The Rise of Money: An Evolutionary Analysis of the Origins of Money.

Kevin Hasker*, and Ahmet Tahmilci.
Economics Department
Bilkent University
06800 Bilkent, Ankara, Turkey

March 1, 2006

Abstract

This paper shows that if there are more goods than money in all trading periods then non-convertible fiat money is evolutionarily successful in complex economies. This result is developed in Kyotaki and Wright [9] search model of money, using a learning algorithm developed by Marimon et. al [13]. When we include an evolutionary model similar to Kandori, Mailath, and Rob [8] we find that fiat money is frequently evolutionarily successful. To be precise let x be the probability someone finds someone who has a good they want—small x represents a complex economy—and let μ be the fraction of people holding money in any trading period. As long as $\mu < \frac{1-2x}{2-2x}$ fiat money will be evolutionarily successful.

1 Introduction

Why do people exchange intrinsically worthless paper money for valuable goods? We have long understood that monetary exchange is Pareto superior to barter, but how did people learn to trust money? Kiyotaki and Wright [9] developed an elegant analytical model that has both the barter exchange equilibrium and a monetary exchange equilibrium but this model can not explain the transition between the two. These are separate equilibria and in the monetary equilibrium people use money only because others use money. In other words if people think money is valuable then they will use money.

To develop a model of how society switches between these two viable equilibria we need to have a model of how people behave out of equilibrium. One that has recently been subjected to analysis in the microeconomics literature is the model of stochastic evolution—first developed by Kandori, Mailath, and Rob [8] and Young [19]. In this model essentially people occasionally try out money, and if a critical mass of people start using fiat money then the rest of the population slowly learns that fiat money is trustworthy, until in the end it is the primary mechanism of trade. We show that

*Coressponding author, e-mail hasker@bilkent.edu.tr

this transition occurs surprisingly frequently. We find that as long as there are more real goods than fiat money produced in each trading period then in complex economies fiat money is evolutionarily dominant (or *stochastically stable* in the terminology of the literature.)

One difficulty that we face is that using money is an equilibrium of a dynamic game—one accepts money today so one can buy things tomorrow. Out of equilibrium calculating the continuation values in this game can be very complicated, so in our model people use the feedback learning model developed by Marimon, McGratten, and Sargent [13]. We have to modify this algorithm for our environment and we show that people using our modified algorithm are able to learn in a wide variety of environments. In our game we can also show that the population will learn to play equilibrium.

Since this model of learning has never been analyzed in the stochastic evolution literature we also provide new results for this literature. We are interested to note that our feedback learning model allows us to avoid two of the three key assumptions of the classic evolutionary model. This model—developed by Kandori, Mailath, and Rob [8]—makes three behavioral assumptions: inertia, myopia, and mutations. Inertia means that not all players change their strategy in every period, in our model this happens naturally since players's beliefs change slowly. Myopia means that they choose the action that is optimal given the current distribution of the population, our learning model replaces this with choosing the action one thinks will give the highest continuation value. The only element of their model that we directly adopt is mutations. Mutations means that very rarely people change their strategy at random, in our model this means they develop radical new guesses about the continuation values. This drives the evolutionary process.

The problem of how people learned to use fiat money has troubled economists for a long time. As Dowd [3] states, economists have long understood the superiority of a monetary equilibrium over a barter one. The majority of the traditional literature has been focused on explaining why this is true. One common conclusion of this literature is that a primary benefit of money is in overcoming the famed double coincidence of wants—an idea first developed by Jevons [6]. In a barter economy you must both find someone who wants what you have and has what you want, while in a monetary economy you only need to find someone who wants what you have. This insight has been used in a spate of recent papers to try and explain the transition from a barter economy to a monetary economy. These papers all use variations of the model developed by Kiyotaki and Wright [9]. For example in Ritter [14] he analyzes the incentives of the body issuing the fiat currency, showing how they can optimally introduce their currency and under what conditions they will do so. However since his is an equilibrium analysis he can not overcome the problem that he has to assume people are willing to use money. He can not analyze or answer the fundamental question, is it possible that people can learn to trust money?

Many other papers in this literature have slightly modified Kiyotaki and Wright [9] to make it more consistent with history. For example Zhou and Green [20] allow for divisible money and goods;

and Williamson and Wright [18] allow for the possibility of private information. Another strain of this literature analyzes whether and how quickly players learn to use money. Marimon, McGratten, and Sargent [13] is the seminal paper in this literature; other examples are Basci [1], Selgin [16], and Williamson and Wright [18]. However, these models can not address the fundamental problem. Since all economies originally used barter how did they learn to trust money? Selgin [15] explicitly shows that if people do not believe money is valuable then there is not a monetary equilibrium.

To model the slow cultural transformation from barter to money we need to take account of beliefs and how they might change, and this requires a model of learning. We utilize a classifier-strength system first developed by Marimon, McGratten, and Sargent [13]. This is a learning system that learns syntactically simple rules to guide its performance in an arbitrary environment. We modify this basic learning algorithm so that it performs better in our environment. Following Basci [1] we require that there is a strength (estimate of the continuation value) for each state and action. Following Watkins [17] we have players experiment in order to be sure they learn. Since the equilibrium can change in our model we follow Benveniste, Metivier and Priouret [2] by not assuming that new information is discounted more and more over time. The resulting algorithm learns in a stable environment almost surely. With the standard algorithm Lettau and Uhlig [12] show people frequently learn to use a sub-optimal strategy. In our model agents learn the equilibrium path from interaction and positive feedback.

This model extends the literature of stochastic evolution because no previous analysis has utilized our learning algorithm. Previous work has always assumed that beliefs are determined by a type of moving average process—players look at the last k observations—in this paper our learning algorithm is more of an auto regressive process—beliefs today are a weighted average of what happens today and what has happened in the past. We note that the standard assumption—for example in Kandori, Mailath, and Rob [8]—would be equivalent to players assuming that they are in a steady state in our environment. Since most of the literature looks at static interactions it has continued to use this assumption. Only one other paper—to the authors knowledge—has looked at a dynamic game. Johnson, Levine, and Pessendorfer [7] look at altruism in a gift giving game where the reward for giving a gift today is receiving a gift tomorrow. Since the players in their model only live for two periods their assumption of how continuation values are calculated is equivalent to the one in this paper.

In the next section we introduce the model in three parts. First we introduce the basic interaction, then our model of learning, and then our model of evolution. The rest of the paper mimics this three step presentation. First we find the steady state equilibria of the model in section 3. Then we show when our learning algorithm will work, and show that this means society will learn in section 4. Finally we show which strategy will be stochastically stable (or evolutionarily successful) in section 5. We then conclude.

2 Model

2.1 The Model of Trading.

The model in Kiyotaki and Wright [9] elegantly represents the famous "double coincidence of wants" that is fundamental to the utility of money. The essential problem is that when people want to barter they must find someone who wants what they have (call this event a) and has what they want (call this event b), while if they use money they only have to find someone who has what they want (event b). Let the probability of event a and b be x (< 1) then barter occurs with probability x^2; money trading with probability x. Since $x > x^2$ there is a utility of money.

