, any pair which gives positive weight to both strategies will have $\sum_i \sum_j n_i n_j r_{ij} < 0$, so it cannot be a maximum.

Now consider the case when G has no monomorphic equilibria. Then we must have that $r_{21} = r_{12} = 0$. Since every symmetric bimatrix game has a symmetric equilibrium, which is unique in this case, there is a polymorphic equilibrium (n_1^*, n_2^*). It is found by setting the expected values of the two strategies equal, i e,

$$n_1^* = \frac{r_{22}}{r_{11} + r_{22}} \tag{5.22}$$

and

$$n_2^* = \frac{r_{11}}{r_{11} + r_{22}} \tag{5.23}$$

Assume (n_1, n_2) is a maximum of $\sum_i \sum_j n_i n_j r_{ij}$. Then

$$\sum_i \sum_j n_i n_j r_{ij} \geq \sum_i \sum_j q_i q_j r_{ij}, \forall q_1, q_2 \geq 0, q_1 + q_2 = 1. \tag{5.24}$$

In particular, this must hold for $q_1 = n_1^*$ and $q_2 = n_2^*$. Substituting and rearranging yields the condition

$$\frac{(n_1 r_{11} - n_2 r_{22})^2}{r_{11} + r_{22}} \geq 0. \tag{5.25}$$

Since the denominator is strictly negative, this can hold only with equality, which is the case when $n_1 = n_1^*, n_2 = n_2^*$. This establishes the truth of Proposition 5.1.

Bibliography

[1] Armen A Alchian. Uncertainty, evolution and economic theory. *Journal of Political Economy*, 58:211–221, 1950.

[2] Norman Barry. The tradition of spontaneous order. *Literature of Liberty*, V(2):7–58, 1982.

[3] Jon Barwise and John Perry. *Situations and Attitudes*. The MIT Press, Cambridge, 1983.

[4] B D Bernheim, B Peleg, and M D Whinston. Coalition-proof equilibria. I. Concepts. *Journal of Economic Theory*, 42:1–12, 1987.

[5] Geoffrey Brennan and James M Buchanan. *The Reason of Rules*. Cambridge University Press, New York, 1985.

[6] Winston C Bush. Individual welfare in anarchy. In Gordon Tullock, editor, *Explorations in the Theory of Anarchy*, pages 5–18, Center for the Study of Public Choice, Blacksburg, Virginia, 1972.

[7] L L Cavalli-Sforza and M W Feldman. *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press, Princeton, 1981.

[8] Ronald H Coase. Industrial organization: a proposal for research. In Victor R Fuchs, editor, *Policy Issues and Research Opportunities in Industrial Organization*, pages 59–73, National Bureau of Economic Research, Cambridge, 1972.

[9] Ronald H Coase. The nature of the firm. *Economica*, 4:386–405, 1937.

[10] Ronald H Coase. The problem of social cost. *Journal of Law and Economics*, 3:1–44, 1960.

[11] Tyler Cowen and Randall Kroszner. The development of the new monetary economics. *Journal of Political Economy*, 95:567–590, 1987.

[12] Carl J Dahlman. The problem of externality. *Journal of Law and Economics*, 22:141–162, 1979.

[13] Partha Dasgupta. Applying game theory. Some theoretical considerations. *European Economic Review*, 33:619–624, 1989.

[14] Richard Dawkins. *The Selfish Gene.* Oxford University Press, Oxford, 1976.

[15] Jon Elster. *Explaining Technical Change: A Case Study in the Philosophy of Science.* Cambridge University Press, Cambridge, 1983.

[16] Jon Elster. *Ulysses and the Sirens: Studies in Rationality and Irrationality.* Cambridge University Press, Cambridge, 1979.

[17] Joseph Farrell. Communication, coordination and Nash equilibrium. *Economics Letters*, 27:209–214, 1988.

[18] Joseph Farrell and Garth Saloner. Installed base and compatibility: innovation, product preannouncements, and predation. *American Economic Review*, 76(5):940–955, 1986.

