

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

May 16, 2008

Abstract

This paper shows that money as a fiat medium of exchange arises naturally in complex economies. An economy is complex if the probability of being able to barter is small.

We show this by having people learn the value of money in a simple search model. We then allow people to occasionally experiment with new strategies (or *mutate*). In this model if there are more goods than money in all trading periods then non-convertible fiat money is evolutionarily dominant (or *stochastically stable*) if the economy is complex enough.

Key words: Fiat Money, Learning, Stochastic Evolution.

JEL codes: E49, C73, D83

1 Introduction

How did people learn to exchange intrinsically worthless paper money for valuable goods? This is a problem that has troubled economists for a long time. The majority of the traditional literature has focused on explaining why money is Pareto superior to barter. One conclusion is that a primary benefit of money is in overcoming the famed double coincidence of wants—an insight first developed by Jevons [12]. In a barter economy one must both find someone who wants what one has and has what one wants, while in a monetary economy one only needs to find someone who has what one wants. This insight has been used in a spate of recent papers that analyze the transition from a barter economy to a monetary economy. Most of these papers are based on the elegant analytic search model in Kiyotaki and Wright [15]. For example Ritter [21] analyzes the incentives of the body issuing the fiat currency, showing when it can optimally introduce currency. However since all these papers use equilibrium analysis they must assume that people trust money, thus bypassing the core issue. This type of analysis can not explain how people learn to trust money.

To understand how society transitions between these viable methods of exchange (money and barter) we need to have a model of how people behave out of equilibrium. One such model that has recently been subjected to analysis in the microeconomics literature is the model of stochastic evolution—first developed by Kandori, Mailath, and Rob [13] and Young [30].

*We would like to thank Erdem Basci, without his inspiration this paper would not have been possible. We would also like to thank Randall Wright and Refet Gurkaynak for their feedback. Finally we would like to thank the participants at the 2006 Federal Reserve Bank of Cleveland Workshop on: Money, Banking and Payments; the 2005 North American Winter Meetings of the Econometric Society; the University of Maynooth and University College, Dublin for their valuable comments. Any remaining mistakes are the responsibility of the authors.

[†]Corresponding author, email hasker@bilkent.edu.tr

This paper adapts this methodology to the study of the evolution of money. Using money is an equilibrium in a dynamic game—one accepts money today so one can buy things tomorrow. Thus we need a well specified model of learning that can enable our agents to learn their optimal strategy. We use a modified version of a reinforcement learning algorithm developed by Marimon, McGrattan, and Sargent [20], and prove that agents using our modified algorithm are able to learn in a wide variety of environments. This contrasts with Lettau and Uhlig [19], who show that the basic algorithm can lead to the use of a sub-optimal strategy. In our model people always learn the optimal strategy from interaction and feedback.

In this paper we put agents using our reinforcement learning algorithm into the classic search model of money developed by Kiyotaki and Wright [15]. Very rarely they “mutate” and try out new strategies. If a critical mass of people start using fiat money, then the rest of the population slowly learns that fiat money is trustworthy. We then drive the probability of mutations to zero to concentrate our prediction. In this case our prediction is that this transition occurs surprisingly frequently. We call an economy “complex” when it is difficult to find someone to barter with. In complex economies, as long as the velocity of money is large enough fiat money is evolutionarily dominant (or *stochastically stable*).

There is a strain of the money literature that analyzes what occurs under various learning dynamics. Unfortunately these models depend on initial conditions. They can not tell us if money will arise, only what the initial conditions must be. In contrast our model is independent of initial conditions, thus can provide an unqualified insight into when and if money will arise. Since they depend on initial conditions these models could be considered models of “static evolution”. The seminal paper on learning is Marimon, McGrattan, and Sargent [20]; and others modify that model to improve on those results. For example Basci [2] includes learning from neighbors and shows people learn more quickly. Other papers use simpler learning algorithms and focus more on the evolutionary aspect of the analysis. Iwai [11] uses a “bootstrap” argument to explain when commodity and fiat money can arise. Wright [29] uses an evolutionary argument to show that barter and fiat money cannot co-exist as methods of exchange. Wright [28] shows that the commodity with the lowest storage cost will naturally become commodity money. Selgin [23] shows that even if fiat money is possible when society starts out using barter people will learn to use the good with the lowest storage cost as commodity money. These articles emphasize our key point. We have to look at models where agents occasionally “mutate” or conduct “grand experiments”. Only by analyzing this type of model can we hope to discover whether people will learn to trust money.