Now populate a world with I infinitely lived agents (I should be even and large). Assume that each period everyone is endowed with one unit of a good that they can not consume, thus they must trade. This good is not durable nor divisible, thus it must be either traded for a unit of consumption good or wasted. Next period each player will be endowed with a unit of consumption good again. This is a basic barter economy.

We will change the endowment of some players to one unit of fiat money (a fraction $\mu = \frac{i}{I}$, $i \in \{4, 6, 8, ...I - 4\}$). This good is not consumable but it is durable, enabling trade across periods. It is also a substitute for barter, thus someone with a unit of money will never be endowed with a unit of consumption good. Now we have a model where players can trade either using barter or fiat money.

Players will be matched with each other by equal likelihood, and when i and j meet, $\{i, j\} \subseteq I$ there are nine possible situations. Player i can have what j wants, call this c_j; or i can not have what j wants, call this c_{-j}; or i can have money, denoted m. Likewise player j can have what i wants—c_i, have a good i does not want—c_{-i}, or be endowed with money m. This means that the nine states of the world are $\Omega = \{c_j, c_{-j}, m\} \times \{c_i, c_{-i}, m\}$.

They will observe the good their trading partner has and then decide whether to trade (T) or not (N). $A_i = A_j = \{T, N\}, A = A_i \times A_j$. If both choose to trade then the exchange takes place, otherwise it does not.

If person i trades for a unit of c_i then he gets utility $U > 0$, if he trades for a unit of c_{-i} then he incurs a transaction cost of $\kappa > 0$, if he trades for a unit of money he gets no utility but does not incur a transaction cost.[1] Next period (which is discounted with a factor of $\beta \in (0, 1)$) someone with money can trade for a unit of his consumption good. Notice that this means that the interaction is dynamic, the value of money today depends on what occurs next period.

[1] If $\kappa = 0$ then our results are simpler wihtout changing in any significant way. However it is traditional to assume $\kappa > 0$.

2.2 The Model of Learning.

In a dynamic interaction the calculation of returns is extremely difficult. In a steady state it is relatively simple but the model will frequently not be in steady state. Out of steady state one must estimate the path of motion of the economy, which requires estimating the other individuals' decisions, and then calculate the optimal strategy. Notice that in order to estimate other individuals' decisions one must calculate their beliefs about the path of motion of the economy. In such a situation it would be surprising if people did not use rules of thumb in such an environment. One model developed by Marimon, McGratten, and Sargent [13] and Lettau and Uhlig [12] does exactly that. People begin with guesses of what the value of various strategies is, and then use an explicit ad-hoc learning algorithm to find the true values.

Each person begins with an initial belief (called a *strength*) of the value of an action pair $\alpha \in A$ at the state $\omega \in \Omega$. We denote this strength $S^0_{\alpha\omega}$, and the column vector of these strengths as S^0. Each period this person observes the state of the world and using S^t choose the action with the highest expected strength. Then he updates the strength of the state that actually occurs. Let $\tau \in (0,1)$ be the weight they put on their old strength, and $U_{\alpha\omega}$ be the instantaneous payoff from the action pair α at the state ω, then if $\{\alpha, \omega\}$ occur:

$$S^{t+1}_{a\omega} = (1-\tau)\, S^t_{a\omega} + \tau \left(U_{\alpha\omega} + \beta M^t_{\alpha\omega} S^t \right). \tag{1}$$

otherwise $S^{t+1}_{\alpha'\omega'} = S^t_{\alpha'\omega'}$. The row vector $M^t_{\alpha\omega}$ is the Markov transition probabilities in period t given $\{\alpha, \omega\}$. These transition probabilities will be a function of the distribution of strategies from last period, this player's strategy, and experiments (discussed below.)

There are several differences between this model and the general model in Lettau and Uhlig [12]. These differences are primarily motivated by a desire to be certain players can learn. We argue that this is a type of "long run rationality." Clearly in the short run agents will not be "rational" because they have the wrong priors, however if they also are not rational in the long run this suggests not only are their priors wrong but their model of the interaction is wrong. Most of the changes we make to the model are motivated by this difference. It might be possible, of course, to show that in some environments people who learn will be driven out of the population, but unless one specifies this model we feel the assumption of rationality should imply agents are able to learn.

The first of the changes we make to the model is that like Basci [1] we have a complete set of strengths and thus our agents choose the optimal action at each state. This means that if S^{t-1} are the true values and the distribution of strategies does not change then these agents are choosing the optimal strategy. The second difference is that we assume agents experiment—there is some fixed $\varepsilon_{ex} > 0$ that they choose the wrong action at every state. As Watkins [17] shows this enables learning because every state is reached with positive probability. Without this assumption a player could reach silly conclusions like never trading is optimal. All results in this paper will be for small ε_{ex}.[2]

A final change is dictated by the environment we are analyzing. In Marimon et al. [13] the weight given to new information (τ) is decreasing over time. As Benveniste, Metivier, and Prioret [2] point out this is only optimal if the state of the world does not change. In our model the state of the world is the distribution of strategies, and since their are multiple equilibria there can be multiple steady states for this distribution. If $\tau \to 0$ this means that players are ignoring new information, and if the new information might reflect a new state of the world the players will make worse and worse decisions over time.

However this last change makes our model unstable and must be compensated for by another change. In the standard model M^t is just what happened to agent i in period t. This is one reason that they assume that $\tau \to 0$, without this assumption the model may never converge. Thus we assume that M^t is based on the population distribution from last period. We are fairly certain our results would hold if M^t was based on a partial sample (like in Young [19]) but are certain our results fail if it is based only on personal experience.

We want to allow for very general priors on the strengths, all we require is that they are drawn from a compact support or that for all $\{\alpha, \omega\}$ $S^0_{\alpha\omega} \in \left[\frac{-\kappa}{1-\beta}, \frac{U}{1-\beta}\right]$.

2.3 The Model of Evolution.

When we add one further refinement to our model of learning it will be an evolutionary model. We will assume that with probability $\varepsilon_m > 0$ people *mutate*, or draw S^{t+1} from the initial distribution. We can justify this as checking to make sure if your historic beliefs are correct (one should consider low τ) but we notice that this addition to the model completely changes results. Because of this society will not settle down at any particular state. Instead society will constantly adopt new strategies. However as $\varepsilon_m \to 0$ society will spend most of their time around a particular state, some situation where everyone is best responding to the current distribution of strategies. This will mean that society is spending most of it's times near *limit sets.*

Definition 1 *A* limit set *is a set of distributions of strategies such that:*

1. *Any distribution of strategies in the limit set can be reached from any other without mutations.*
2. *Any distribution of strategies that is not in the limit set can only be reached via mutations.*

Intuitively this is a "best response cycle." I.e. the distribution of strategies is currently at one distribution in the limit set, and then everyone best responds to that and the distribution of

[2] We assume that $\varepsilon_{ex} < \min\left\{\frac{1}{2}, \frac{xU}{\varepsilon(1-x)+xU}\right\}$.

strategies is in a another. Clearly any Nash equilibrium is a limit set, but there might be others. In this interaction, however, there will not be.