[19] Joseph Farrell and Garth Saloner. Standardization, compatibility, and innovation. *Rand Journal of Economics*, 16(1):70–83, 1985.

[20] James W Friedman and Robert W Rosenthal. A positive approach to non-cooperative games. *Journal of Economic Behavior and Organization*, 7:235–251, 1986.

[21] Milton Friedman. The methodology of positive economics. In *Essays in Positive Economics*, pages 3–43, University of Chicago Press, Chicago, 1953.

[22] Drew Fudenberg and Eric Maskin. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54(3):533–554, May 1986.

[23] Mark Granovetter and Roland Soong. Threshold models of interpersonal effects in consumer demand. *Journal of Economic Behavior and Organization*, 7:83–99, 1986.

[24] John C Harsanyi and Reinhard Selten. *A General Theory of Equilibrium Selection in Games*. MIT Press, Cambridge, 1988.

[25] Oliver Hart. An economist's perspective on the theory of the firm. *Columbia Law Review*, 89(7):1757–1774, 1989.

[26] Friedrich A Hayek. *The Constitution of Liberty*. The University of Chicago Press, Chicago, 1960.

[27] Friedrich A Hayek. *The Counterrevolution of Science*. Liberty Press, Indianapolis, 1979.

[28] Friedrich A Hayek. *The Fatal Conceit. The Errors of Socialism*. Routledge, London, 1988.

[29] Friedrich A Hayek. *Law, Legislation and Liberty*. The University of Chicago Press, Chicago, 1973–1979. 3 volumes.

[30] Ronald A Heiner. The origin of predictable behavior. *American Economic Review*, 73:560–595, 1983.

[31] Jack Hirshleifer. Evolutionary models in economics and law: cooperation versus conflict strategies. *Research in Law and Economics*, 4:1–60, 1982.

[32] Robert A Jones. The origin and development of media of exchange. *Journal of Political Economy*, 84:757–775, 1976.

[33] Michael L Katz and Carl Shapiro. Technology adoption in the presence of network externalities. *Journal of Political Economy*, 94(4):822–841, 1986.

[34] Nobuhiro Kiyotaki and Randall Wright. Money and specialization. 1988. Working paper.

[35] Nobuhiro Kiyotaki and Randall Wright. On money as a medium of exchange. *Journal of Political Economy*, 97(4):927–954, 1989.

[36] Richard N Langlois. Coherence and flexibility: social institutions in a world of radical uncertainty. In Israel M Kirzner, editor, *Subjectivism, Intelligibility and Economic Understanding*, New York University Press, New York, 1986.

[37] Richard N Langlois. Rationality, institutions, and explanation. In Richard N Langlois, editor, *Economics as a Process: Essays in the New Institutional Economics*, Cambridge University Press, Cambridge, 1986.

[38] David K Lewis. *Convention. A Philosophical Study.* Basil Blackwell, Oxford, 1986.

[39] John Locke. *Two Treatises of Government.* Cambridge University Press, New York, 1967.

[40] Alfred Marshall. *Principles of Economics.* Macmillan, London, 1920. Eighth edition.

[41] John Maynard Smith. *Evolution and the Theory of Games.* Cambridge University Press, Cambridge, 1982.

[42] John Maynard Smith and G R Price. The logic of animal conflict. *Nature*, 246:15–18, 1973.

[43] Carl Menger. *Investigations into the Method of the Social Sciences with Special Reference to Economics.* New York University Press, New York, 1985. Trans by Francis J Nock.

[44] Carl Menger. *Principles of Economics.* New York University Press, New York, 1976. Trans James Dingwall and Bert F Hoselitz.

[45] Paul Milgrom and John Roberts. Economic theories of the firm: past, present, and future. *Canadian Journal of Economics*, XXI(3):444–458, August 1988.

[46] John F Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences of the USA*, 36:48–49, 1950.

[47] Richard R Nelson and Sidney G Winter. *An Evolutionary Theory of Economic Change*. Harvard University Press, Cambridge, 1982.