Our reinforcement learning model allows us to avoid two of the three key assumptions of the classic evolutionary model. That model—as developed by Kandori, Mailath, and Rob [13]—makes three behavioral assumptions: inertia, myopia, and mutations. Inertia means that not all agents change their strategy in every period, in our model this happens naturally since agents’ beliefs change slowly. Myopia means that one chooses the action that is optimal given the current distribution of the population, our learning model explicitly explains how beliefs are formed and with our explicit model one just chooses the optimal action given one’s beliefs. The only element of their model that we directly adopt is mutations. Mutations mean that very rarely people change their strategy at random, in our model very rarely people develop radical new guesses about the continuation values.

Since mutations are new to the money literature we will discuss this element of the model in detail. What are usually referred to as “mutations” in the literature on stochastic evolution can also be thought of as experiments. One should recognize that simple experiments are necessary for learning. In our model, for example, if someone does not occasionally experiment with trading he could believe that never trading is best. We handle this type of simple experiment by having people occasionally take the wrong action by

mistake.

In complicated environments one also needs to try out more complicated and long-lasting experiments to find the globally optimal strategy. Examples of this are easily observed in real life. One simple but unstructured example of this is an entrepreneur. An entrepreneur commits three to five years of his life to develop a new good or service, and in our model this would be a mutation. Clearly others do not share his belief that this good or service is valuable or they would already be providing it. A more concrete example is the strategy of allowing customers to return goods. Clearly the first impact of such a policy is that the store will lose revenue when goods are returned in poor condition. It is only over time, as individuals realize they can buy goods that they are not sure that they need, that the benefit arises. This has become standard policy in the United States, thus consumers in the United States may not be aware of this incentive, but in Turkey this policy is only now being adopted and the authors can state, based on personal experience, that the secondary effect makes this a profitable strategy. Clearly accepting money is a similar experiment. Initially each time one accepts money one will have a hard time using it to buy goods. Only over time, as more and more people decide to use money, does the benefit come. This more complicated type of experiment is a mutation.

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 heart of the paper mimics this three step presentation. We first find the steady state equilibria of the model in Section 3. Following this we show that 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 dominant) in Section 5.

Having finished the main argument of the paper we then discuss how well our model fits the history of money; analyze its implication for welfare; and consider several extensions of the model in Section 6. In Section 6.3 first we allow for a more general covariance between the preferences of individuals; second we allow for a flow value for money, and finally we consider a model where there is a cost to production. Section 7 concludes our paper.

2 Model

2.1 The Model of Trading

The model in Kiyotaki and Wright [15] 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 trade using 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 each of these exogenous events be $x < 1$ then barter occurs with probability x^2 ; monetary exchange 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 each agent is endowed with one unit of a good that he can not consume, thus he must trade. This good is neither durable nor divisible, thus it must be either traded for a unit of consumption good or wasted. Next period each agent will be endowed with a unit of consumption good again. This is a basic barter economy.

We will change the endowment of some agents to one unit of fiat money (a fraction $\mu = \frac{i}{I}$, $i \in \{3, 5, 7, \dots, I - 3\}$). This good is not consumable but it is durable, enabling trade across periods. It is also a substitute for barter, thus someone holding a unit of money will not have a unit of consumption good. Now we have a model where agents can trade either using barter or monetary exchange.

Agents 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 has what j wants, call this c_j ; or i does not have what j wants, call this c_{-j} ; or i has money, denoted m . Likewise agent j may 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 $\Omega = \{c_j, c_{-j}, m\} \times \{c_i, c_{-i}, m\}$. Given that i has met with j the exogenous probability that j is in each of her three states are μ , $(1 - \mu)x$, and $(1 - \mu)(1 - x)$ for states m , c_i , and c_{-i} respectively.

Each agent will observe the good that both he and his 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. To facilitate learning players will "experiment." This experimentation will take the form of taking the wrong action with an exogenous probability $\chi \in (0, \frac{1}{2})$.