We will be interested in the system when ε_m is very small. As $\varepsilon_m \to 0$ the relative likelihood of any other event dwarfs the likelihood of a mutation. Thus the number of mutations needed to transition from one limit set to another will completely determine how likely the transition is. Here this will be some fraction, η, of the population, and for given I we will need $\lceil \eta I \rceil$ mutations.[3] This dependency on I will be unimportant for the analysis, and we will say that a fraction η of the population must mutate.

The only transition that will matter in this paper is the transition that takes the least mutations, this transition determines the *radius*.

Definition 2 *The* radius *of a limit set π is the least fraction of the population that must change their strategy before in the long run a player can not use a strategy in π, we denote this $r^*(\pi)$.*

Intuitively the reader might find it easier to think of the radius as a "security level." In this game if fewer than this fraction stop using the current strategy then it can be ignored. A person can be secure that they are using the right strategy without qualification.

3 The Steady State Equilibria.

In this section we will ignore the model of learning and evolution for a while, and focus on the steady state equilibria of the underlying game. A simple manner in which to write down a strategy is to list the states at which this strategy trades. Then, for example, the *always trade* strategy is

$$AT = \left\{ \begin{array}{lll} \{c_j, c_i\} & \{c_j, c_{-i}\} & \{c_j, m\} \\ \{c_{-j}, c_i\} & \{c_{-j}, c_{-i}\} & \{c_{-j}, m\} \\ \{m, c_i\} & \{m, c_{-i}\} & \{m, m\} \end{array} \right\}$$

and the *never trade* strategy is:

$$NT = \{\emptyset\}$$

This is an equilibrium when $\varepsilon_{ex} = 0$, but for all $\varepsilon_{ex} > 0$ it is a dominated strategy.

Lemma 1 *Trading for c_i is a dominant action. So is not trading at the states $\{c_j, c_{-i}\}$ and $\{c_{-j}, c_{-i}\}$.*

Proof. See the appendix on page 14. ■

This pins down an optimal strategy at five of the nine states of the world. Furthermore the strategy at $\{m, m\}$ does not affect payoffs so we will ignore it. This leaves only three states left to

[3] For a real number x, $\lceil x \rceil$ is the smallest integer which is larger than x, and $\lfloor x \rfloor$ is the largest integer that is smaller.

specify a strategy at: $\{c_j, m\}$, $\{c_{-j}, m\}$, and $\{m, c_{-i}\}$. The important steady state equilibria each trade at a subset of these states. These are:

$$B = \left\{ \begin{array}{ll} \{c_j, c_i\} & \\ \{c_{-j}, c_i\} & \\ \{m, c_i\} & \{m, c_{-i}\} \end{array} \right\}, F = \left\{ \begin{array}{ll} \{c_j, c_i\} & \{c_j, m\} \\ \{c_{-j}, c_i\} & \{c_{-j}, m\} \\ \{m, c_i\} & \end{array} \right\}$$

B is a barter strategy, if you are using this strategy and holding money you will take anything to get rid of it. F is the fiat money trading strategy, in this strategy you will trade anything in order to hold money. If players make mistakes too frequently, or β is too low, then trading for c_{-i} when holding money might not be optimal. In this case instead of the B strategy everyone will use the $\tilde{B}$ strategy:

$$\tilde{B} = \left\{ \begin{array}{l} \{c_j, c_i\} \\ \{c_{-j}, c_i\} \\ \{m, c_i\} \end{array} \right\}$$

When we call something an "equilibrium" we mean it is a Nash equilibrium, but notice that since people make mistakes with strictly positive probability this is also a Subgame Perfect equilibrium and a Sequential equilibrium. Furthermore, since people base their decisions on strengths it is probability zero that they will be indifferent between actions, thus we assume people use pure strategies.[4] When we call a strategy a "pure population equilibrium" we mean that everyone in the population is using the same pure strategy. When we call a strategy "mixed population equilibrium" this means that some people are using different pure strategies.

Characterizing the equilibria is easiest when one knows the continuation values of holding a consumption good (c) or money (m).

Lemma 2 (Charcterization) *The continuation values from holding money and a consumption good are respectively:*

$$\begin{aligned} V(m) &= \frac{1}{\Delta}\left(\left((1-\beta)+\beta P_m(c)\right)\left(P_{c_i}(m)U - P_{c_{-i}}(m)\kappa\right) + \beta\left(P_{c_i}(m)+P_{c_{-i}}(m)\right)\left(P_{c_i}(c)U - P_{c_{-i}}(c)\kappa\right)\right) \\ V(c) &= \frac{1}{\Delta}\left(\beta P_m(c)\left(P_{c_i}(m)U - P_{c_{-i}}(m)\kappa\right) + \left((1-\beta)+\beta\left(P_{c_i}(m)+P_{c_{-i}}(m)\right)\right)\left(P_{c_i}(c)U - P_{c_{-i}}(c)\kappa\right)\right) \end{aligned}$$

where $P_g(s)$ *is the probability of trading for good* $g \in \{c_i, c_{-i}, m\}$ *when holding* $s \in \{c, m\}$ *and* $\Delta = (1-\beta)(1-\beta) + \beta\left(P_{c_i}(m) + P_{c_{-i}}(m) + P_m(c)\right)$

Proof. One can easily calculate that the flow values of these states are:

$$\begin{aligned} V(m) &= P_{c_i}(m)(U+\beta V(c)) + P_{c_{-i}}(m)(-\kappa+\beta V(c)) + \left(1-P_{c_i}(m)-P_{c_{-i}}(m)\right)(0+\beta V(m)) \\ V(c) &= P_{c_i}(c)(U+\beta V(c)) + P_{c_{-i}}(c)(-\kappa+\beta V(c)) + P_m(c)(0+\beta V(m)) \\ &\quad + \left(1-P_{c_i}(c)-P_{c_{-i}}(c)-P_m(c)\right)(0+\beta V(c)) \end{aligned}$$

and after algebraic manipulation these can be turned into the continuation values above. ■

[4] If a player is indifferent assume that they will trade.

Lemma 3 *F and B are the pure population equilibria unless players are impatient or experiment a lot. In that case if $\beta < \frac{\kappa}{(1-\mu)x^2U+\kappa}$ or $\varepsilon_{ex} > \varepsilon_{ex}^*$ then $\tilde{B}$ is an equilibrium instead of B.*

Proof. See the appendix on page 14. ■

In the rest of the paper we will assume that B is the pure population barter equilibrium without loss of generality, indeed our proofs would be easier if trading at $\{m, c_{-i}\}$ were never a best response.