[48] Jürg Niehans. *The Theory of Money*. The Johns Hopkins University Press, Baltimore, 1978.

[49] Robert Nozick. *Anarchy State and Utopia*. Basic Books, New York, 1974.

[50] Gerald P O'Driscoll, Jr. Money: Menger's evolutionary theory. *History of Political Economy*, 18:601–616, 1986.

[51] Seonghwan Oh. A theory of a generally acceptable medium of exchange and barter. *Journal of Monetary Economics*, 23:101–119, 1989.

[52] Michael Polanyi. *The Logic of Liberty. Reflections and Rejoinders*. The University of Chicago Press, Chicago, 1951.

[53] Michael Polanyi. *Personal Knowledge. Towards a Post-Critical Philosophy*. Routledge & Kegan Paul, London, 1958.

[54] Michael Polanyi. The span of central direction. In *The Logic of Liberty*, pages 111–137, University of Chicago Press, Chicago, 1951.

[55] Karl R Popper. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Harper Colophon, New York, 1965.

[56] Thomas C Schelling. *Micromotives and Macrobehavior*. Norton, New York, 1978.

[57] Andrew Schotter. *The Economic Theory of Social Institutions*. Cambridge University Press, Cambridge, 1981.

[58] Joseph A Schumpeter. *The Theory of Economic Development*. Harvard University Press, Cambridge, 1934.

[59] Reinhard Selten. Re-examination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4:25–55, 1975.

[60] Herbert A Simon. *Reason in Human Affairs.* Stanford University Press, Stanford, 1983.

[61] Robert Sugden. *The Economics of Rights, Cooperation and Welfare.* Basil Blackwell, Oxford, 1986.

[62] Peter D Taylor and Leo B Jonker. Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences*, 40:145–156, 1978.

[63] Henri Theil. *Economics and Information Theory.* North-Holland, Amsterdam, 1967.

[64] Edna Ullmann-Margalit. *The Emergence of Norms.* The Clarendon Press, Oxford, 1977.

[65] Edna Ullmann-Margalit. Invisible hand explanations. *Synthèse*, 39:263–291, 1978.

[66] John Umbeck. Might makes rights: a theory of the foundation and initial distribution of property rights. *Economic Inquiry*, XIX(1):38–59, January 1981.

[67] Eric van Damme. Comment on W Güth and M Hellwig: Competition versus monopoly in the supply of public goods. In Rüdiger Pethig and Ulrich Schlieper, editors, *Efficiency, Institutions, and Economic Policy*, pages 219–225, Springer-Verlag, Berlin, 1987.

[68] Eric van Damme. *Stability and Perfection of Nash Equilibria.* Springer-Verlag, Berlin, 1987.

[69] Philippe Van Parijs. *Evolutionary Explanation in the Social Sciences. An Emerging Paradigm.* Tavistock Publications, New York, 1981.

[70] Viktor Vanberg. Spontaneous market order and social rules: A critical examination of F A Hayek's theory of cultural evolution. *Economics and Philosophy*, 2:75–100, 1986.

[71] Neil Wallace. A legal restrictions theory of the demand for 'money' and the role of monetary policy. *Federal Reserve Bank of Minneapolis Quarterly Review*, 7:1–7, 1983.

[72] Lawrence H White. Accounting for non-interest-bearing currency: a critique of the legal restrictions theory of money. *Journal of Money, Credit, and Banking*, 19:448–456, 1987.

[73] Lawrence H White. Competitive payments systems and the unit of account. *American Economic Review*, 74:699–712, 1984.

[74] Lawrence H White. A subjectivist perspective on the definition and identification of money. In Israel M Kirzner, editor, *Subjectivism, Intelligibility and Economic Understanding*, New York University Press, New York, 1986.

[75] Ulrich Witt. The evolution of economic institutions as a propagation process. *Public Choice*, 62(2):155–172, August 1989.

[76] E C Zeeman. Population dynamics from game theory. In Z Nitecki and C Robinson, editors, *Global Theory of Dynamical Systems*, pages 471–497, Springer-Verlag, Berlin, 1980.