If agent 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. The following period (which is discounted at a rate of $\rho > 0$) someone with money can trade for a unit of consumption good if she finds a willing partner. Notice that this means that the interaction is dynamic, the value of money today depends on what occurs next period.

In the model x is an exogenous parameter, representing the probability that person i wants the good j has, however—as Kiyotaki and Wright [15] suggests—we can interpret x as a measure of the complexity of the economy. We could generate x by assuming each agent produces one of M goods and each period needs to consume one of $M_c \leq M$ goods, then $x = \frac{M_c}{M}$. Then the more agents specialize in production the smaller x will be—in modern economies one can easily argue that x is nearly zero.

We would like to specifically mention the two simplifications in this model that make it the most divorced from reality. First of all there is no body that issues the fiat money and controls the quantity of money. This allows us to ignore the incentives of the body that is issuing the fiat money and the issue of seigniorage. For articles that address these issues one can refer to Ritter [21] and Berentsen [4]. Ritter's key result is that the body issuing the fiat money has to represent a large enough fraction of the population for there to be a fiat money equilibrium. Berentsen shows that if consumers discipline the money issuer then the issuer can be purely profit motivated.

Secondly neither money nor goods are divisible. If money is divisible it means that an important state variable is the distribution of money and this makes the model intractable for evolutionary analysis. All the papers that have overcome this problem have either used a continuum of agents (Shi [25] and Green and Zhou [9]) or a centralized market to balance the money stock (Lagos and Wright [18]). If goods are divisible then this is the Shi-Trejos-Wright model ([24] and [26]). These papers may be of interest for future research, but in this analysis we want to focus on the method of trade rather than the terms of trade.

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 economy will usually 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. It would be surprising if people did not use rules of thumb in such an environment. One model developed by Marimon, McGrattan, and Sargent [20] and Lettau and Uhlig [19] 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.

While this type of model has only recently been analyzed in economics we should mention that it has a long history in other literatures. One of the ground breaking works in this field was Chris Watkins' dissertation in psychology [27]. Computer scientists have long analyzed models in this class as well; see Barto, Bradtke, and Singh [1] for a review of this literature and some new results. Notice that the psychologists are seeking a reasonable model for human behavior while the computer scientists are seeking the most efficient learning algorithm in a complex environment. Despite these divergent goals they both have analyzed the type of model we consider here.

The model is a formulation of a simple feedback response mechanism. Each period one takes the action that currently seems to be optimal, whatever action one takes one re-estimates that action's value, and then the following period one repeats the algorithm. This simple algorithm is slightly complicated by the fact that money has a dynamic value. Thus when one re-estimates the action's value one has to also consider the current belief about the continuation payoffs.

In this model each person begins with an initial belief of the value (called a *strength*) of an action pair $\alpha \in A$ at the state $\omega \in \Omega$. We denote this strength $S_{\alpha\omega}^0$, 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.¹ Then he updates the strength of the state that actually occurs. Let $1 - \tau \in (0, 1)$ 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:

$$S_{\alpha\omega}^{t+1} = (1 - \tau) S_{\alpha\omega}^t + \tau \left(U_{\alpha\omega} + \frac{1}{1 + \rho} P_{\alpha\omega}^t S^t \right). \quad (1)$$

otherwise $S_{\alpha'\omega'}^{t+1} = S_{\alpha'\omega'}^t$. The row vector $P_{\alpha\omega}^t$ is the Markov transition probabilities in period t given $\{\alpha, \omega\}$. We construct $P_{\alpha\omega}^t$ by having each person observe the distribution of strategies last period but not who is holding money. People then deduce that all people who trade for money are holding money, and only if there are too many will the rest be holding a consumption good. Notice that when calculating the probability of a trade people are aware that players experiment but do not take into consideration the impact of their own strategy. One should assume that I is large enough that this last assumption does not matter. We will discuss our assumptions and alternative assumptions in section 4.

There are several differences between this model and the general model in Lettau and Uhlig [19]. These differences are primarily motivated by a desire to be certain agents can learn. We argue that this is a type of "long run rationality". Clearly in the short run agents will make mistakes because they have the wrong priors, however if they also are making mistakes in the long run this suggests that 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.