Quite interestingly there is a continuum of mixed population equilibria, and in none of them do players use the strategy B. The reason for this is that B trades at the state $\{m, c_{-i}\}$ because the value of holding money is so low that the cost of κ is worth the benefit of holding a consumption good. In a mixed population equilibrium the value of holding money is higher, and thus it is not a good idea to trade money for c_{-i}. There is a continuum of mixed population equilibria because in such an equilibrium players are indifferent between using the strategy F and $\tilde{F}$:

$$\tilde{F} = \left\{ \begin{array}{ll} \{c_j, c_i\} & \{c_j, m\} \\ \{c_{-j}, c_i\} & \\ \{m, c_i\} & \end{array} \right\}$$

but the value of holding money is decreasing the more people use the strategy F, and thus for each mixture we have a different equilibrium.

Notice that players might also use strategies other than $\tilde{B}$, F, and $\tilde{F}$ in a mixed population equilibrium. We might, for example, have some people who only trade for money when holding c_{-j}. Thus the entire space of equilibria is complicated, and not worth specifying. The only important equilibria in our analysis are the extremes, where all players who trade for money are using either the strategies F or $\tilde{F}$. Let $\pi_m(f)$ be the fraction of the people holding money who use strategy $f \in \left\{F, \tilde{F}\right\}$ and $\pi_c(f)$ be the equivalent portion holding the consumption good.

Lemma 4 *The extreme mixed population equilibria are $\pi_c(f) = x(1 - o_c(\varepsilon_{ex}|f))$, $\pi_m(f) = 1 - o_m(\varepsilon_{ex}|f)$ where $\lim_{\varepsilon_{ex}\to 0} o_c(\varepsilon_{ex}|f) = 0$, $\lim_{\varepsilon_{ex}\to 0} o_m(\varepsilon_{ex}|f) = 0$ and*

$$\begin{aligned} o_c\left(\varepsilon_{ex}|\tilde{F}\right) &= \frac{\varepsilon_{ex}}{1-\varepsilon_{ex}}\frac{(1-x)\kappa}{xU} \\ o_c(\varepsilon_{ex}|F) &= \frac{1}{2}\left(1+\frac{1}{x}\frac{o_{io}(\varepsilon_{ex}|F)}{(1-o_{io}(\varepsilon_{ex}|F))}\right) - \sqrt{\frac{1}{4}\left(1+\frac{1}{x}\frac{o_{io}(\varepsilon_{ex}|F)}{(1-o_{io}(\varepsilon_{ex}|F))}\right)^2 + \frac{\varepsilon_{ex}}{1-\varepsilon_{ex}}\frac{(1-x)^2\kappa}{x^2U}} \\ o_m(\varepsilon_{ex}|f) &= \frac{o_{io}(\varepsilon_{ex}|f)((1-\pi_c(f)))}{\pi_c(f)+o_{io}(\varepsilon_{ex}|f)(1-\pi_c(f))} \end{aligned}$$

where:

$$o_{io}(\varepsilon_{ex}|f) = \varepsilon_{ex}\frac{x(1-\varepsilon_{ex})+(1-x)\varepsilon_{ex}}{x(1-\varepsilon_{ex})^2+(1-x)\varepsilon_{ex}\alpha(f)}.$$

where $\alpha(F) = (1-\varepsilon_{ex})$, $\alpha\left(\tilde{F}\right) = \varepsilon_{ex}$ and $\lim_{\varepsilon_{ex}\to 0} o_{io}(\varepsilon_{ex}|f) = 0$. Furthermore $\pi_c(F) > \pi_c\left(\tilde{F}\right)$.

Proof. See the appendix on page 15. ■

Notice two interesting facts about these equilibria. First they are not stable, if in any period less than $\pi_c(f)$ fiat money traders are in the consumption state a barter strategy is a strict best response. As well notice that the equilibrium condition depends only on $\pi_c(f)$. In fact (as we will have to assume later) they could have incorrect beliefs about $\pi_m(f)$ and as long as $\pi_c(f)$ is at it's critical value this will be sufficient.

4 Learning.

In this section we will specify the steady states of our learning algorithm. The first proposition merely establishes that we have met our criteria of "long run rationality": players can learn. If this result did not hold then we would argue that we had misspecified the learning algorithm. The second proposition is not a result we assumed a priori, this proposition shows that the population can learn simultaneously. The results of the first proposition we are certain work in very general environments since our learning algorithm is similar to a Gauss-Seidel learning algorithm, the results of the second proposition may not hold in all games.

One key decision we have to make at this point is what information people will have about the current distribution of strategies, and how they should interpret this information. One option is to have them know the distribution of strategies dependent on whether people are holding money or not. This would be the most information we would want them to have, and we would argue too much. While explicitly they will know the distribution in each period implicitly we wish the distribution to represent some general information they have gathered about the world. This information might be old and not completely trustworthy. Thus they should not be able to deduce exactly how many people are using a certain strategy today, and whether those people are holding a consumption good or a money. Furthermore is it really reasonable to assume they know the distribution over strategies? A strategy is not directly observable, it requires many observations of a person to know his entire strategy. Thus we will only let them know the fraction of people who trade in each state, $\eta = \{\eta_\omega\}_{\omega \in \Omega}$.

This assumption, however, presents us with a secondary problem. One which will not strongly affect analysis but must be addressed. Given the distribution of actions what are players' beliefs about the distribution of strategies? The simplest assumption would be that this distribution is independent, but these beliefs would be wrong any time the society is not in an equilibrium, or they fail a "reasonability" test. We notice that our learning algorithm induces a correlation between actions, thus we assume that players understand this correlation. This leads us to four simple conclusions. First if a player is taking one dominant action then they are taking all dominant actions. If they are not taking all of their dominant actions and they trade for c_i when holding c_j

(or c_{-j}) then they also trade when holding c_{-j} (or c_j). If they trade at $\{m, c_{-i}\}$ then they do not trade at $\{c_j, m\}$ or $\{c_{-j}, m\}$. Finally if they trade at $\{c_j, m\}$ then they trade at $\{c_{-j}, m\}$ if possible.

These are all fairly transparent, the final issue we must address is the strategies of players who are holding money. We have two options, either we could have them calculate $\pi_m(f)$ or we could simply have them assume that with probability one someone who is holding money is using a fiat money strategy. The former assumption requires them to be able to do complicated calculations, and furthermore would not be that different from the latter when ε_{ex} is small. Thus we assume that everyone who trades at $\{c_j, m\}$ and $\{c_{-j}, m\}$ is holding money, and then people who only trade at $\{c_j, m\}$ are next most likely, finally this population is made up of people (if any) who trade only at $\{c_{-j}, m\}$. Technically if $P(T|c., m)$ is the probability that someone holding a consumption good will trade it for money, then $P(T|c., m) = \max\left\{\max\left\{\eta_{c_j,m}, \eta_{c_{-j},m}\right\} - \mu, 0\right\}$.