EFI reports since 1987

Published in the language indicated by the title

1987

Andersson, T., Ternström, B., Private Foreign Investments and Welfare Effects: A comparative study of five countries in southeast Asia. Research report

Andersson, T., Ternström, B., External Capital and Social Welfare in South-East Asia. Research report

Benndorf, H., Marknadsföringsplanering och samordning mellan företag i industriella system. EFI/MTC

Bergman, L., Mäler, K-G., Ståhl, I., Överlåtelsebara utsläppsrätter. En studie av kolväteutsläpp i Göteborg. Research report

Björkegren, D., Mot en kognitiv organisationsteori. Research report

Claesson, K., Effektiviteten på Stockholms Fondbörs

Davidsson, P. Growth Willingness in Small Firms. Entrepreneurship - and after? Research report

Engshagen, I., Finansiella nyckeltal för koncern versus koncernbolag. En studie om finansiella måltal. Research report

Fredriksson, O., Holmlöv, PG., Julander, C-R., Distribution av varor och tjänster i informationssamhället.

Hagstedt, P., Sponsring - mer än marknadsföring. EFI/MTC

Järnhäll, B., On the Formulation and Estimation of Models of Open Economies. Research report

Kylén, B., Digitalkartans ekonomi. Samhällsekonomiska modeller för strategiska val. Delrapport 2 i "Nordisk Kvantif"

Lundgren, S., Elpriser: Principer och praktik. Research report

Schwarz, B., Hederstierna, A., Socialbidragens utveckling - orsaker eller samvariationer? Research report

Schwartz, E., Begreppet federativ organisation belyst i ett organisationsteoretiskt, juridiskt och empiriskt perpektiv. Research report.

Westlund, A.H. Öhlén, S., Business Cycle Forecasting in Sweden; A problem analysis. Research report

1988

Andréasson, I-M., Costs of Controls on Farmers' Use of Nitrogen. A study applied to Gotland.

Björkegren, D., Från sluten till öppen kunskapsproduktion. Delrapport 3 i forskningsprojektet Lärande och tänkande i organisationer. Research report

Björkegren, D., Från tavistock till Human-relation. Delrapport 4 i forskningsprojektet Lärande och tänkande i organisationer. Research report

Björklund, L., Internationell projektförsäljning. En studie av ad hoc-samverkan mellan företag vid internationella projekt. Research report

Bojö, J., Mäler, K-G., Unemo, L., Economic Analysis of Environmental Consequences of Development Projects. Research report

Brynell, K., Davidsson, P., Det nya småföretagandet? - en empirisk jämförelse mellan små high-tech företag och konventionella småföretag. Research report

Dahlgren, G., Witt, P., Ledning av fusionsförlopp. En analys av bildandet av Ericsson Information Systems AB.

Forsell, A., Från traditionell till modern sparbank. Idé organisation och verksamhet i omvandling. Research report

Hirdman, V., The Stability and the Interest Sensitivity of Swedish Short Term Capital Flows. Research report

Hultén, S., Vad bestämmer de svenska exportmarknadsandelarnas utveckling?

Häckner, J., Biobränslenas konkurrenskraft i ett framtida perspektiv. Research report

Jennergren, P.,Näslund, B., If the Supreme Court had Known Option Theory: A Reconsideration of the Gimo Case. Research report

Jennergren, P., Näslund, B., The Gimo Corporation Revisited: An Exercise in Finance Theory. Research report

Jennergren, P., Näslund, B., Valuation of Debt and Equity Through One- and Two-State Contingent Claims Models - with an Application to a Swedish Court Case. Research report

Jonung, L., Laidler, D., Are Perceptions of Inflation Rational? Some Evidence for Sweden. Research report.

Lagerstam, C., Business Intelligence. Teori samt empirisk studie av elektronik- och bioteknikbranschen i Japan. Research report

Liljegren, G., Interdependens och dynamik i långsiktiga kundrelationer. Industriell försäljning i nätverksperspektiv.