The first of the changes we make to the model is that like Basci [2] we have one strength for each action/state pair and thus our agents choose an optimal action at each state instead of optimal strategies. Notice that optimal strategies are made up of optimal actions at each state, thus in our model if S^{t-1} is correct and the distribution of strategies is not changing agents will choose the optimal strategy. If we instead had them choosing strategies then either we rule out a priori players using certain strategies or we have a model which is equivalent to the one here.² Since ruling out strategies a priori is anathema to unbiased

¹To calculate this agents need an estimate for $P^t(T|\omega)$, we will assume that it is $P^{t-1}(T|\omega)$. The value for the initial period will not affect analysis, but for historic consistency we can assume that $P^0(T|\omega)$ is generated assuming the population is in the barter equilibrium.

²Notice that the optimal strategy may be available and still not be chosen due to the other strategies allowed. For an example see Lettau and Uhlig [19].

analysis we favor the simpler model. Our assumption that they have a different strength depending on the action of their opponent is made for technical convenience.

The second difference is that we assume agents experiment—there is some fixed probability $\chi > 0$ that they choose the wrong action at every state. As Watkins [27] shows this enables learning because every state is reached with positive probability. Without this assumption an agent could reach suboptimal conclusions like never trading is optimal. All results in this paper will be for small χ .

Two final changes are dictated by the environment we are analyzing. In Marimon et al. [20] the update factor (τ) is decreasing over time. As Benveniste, Metivier, and Priouret [3] point out this is only sensible if the optimal strategy cannot change over time. In our model there are multiple equilibria and thus the optimal strategy can change. If $\tau \rightarrow 0$ this means that agents are ignoring new information, and thus the agents might be making worse and worse decisions as time goes by. To avoid this difficulty we assume that τ is constant.

However this last change makes our model unstable and must be compensated for by another change. In the standard model P^t is just what happened to agent i in period t . This is one reason that they assume that $\tau \rightarrow 0$, without this assumption the model may never converge. Thus we assume that P^t is based on the population distribution of strategies from last period. Our results would hold if P^t was based on old information but are certain our results fail if it is based only on personal experience or a partial sample of the true distribution.

We want to allow for very general priors on the strengths, all we require is that they are drawn from a compact and continuous support that includes all feasible payoffs, or that any $S^0 \in \left[-\kappa \frac{1+\rho}{\rho}, U \frac{1+\rho}{\rho}\right]^{|A||\Omega|}$ is possible.

2.3 The Model of Evolution

Adding one refinement to our model of learning yields a model of stochastic evolution. We will assume that with probability $\varepsilon > 0$ people *mutate*, or draw S^{t+1} from the initial distribution. Notice that this behavior is not a priori irrational. In any model with multiple equilibria if enough people switch their strategy at one time this will change the equilibrium. The only difference is that in our model people may be wrong. We can also justify this as checking to make sure if one’s beliefs are correct. Using our updating rule beliefs will only be correct in the long run, and so mutants could be players who think it worthwhile to check these beliefs. They could, for example, be responding to the fact that everyone they have interacted recently has not used the correct strategy (either by mistake or because they themselves are mutants).

Regardless of why one accepts these mutations, the impact of these mutations is to completely change the dynamics of the game. Because of this assumption society will not settle down to any particular state, instead it will constantly adopt new strategies. However as $\varepsilon \rightarrow 0$ society will spend most of its 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 its 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. if the distribution of strategies is currently at one distribution in the limit set, and everyone best responds to that distribution then the distribution of strategies

is at another. Clearly any Nash equilibrium is a limit set, but there might be others. In this interaction, however, there will not be. To be precise the limit sets in our model will be the steady state Nash equilibria of the underlying game. Most of the time society will be in one of these equilibria, and infrequently it will transition between them.

We focus our analysis on the case when ε is very small. As $\varepsilon \rightarrow 0$ 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.

In Section 5 we will show that the only transitions that matter in this paper are the transitions that takes the least mutations, these transitions determines the *radius*.

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

Intuitively the reader might find it easier to think of the radius as a “security level”. In this interaction if fewer than this fraction stop using a strategy in the limit set then this change 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 set aside the model of learning and evolution for a while, and focus on the steady state equilibria of the game. This is essential because most of the time society will be resting in one of these equilibria, with only mutants temporarily using a different strategy.