With these decisions made we can now focus on our learning algorithm. The convergence of our learning algorithm is immediate since $\varepsilon_{ex} > 0$. With this assumption our problem is equivalent to the Gauss-Seidel method of value calculation. Thus convergence is guaranteed.

Proposition 1 *If only one person learns then he will learn all of his strict best responses.*

Proof. See the appendix on page 7. ■

As we said above there is no guarantee that one person being able to learn implies that society is able to learn. Kushner [11] has worked on games where the population has two players and proved that the Gauss-Seidel algorithm converges to a steady state, but there are no results appropriate for our environment. Thus here we address the simpler problem of showing that in this game the population learns.

Proposition 2 *There is an $\underline{I}$ such that if $I \geq \underline{I}$ society almost surely converges to an equilibrium.*

Proof. See the appendix on page 7. ■

Notice that this proposition would be simpler if players knew the current distribution of strategies. In that case we could have all of one group trade for money, and at the current distribution either a fiat or barter strategy would be a best response.

5 Evolution.

We now will show when evolution will result in a society using fiat money. The first step of our analysis is to find the radii of our equilibria. Notice that in our model we do not only need to know what strategies people are using but what their strengths are. However as $\varepsilon_m \to 0$ if a society is in a limit set they will be there for a long time, thus their strengths will be nearly equal to the true values.

Lemma 5 *$\forall I\ \exists \bar{\varepsilon}_{ex} > 0$ such that if $\varepsilon_{ex} \leq \bar{\varepsilon}_{ex}$ then $r^*(F) = (1-\pi_c(F))(1-\mu)$ and $r^*(B) = (1-\mu)\pi_c\left(\tilde{F}\right) + \mu$, and if π^* is a mixed population equilibrium $r^*(\pi^*) = \lim_{\xi \to 0} \xi$.*

Proof. See the appendix on page 16. ■

In this lemma one can see that if we are interested in the rise of fiat money then we have made the worst assumption about how players estimate the fraction of people using the fiat money strategy that are holding money. We have assumed that they always over estimate the fraction holding money, if we required them to calculate the correct probability or believe that strategy and the good being held were independent this would make the transition from barter to fiat money easier.

The final result can be explained very succinctly. The most interesting case is when ε_{ex} is nearly zero and the society is complex—x is nearly zero—or it is hard to find someone who wants to trade with you. In this case the only important parameter is μ, and if $\mu < \frac{1}{2}$ then fiat money will be evolutionarily successful—technically *stochastically stable.*

Theorem 1 *If $I \geq \underline{I}$ and $\varepsilon_{ex} \leq \bar{\varepsilon}_{ex}$ and there is less money than goods ($\mu < \frac{1}{2}$) then in complex societies non-convertible fiat money is stochastically stable when ε_{ex} is nearly zero. More generally this is true when:*

$$\mu < \frac{1 - 2x\left(1 - o\left(\varepsilon_{ex}\right)\right)}{2 - 2x\left(1 - o\left(\varepsilon_{ex}\right)\right)} \tag{2}$$

where $o\left(\varepsilon_{ex}\right) = \frac{o_c(\varepsilon_{ex}|F) + o_c\left(\varepsilon_{ex}|\tilde{F}\right)}{2}$.

Proof. See the appendix on page 17. ■

Notice that almost all trades occur using money. Especially in complex societies the number that use barter will be trivial. The general formula in the fiat money equilibrium is:

$$P\left(\text{trade uses money}\right) > \frac{2\mu s}{2\mu x + x^2\left(1-\mu\right)} \tag{3}$$

thus 95% of trades will use money if $\mu \geq \frac{19x}{2+19x}$. For example if $x \leq \frac{1}{20}$ then $\mu = \frac{1}{3}$ is sufficient.

6 Conclusion.

It has long been wondered how non-convertible fiat money came to be trusted. We all recognize the optimality of money, this has been explained in many frameworks and many different models. The question is how people came to trust a worthless piece of paper, enough that they would give up real goods for these wooden nickels. We show that in complex economies as long as the market is not flooded with this worthless currency people will learn to trust it.

In fact the most surprising fact, given our analysis, is that non-convertible fiat money has taken so long to arise. As Dowd **??** points out fiat money has never arisen in a barter economy historically.

It has always been preceded by commodity money—with gold being the primary example, then by fractionally based commodity money, and then pure fiat money has only arisen through government fiat—hence the name. This motivates several extensions of this model. First would be to also allow for a commodity money—or money that is at least partially backed by real goods, but such a model must also explain why people don't just use currency and fiat money. This relates to another issue that Dowd [3] raises, how do people choose between currencies? We believe that it is uncommon in history that multiple currencies are commonly accepted—early United States history being a counter example. Why is this? One part of this analysis would have to be exchange rates, and for that we would need to have divisible goods and currency as in Zhou and Green [20]. But we also wonder if perhaps such a model should include a cost of accepting multiple currencies. How could this be included in the model? These are open questions that should be pursued. Another line of analysis is looking at the incentives of the party issuing the currency. In our model the stock of currency is a given, but in Ritter [14] this depends on the incentives of the issuing body. Our results here would require that this body represents a significant fraction of the population. We hope that the our basic results add to the understanding of this issue, and that with further analysis we can develop a complete understanding of the rise of money.

7 Appendix

.

Proof of Lemma 1. If i trades at $\{c_{-j}, c_{-i}\}$ and j trades with probability $\eta_{j\cdot} > 0$ the trade will take place with probability $(1-\varepsilon_{ex})\,\eta_{j\cdot}$. If i does not trade then the trade will take place with probability $\varepsilon_{ex}\eta_{j\cdot} < (1-\varepsilon_{ex})\,\eta_{j\cdot}$. Thus trading is not optimal since it gives an instantaneous loss of κ and in the next period has the same continuation value as not trading. The same argument also explains not trading at $\{c_j, c_{-i}\}$ and inverting the argument explains trading at $\{c_j, c_i\}$ and $\{c_{-j}, c_i\}$.

Now consider not trading at $\{m, c_i\}$. To clarify the following argument given a strategy σ let the value of holding a unit of money between periods be $V(m|\sigma)$, and $V(c|\sigma)$ be the value of holding a unit of a consumption good. Assume that i meets someone with a unit of c_i, then if i trades they get $U + \beta V(c|\sigma)$ if they do not trade then they get $0 + \beta V(m|\sigma)$. The proof will be done when we establish upper bounds for $V(m|\sigma)$.