Näslund, B., Några synpunkter på börsfallet i oktober 1987. Research report

Olsson, C., The Cost-Effectiveness of Different Strategies Aimed at Reducing the Amount of Sulphur Deposition in Europe. Research report

Philips, Å., Eldsjälar. En studie av aktörsskap i arbetsorganisatoriskt utvecklingsarbete.

Sellstedt, B., Produktionsstrategier. En diskussion med utgångspunkt från litteraturen. Research report.

Skogsvik, K., Prognos av finansiell kris med redovisningsmått. En jämförelse mellan traditionell och inflationsjusterad redovisning.

Wahlund, R., Skatteomläggningen 1983 - 1985. En studie av några konsekvenser. Research report

Wahlund, R., Varför och hur olika svenska hushåll sparar. Research report

Vredin, A., Macroeconomic Policies and the Balance of Payments.

Åkerman, J., Economic Valuation of Risk Reduction: The Case of In-Door Radon. Research report.

1989

Andersson, T., Foreign Direct Investment in Competing Host Countries. A Study of Taxation and Nationalization.

Björkegren, D., Skönhetens uppfinnare. Research report

Björkegren, D., Hur organisationer lär. Studentlitteratur.

Blomström, M., Transnational Corporations and Manufactoring Exports from Developing Countries. Research report.

Carlsson, A., Estimates of the Costs of Emission control in the Swedish Energy Sector. Research report

Davidsson, P., Continued Entrepreneurship and Small Firm Growth.

DeJuan, A., Fiscal Attitudes and Behavior. A study of 16 - 35 years old Swedish citizens. Research report

Edlund, P-O., Preliminary Estimation of Transfer Function Weights. A Two-Step Regression Approach.

Fridman, B., Östman, L., eds. Accounting Development - some perspectives. In honour of Sven-Erik Johansson.

Gadde, L-E., Håkansson, H., Öberg, M., Stability and Change in Automobile Distribution. Research report

Jonung, L., The Economics of Private Money. Private Bank Notes in Sweden 1831 - 1902. Research report

Jonung, L., Batchelor, R.A., Confidence about Inflation Forecasts: Tests of Variance Rationality. Research report.

Kylén, B., Hur företagschefer beslutar innan de blir överraskade. Ett försök till förklaring av svarsmönster i svagsignalsitutationer.

Lagerstam, C., Jämförelse av skydd med samlad valuta/aktieoption och skydd med separata aktie- och valutaoptioner samt härledning av optionspriser. Research report

Lagerstam, C., On the Pricing and Valuation of Forwards and Options on Futures and their Relationsship. Research report.

Larsson, B., Koncernföretaget. Ägarorganisationer eller organisation för ägare?

Löwstedt, J., Föreställningar, ny teknik och förändring. Tre organisationsprocesser ur ett kognitivt aktörsperspektiv. Doxa Förlag

Löwstedt, J., (red) Organisation och teknikförändring. Studentlitteratur.

Tollgerdt-Andersson, I., Ledarskapsteorier, företagsklimat och bedömningsmetoder.

Schuster, W., Ägandeformens betydelse för ett företag - en studie av ICA-rörelsen.

Schwartz, B., Företaget som medborgare. Samhällskontakter och reklam som legitimeringsinstrument. Research report

Spets, G., Vägen till nej. Anade och oanade konsekvenser av en OS-satsning. Research report.

Stymne, B., Information Technology and Competence Formation in the Swedish Service Sector. IMIT/EFI

1990

Brunsson, N., Forsell, A., Winberg, H., Reform som tradition. Administrativa reformer i Statens Järnvägar.

Drottz-Sjöberg, B-M., Interests in Humanities, Social Science and Natural Science. A Study of High School Students.

Eklöf, J.A., Macro Ecoonomic PLanning with Quantitative Techniques - and Effects of Varying Data Quality.

Steiner, L., Ledningsfunktionen i tillväxtföretag. Ledningsteamens sammansättning och funktion i tillverkande företag inom informationsteknologiindustrin.