In order to carefully analyze learning and out of equilibrium behavior we need a simple method of representing every possible strategy. One method is to list the states at which this strategy trades. Then, for example, the *always trade* strategy is:

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

and the *never trade* strategy is:

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

Notice that NT is a steady state of the learning algorithm if $\chi = 0$. If one starts out with the beliefs that not trading is optimal then one will never have any experience that will change one’s mind—even if everyone else is trading. Of course if we included learning from neighbors like Basci [2] if some trade then eventually all will, but it would still be an equilibrium for everyone not to trade. This is an example of why experimentation is important, since we assume that people experiment ($\chi > 0$) this strategy will be strictly dominated.

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. ■

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 to specify a strategy

³For a real number x , $\lceil x \rceil$ is the smallest integer which is larger than x .

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:

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

B is a barter strategy, if a person is using this strategy and holding money then they will take anything to get rid of it. F is the monetary exchange strategy, in this strategy one will trade anything in order to hold money. Notice that everyone barter if it is feasible, one always is willing to trade for c_i .

If agents make mistakes too frequently, or are too impatient, then trading for c_{-i} when holding money might never 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 steady state 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 with 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). We find that this can be done very elegantly in terms of the expected per-period (or *flow*) utility of holding either a consumption good or money. In effect the two values are nothing more than a weighted average of these two flow utilities. Establishing the proper notation and some simple algebra makes this clear.

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

$$\begin{aligned} \Pi(m) &= \frac{1+\rho}{\rho} \frac{1}{\Delta} [(\rho + P(m|c))U(m) + P(c|m)U(c)] \\ \Pi(c) &= \frac{1+\rho}{\rho} \frac{1}{\Delta} [P(m|c)U(m) + (\rho + P(c|m))U(c)] \end{aligned}$$

where $U(z)$ is the expected flow utility of a person holding good $z \in \{c, m\}$; $P(z'|z)$ is the probability of starting the period holding good z and ending the period holding $z' = \{c, m\} \setminus z$ and $\Delta = \rho + P(m|c) + P(c|m)$.

Proof. Let $P_z(z')$ be the probability of trading for good $z \in \{c_i, c_{-i}, m\}$ given one is holding $z' \in \{c, m\}$. Then the flow value of these states are:

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

⁴If an agent is indifferent we assume that he will trade.

When one realizes $P(m|c) = P_m(c)$ and $P(c|m) = P_{c_i}(m) + P_{c_{-i}}(m)$, $U(c) = P_{c_i}(c)U - P_{c_{-i}}(c)\kappa$, and $U(m) = P_{c_i}(m)U - P_{c_{-i}}(m)\kappa$ then one can solve for the above values easily. ■

With this result we can easily find the pure population equilibria.

Lemma 3 *F and B are the pure population equilibria unless agents are impatient or experiment a lot. If $(1 - \mu)x^2 \frac{U}{\kappa} < \rho$ or $\chi > \chi^*$ then \tilde{B} is an equilibrium instead of B.*

Proof. See the appendix. ■

In the rest of the paper we will assume that B is the pure population barter equilibrium without loss of generality.

Quite interestingly there is a continuum of mixed population equilibria, and in none of them do agents 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 paying κ to hold a consumption good is worth it. The essence of any mixed population equilibrium is that the value of holding money or a consumption good is the same; thus it is not worth paying κ to hold a consumption good. There is a continuum of mixed population equilibria because in such an equilibrium agents are indifferent between using the strategy F and \tilde{F} :

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

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 agents 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, but fortunately it is not important for our analysis. The radius of the barter equilibrium and the fiat money equilibrium will be determined by the extreme distributions, where all agents 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 $f \in \{F, \tilde{F}\}$ and $\pi_c(f)$ be the equivalent portion holding the consumption good.

For notational convenience we write the equilibrium using two *experimentation functions*. Due to experimentation the equilibria will be slightly modified from a model without experimentation, but these modifications will be small and it is most convenient to summarize their affect as two types of functions, $L(\chi)$ and $O(\chi)$. The $O(\chi)$ is the order function from statistical analysis, as $\chi \rightarrow 0$ $O(\chi) \rightarrow \chi K$ for some positive constant K . The $L(\chi)$ functions are equivalent functions that converge to one instead of zero.