If the strategy calls for trading at $\{m, c_{-i}\}$ then by optimality we know that $-\kappa + \beta V(c|\sigma) \geq 0 + \beta V(m|\sigma)$ and we are done. Thus the strategy must require no trading at $\{m, c_{-i}\}$ and $\{m, c_i\}$. If $\varepsilon_{ex} = 0$ then $V(m|\sigma) = 0 < U + \beta V(c|\sigma)$ since $V(c|\sigma) \geq 0$ by optimality. Thus if there is a ε_{ex} such that $0 + \beta V(m|\sigma) > U + \beta V(c|\sigma)$ $V(m|\sigma)$ must be increasing in ε_{ex}. But

$$\begin{aligned} V(m|\sigma) &= (1-(1-\mu)\,\varepsilon_{ex})\,(0+\beta V(m|\sigma)) + (1-\mu)\,\varepsilon_{ex}\,(xU-(1-x)\,\kappa+\beta V(c|\sigma)) \quad (4)\\ \frac{\partial V(m|\sigma)}{\partial \varepsilon_{ex}} &= (1-\mu)\,((xU-(1-x)\,\kappa+\beta V(c|\sigma)) - (0+\beta V(m|\sigma))) \end{aligned}$$

and $\frac{\partial V(m|\sigma)}{\partial \varepsilon_{ex}} \leq 0$ when $xU - (1-x)\,\kappa + \beta V(c|\sigma) \leq 0 + \beta V(m|\sigma)$, thus the highest $0 + \beta V(m|\sigma)$ can be is $xU - (1-x)\,\kappa + \beta V(c|\sigma)$ which is strictly less than $U + \beta V(c|\sigma)$. Therefore for all ε_{ex}, $U + \beta V(c|\sigma) > 0 + \beta V(m|\sigma)$. ■

.

Proof of Lemma 3.

In order for the F strategy to be a best response we need that:

$$V(m|F) \geq V(c|F) \quad (5)$$

or

$$(P_{c_i}(m) - P_{c_i}(c))\,U \geq \left(P_{c_{-i}}(m) - P_{c_{-i}}(c)\right)\kappa \quad (6)$$

and this requires that:

$$\begin{aligned} (P_{c_i}(m) - P_{c_i}(c))\,U &\geq \left(P_{c_{-i}}(m) - P_{c_{-i}}(c)\right)\kappa \quad (7)\\ (1-\mu)\,(1-\varepsilon_{ex})\,x\,(1-x)\,(1-2\varepsilon_{ex})\,U &\geq -(1-\mu)\,(1-x)\,x\varepsilon_{ex}\,(1-2\varepsilon_{ex})\,\kappa \end{aligned}$$

since the right hand side is negative this will be true as long as $\varepsilon_{ex} < \frac{1}{2}$. In order for $\tilde{B}$ to be a best response we need that:

$$\begin{aligned} (P_{c_i}(c) - P_{c_i}(m))\,U &\geq \left(P_{c_{-i}}(c) - P_{c_{-i}}(m)\right)\kappa \quad (8)\\ (1-\mu)\,(1-\varepsilon_{ex})\,x^2\,(1-2\varepsilon_{ex})\,U &\geq (1-\mu)\,(1-x)\,x\varepsilon_{ex}\,(1-2\varepsilon_{ex})\,\kappa \end{aligned}$$

and this requires that $\varepsilon_{ex} < \min\left\{\frac{1}{2}, \frac{xU}{xU+(1-x)\kappa}\right\}$ in order for B to be a best response we also need that:

$$\beta V(m) \leq -\kappa + \beta V(c) \quad (9)$$

or

$$\left(P_{c_{-i}}(c)+\frac{1-\beta}{\beta}+P_{c_i}(m)+P_m(c)\right)\kappa \leq \left(P_{c_i}(c)-P_{c_i}(m)\right)U \tag{10}$$

If $\varepsilon_{ex}=0$ and $\beta>\frac{\kappa}{(1-\mu)x^2U+\kappa}$ one can easily show this is true. One can also show that the left hand side is increasing in ε_{ex} and the right hand side is decreasing and that when $\varepsilon_{ex}=\frac{1}{2}$ the right hand side is zero and the left hand side positive. Thus there is a maximal ε_{ex} for each β such that this holds, and this is the ε_{ex}^* in the statement of the lemma. ■

.

Proof of Lemma 4. In a mixed population equilibrium one can show that:

$$P_{c_i}(m)U-P_{c_{-i}}(m)\kappa=P_{c_i}(c)U-P_{c_{-i}}(c)\kappa \qquad \text{(Mixed Pop Condition)}$$

After solving for the probabilities and substitution this gives us:

$$\pi_c(f)=x+\frac{\varepsilon_{ex}}{1-\varepsilon_{ex}}\frac{(1-x)\kappa}{xU}\left(\frac{\alpha(f)-\varepsilon_{ex}}{1-2\varepsilon_{ex}}\pi_m(f)-x\right) \tag{11}$$

where $\alpha(F)=(1-\varepsilon_{ex})$, $\alpha\left(\tilde{F}\right)=\varepsilon_{ex}$. Note that from this equation we can see that $\pi_c(F)=\pi_c\left(\tilde{F}\right)+\frac{\varepsilon_{ex}}{1-\varepsilon_{ex}}\frac{(1-x)\kappa}{xU}\pi_m(F)$, and that $\pi_c\left(\tilde{F}\right)$ is independent of $\pi_m\left(\tilde{F}\right)$, thus we know that $\pi_c(F)<\pi_c\left(\tilde{F}\right)$. The equation equalizing the inflow and outflow of players in the money state is:

$$\begin{aligned}\pi_m(f)(1-\pi_c(f))\left(x(1-\varepsilon_{ex})\varepsilon_{ex}+(1-x)\varepsilon_{ex}^2\right) &= \pi_c(f)(1-\pi_m(f))\left(x(1-\varepsilon_{ex})^2+(1-x)\varepsilon_{ex}\alpha(f)\right) \\ \frac{\pi_m(f)}{1-\pi_m(f)}o_{io}(\varepsilon_{ex}|f) &= \frac{\pi_c(f)}{1-\pi_c(f)}\end{aligned} \tag{12}$$

and after manipulation these are equivalent to the conditions in Lemma 4. ■

.

Proof of Proposition 1.

The proof is done when we rewrite our algorithm so that it is transparently a Gauss-Seidel learning algorithm. To do this choose any arbitrary sequence over Ω and any arbitrary sequence over A for each Ω, being sure that the last action pair is $\{T,T\}$, since $\varepsilon_{ex}>0$ this sequence has strictly positive probability. Then note that since others are not changing their strategy $M^t=M(\sigma_i)$ where σ_i is this player's strategy. Then we can rewrite our formula for $S_{\alpha\omega}^{t+1}$ as:

$$\begin{aligned}S_{a\omega}^t &= (1-\tau)S_{a\omega}^{t-1}+\tau\left(U_{\alpha\omega}+\beta M_{\alpha\omega}(\sigma_i)S^{t-1}\right) \\ S_{a\omega}^t &= \tau U_{\alpha\omega}+D_{\alpha\omega}(\sigma_i)S^{t-1}\end{aligned} \tag{13}$$

where

$$D(\sigma_i)=(1-\tau)I_{|A|*|\Omega|}+\tau\beta M(\sigma_i) \tag{14}$$

(I_Z is an identity matrix with Z rows) and our agent chooses $a_i(\omega)$ as

$$a_i(\omega)\in\arg\max_{\alpha_i\in\{T,N\}}E_{\alpha_j}\left[S_{\alpha_i\alpha_j\omega}^t\right] \tag{15}$$

combining these two facts we can write:

$$S_\omega^k = \max_{\alpha_i \in \{T,N\}} E_{\alpha_j}\left[\tau U_{\alpha_i \alpha_j \omega} + D\left(\sigma_i\left(\alpha_i\right)\right) S^{t-1}\right] \tag{16}$$

where $k = \left\lceil \frac{t}{|A|*|\Omega|} \right\rceil$ where $\sigma_i\left(\alpha_i\right)$ reflects the dependence of σ_i on the player choosing α_i at state ω. As Kushner [10] shows since $\Sigma_j D_{ij} = (1-\tau) + \tau\beta < 1$ the vector $\left\{S_\omega^k\right\}_{\omega\in\Omega}$ converges to the optimal value function, thus our $\{S_{a\omega}^t\}_{\alpha\omega\in A\times\Omega}$ converges to the true values of the optimal strategies. (Note that the strategy will converge in finite time. Once the strategy converges we can ignore the σ_i in equation (13) then we can be sure that the $\{S_{a\omega}^t\}_{\alpha\omega\in A\times\Omega}$ converge to their true value.)

Convergence to the optimal strategy will happen in finite time on this sequence. Let $\zeta > 0$ be the least difference in continuation values for actions that have a strict best response. Then since the strengths converge to the true values there is a finite T such that the difference between all optimal values and the current strengths are strictly less than $\frac{1}{2}\zeta$ and will be forever after. At this point the player will be using the optimal strategy.

Since T is finite this subsequence is finite and in each period a player has a strictly positive probability of entering it. Thus a player learns almost surely. ■

.

Proof of Proposition 2. Choose $\underline{I}$ large enough so that if a player's best response is not independent of what four other are doing then there are at least four people who are trading at $\{c_j, m\}$ and four people who are not.

Notice that in order to implement Proposition 1 we need to have a group of two people who are holding money and two people who are holding a consumption good. For this Proposition we also need them to be using the right strategies.

If the best response of any player is independent of what four people are doing then we can choose the group at random, being sure that as many as possible are not best responding at the state $\{c_j, m\}$. Otherwise we need the four people in the group to all not be best responding at the state $\{c_j, m\}$.

To achieve a group of this sort implement a finite sequence of trades so that there are two groups—one of whom is trading at $\{c_j, m\}$ and one of whom is not. If this is not possible in a finite number of trades then everyone must be either trading or not at $\{c_j, m\}$ and we are done, thus assume that it is possible. Therefore without loss of generality select the group that is not best responding at $\{c_j, m\}$.

Have the people in the group go into a finite subsequence as specified in Proposition 1, at the end of this subsequence all players in the group will have learned to best respond at all states. While this group is in this subsequence have everyone not in this group be in matches that either reinforce dominant actions or involve trading money.

At this point everyone's best response will be independent of what four people do, thus we can form a random group of four with as many non-best responders as possible and repeat this process. By iteration we can be sure everyone will best respond, and then we can have everyone learn one final time so that their value functions will coincide with the fact that everyone is now using a best response. ■

.

Proof of Lemma 5. We first note that if we are in a mixed population equilibrium we can have one person switch to either F or $\tilde{B}$ and then we can use Proposition 2 to converge to either the F or B equilibrium.

First notice that if all players are using the B strategy that for all $\lambda > 0$ and $\kappa > 0$ there is a $\varepsilon_m > 0$ such that if the probability of a mutation is less than this ε_m $P\left(\left[S^t - V\left(B\right)\right] < \kappa\right) > 1 - \lambda$

for all people in the society. This is an immediate implication of Proposition 1 and the fact that mutations only occur with probability ε_m. Thus we can assume all people are using the strategy B and that $[S^t - V(B)] < \kappa$, where the value of κ will depend on the population size and be selected below.

Then clearly if $\left\lceil I\left((1-\mu)\pi_c\left(\tilde{F}\right)+\mu\right)\right\rceil$ people mutate to the $\tilde{F}$ strategy then this is sufficient by Proposition 2. Therefore assume, in contradiction, that less than $\left\lceil I\left((1-\mu)\pi_c\left(\tilde{F}\right)+\mu\right)\right\rceil$ people mutate. We will show that behavior will return to all people using B.

Clearly the most difficult case will be when mutants do not learn their true values, so have them either trade using dominant actions or trade money. Furthermore it will be insufficient if someone using the B strategy switches to the $\tilde{B}$ strategy, so what we must show is that there is no sequence where someone using the strategy $\tilde{B}$ will switch to $\tilde{F}$.

Consider the most difficult case, when $\left\lceil I\left((1-\mu)\pi_c\left(\tilde{F}\right)+\mu\right)\right\rceil - 1$ people mutate to the $\tilde{F}$ strategy. Then we know that since I is discrete $V\left(c|\tilde{B}\right) - V\left(m|\tilde{F}\right) \geq 2\kappa > 0$ for all ε_{ex}. Let this be the κ we used above. Notice as well that in this case the value of every strength will increase since when some people are using the $\tilde{F}$ strategy this increases the value of money. This means that for all $\{\alpha\omega\}$ $S^t_{a\omega} - S^*_{a\omega}$ is negative, thus every time $\{\alpha\omega\}$ is updated $S^t_{a\omega} - S^*_{a\omega}$ will be negative, or all strengths will approach their new limit from below.

What we must show is that trading at the states $\{c_j, m\}$ and $\{c_{-j}, m\}$ is not optimal for someone who is using the B strategy. Now for someone using the B strategy the initial value of the strengths $\left\{S^t_{\{N,T\}\{c_j,m\}}, S^t_{\{N,N\}\{c_j,m\}}, S^t_{\{N,T\}\{c_j,m\}}\right\}$ are all at least $(1-\varepsilon_{ex})\beta x^2 \frac{(1-\mu)}{1-\beta} U + \varepsilon_{ex}\beta Z_1(\varepsilon_{ex}) - \kappa$ where $Z_1(\varepsilon_{ex})$ is a function such that $\lim_{\varepsilon_{ex}\to 0} Z_1(\varepsilon_{ex}) = 0$. As well given the new distribution of strategies $V\left(c|\tilde{B}\right) = (1-\varepsilon_{ex}) x^2 \frac{(1-\mu)}{1-\beta} U + \varepsilon_{ex} Z_2(\varepsilon_{ex})$ where $\lim_{\varepsilon_{ex}\to 0} Z_2(\varepsilon_{ex}) = Z_2 > 0$.

Since $V\left(m|\tilde{F}\right) \leq V\left(c|\tilde{B}\right) - 2\kappa \leq (1-\varepsilon_{ex}) x^2 \frac{(1-\mu)}{1-\beta} U + \varepsilon_{ex} Z_2(\varepsilon_{ex}) - \kappa$ and the limiting value of $S^t_{\{T,T\}\{c_j,m\}}$ is $\beta V\left(m|\tilde{F}\right)$ there clearly is a $\bar{\varepsilon}_{ex}$ such that

$$\begin{aligned}
(1-\bar{\varepsilon}_{ex})\beta x^2 \frac{(1-\mu)}{1-\beta} U + \bar{\varepsilon}_{ex}\beta Z_2(\bar{\varepsilon}_{ex}) - \beta\kappa &< (1-\bar{\varepsilon}_{ex})\beta x^2 \frac{(1-\mu)}{1-\beta} U + \bar{\varepsilon}_{ex}\beta Z_1(\bar{\varepsilon}_{ex}) \\
\bar{\varepsilon}_{ex}\left(Z_2(\bar{\varepsilon}_{ex}) - Z_1(\bar{\varepsilon}_{ex})\right) &< \kappa
\end{aligned} \tag{17}$$

and if this is true then even if $S^t_{\{N,T\}\{c_j,m\}}$ is at it's initial value and $S^t_{\{T,T\}\{c_j,m\}}$ increases to it's limiting value it will still be true that $S^t_{\{T,T\}\{c_j,m\}} < S^t_{\{N,T\}\{c_j,m\}}$.

Thus all people will choose either the strategy B or $\tilde{B}$ even if the learning process favors the strategy $\tilde{F}$. This argument is clearly reversible because the only difference between $\tilde{B}$ and $\tilde{F}$ is whether you trade at the state $\{c_j m\}$. Note that if people are using F instead of $\tilde{F}$ this only has an impact on order of ε_{ex}, thus for small enough ε_{ex} the statement is correct. ■

.

Proof of Theorem 1. From Young [19] we know that a limit set (equilibrium in this game) is stochastically stable if it has the minimum stochastic potential. From Hasker [5] we know that the stochastic potential of an equilibrium σ^* is:

$$c^*_{\sigma^*} = c^*(E) - r^*(\sigma^*) + ca(\sigma^*) \tag{18}$$

where E is the emergent seed and $ca(\cdot)$ is the core attraction rate. Notice that in order to find the stochastically stable strategy we do not actually need to find the emergent seed (which we will not), and we will show that for the relevant equilibria $ca(\sigma^*) \leq \frac{1}{I}$.

First we will dispatch with the possibility that a mixed population equilibrium could be stochastically stable. Assume to the contrary that π^* has the lowest stochastic potential, notice that we can transition from the π^* equilibrium to F with only one mutation, this means that:

$$\begin{aligned} c_F^* &\leq c_{\pi^*}^* + \frac{1}{I} - r^*(F) \\ &\leq c_{\pi^*}^* + \frac{1}{I} - \frac{\lceil I(1-\pi_c(F))(1-\mu)\rceil}{I} \end{aligned} \quad (19)$$

and clearly this is negative for reasonable I.

Now the core attraction rate is the number of mutations to get from the core to the given strategy. The core must contain more than one limit set thus it either contains both F and B or it contains a mixed strategy equilibrium. This tells us that the maximum of the core attraction rate is $\frac{1}{I}$ as claimed above.

Given this F must be stochastically stable if:

$$-\frac{\lceil I(1-\pi_c(F))(1-\mu)\rceil}{I} + \frac{1}{I} < -\frac{\left\lceil I\left((1-\mu)\pi_c\left(\tilde{F}\right)+\mu\right)\right\rceil}{I} \quad (20)$$

which will be true for large enough I if $(1-\pi_c(F))(1-\mu) > (1-\mu)\pi_c\left(\tilde{F}\right)+\mu$, which is the condition above. ■

References

[1] Basci, Erdem. (1999) "Learning by Imitation"; *Journal of Economic Dynamics and Control*, 23:1569-85.

[2] Benveniste, Albert; Michel Metivier; and Pierre Priouret. (1987). *Adaptive Algorithms and Stochastic Approximations.* Springer-Verlag; Berlin.

[3] Dowd, K. (2001) "The Emergence of Fiat Money: A Reconsideration" *Cato Journal* 20: 467-76.

[4] Clower, Robert W.(1984) *Money and Markets: Essays by Robert W. Clower Edited by Donald A. Walker.* Cambridge: Cambridge University Press.

[5] Hasker, Kevin. (2004) "The Emergent Seed: Simplifying the Analysis of Dynamic Evolution." Bilkent Economics Department Discussion Paper 04-06.

[6] Jevons, W. Stanley. (1875). *Money and the Mechanism of Exchange.* King: London.

[7] Johnson, Philip; David Levine, and Wolfgang Pessendorfer. (2001): "Evolution and Information in a Gift Giving Game." *Journal of Economic Theory,* 100:1-21.

[8] Kandori, Michihiro; George Mailath, and Rafael Rob. (1993): "Learning Mutation, and Long Run Equilibria in Games." *Econometrica*, 61:9-56

[9] Kiyotaki, Nobuhiro and Randall Wright. (1993). "A Search-Theoretic Approach to Monetary Economics. *American Economic Review.* 83:63-77.

[10] Kushner, Harold J. (1972) *Introduction to Stochastic Control Theory.* Holt, Reinhart, and Winston; New York.

[11] Kushner, Harold J. (2005) "Numerical Approximations for Non-Zero-Sum Stochastic Differential Games." *mimeo*, available at: http://www.dam.brown.edu/lcds/papers/Num-Games-non-zero.pdf

[12] Lettau, Martin and Harald Uhlig. (1999). "Rules of Thumb versus Dynamic Programming." *American Economic Review.* 89:148-174.

[13] Marimon, R.; E. McGratten; T.J. Sargent. (1990). "Money as a Medium of Exchange in an Economy with Artificially Intelligent Agents." *Journal of Economic Dynamics and Control.* 14:329-373.

[14] Ritter, Joseph. (1995). "The Transition from Barter to Fiat Money." *American Economic Review.* 85:134-149

[15] Selgin, G.A. (1994) "On Ensuring the Acceptability of a New Fiat Money." *Journal of Money, Credit, and Banking* 26: 808-26.

[16] Selgin, George. (2003)."Adaptive Learning and the Transition to Fiat Money". *Economic Journal*, 113:147-65.

[17] Watkins, C. (1989). "Learning with Delayed Rewards." Ph.D. thesis, Cambridge University, Psychology Department.

[18] Williamson, Steve; and Randall Wright. (1994) "Barter and Monetary Exchange under Private Information." *American Economic Review*, 84:104-23.

[19] Young, Peyton. (1993): "The Evolution of Conventions." *Econometrica,* 61:29-56.

[20] Zhou, Ruilin; and Edward J. Green. (2002) "Dynamic Monetary Equilibrium in Random Matching Economy " *Econometrica,* 70:929-969.