Lemma 4 *The extreme mixed population equilibria are*

$$\begin{aligned} \pi_c(f) &= xL_c(\chi|f) - O_c(\chi|f) \\ \pi_m(f) &= L_m(\chi|f) \end{aligned}$$

where if $\delta(F) = 1$ and $\delta(\tilde{F}) = 0$ then

$$\begin{aligned} L_c(\chi|f) &= \frac{(1 - \chi)^2 + \chi^2 \frac{\kappa}{U}}{(1 - 2\chi)(1 - \chi - \chi \frac{\kappa}{U} \frac{1 - x}{x} \delta(f))} > 1 \\ O_c(\chi|f) &= \chi \frac{\kappa}{U} \frac{1}{1 - \chi - \chi \frac{\kappa}{U} \frac{1 - x}{x} \delta(f)} \\ L_m(\chi|f) &= \frac{1}{1 + O_{io}(\chi|f) \frac{1 - \pi_c(f)}{\pi_c(f)}} < 1 \end{aligned}$$

where

$$O_{io}(\chi|f) = \chi \frac{x(1-2\chi) + \chi}{x(1-x) + (x-\chi)^2 + \delta(f)\chi(1-2\chi)(1-x)}.$$

Note that $\lim_{\chi \rightarrow 0} L_c(\chi|f) = \lim_{\chi \rightarrow 0} L_m(\chi|f) = 1$, $\lim_{\chi \rightarrow 0} O_c(\chi|f) = \lim_{\chi \rightarrow 0} O_{io}(\chi|f) = 0$. Furthermore $\pi_c(F) > \pi_c(\tilde{F})$.

Proof. See the appendix. ■

The equilibrium condition is easiest to understand when $\chi = 0$. In this case it is $\pi_c(f) = x$. The intuition behind this condition is that someone who is choosing between trading for a unit of money or not must be indifferent between holding money and a consumption good next period. Since when one is holding a consumption good the probability someone wants what one has is x , the probability they want what one has when holding money better be the same.

Notice several interesting facts about this equilibrium. First, as Wright [29] shows the mixed population equilibrium requires more people to be using money than a mixed strategy equilibrium. In a mixed strategy equilibrium anyone will trade for money thus $\pi_m(f) \approx \pi_c(f)$, while in a mixed population equilibria essentially everyone who is holding money will be using a monetary exchange strategy. Secondly notice these equilibria are not stable, if in any period less than $\pi_c(f)$ money traders are in the consumption state a barter strategy is a strict best response. Finally notice that the equilibrium condition depends only on $\pi_c(f)$. From now on we will assume that $\pi_m(f) = 1$, and this will not affect our analysis as long as $\pi_c(f)$ is at it's critical value.

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”: agents can learn. With any proper learning algorithm this result should hold. The second proposition is not a result we required a priori, this proposition shows that the population can learn simultaneously. The results of both of these propositions are very general since our method is to show that our algorithm collapses to a Gauss-Seidel learning algorithm.

Before launching on the proofs in this section we would like to pause and explain exactly why we have modeled the transition matrix as we have. We assume that players observe the distribution of strategies and then deduce that everyone who will trade for money is holding money—unless there are too many of these people. Notice we could, without loss of generality, assume that people only observe the fraction of people who will trade at each state— ω . This is a more appealing description because now we are only requiring them to know observable information. If we used this more primitive assumption we would then have to have players deduce the distribution of strategies from this information, but other than being technically complicated this could be done in a straightforward way.

Alternatively we could assume that they know more information—the distribution of strategies dependent on whether people are holding money, however our assumption is less information intensive. Implicitly we can allow them to be using summarized information gathered from several periods. As well if they knew the distribution dependent on whether people are holding money then mass experiments (everyone choosing the wrong action) could cause people to change what strategy they think is optimal. Our assumption requires that the fundamentals of the model—the strategies of other players—change before people change their strategy. This alternative would result in more frequent transitions to fiat money. Another alternative is to have them believe the strategies and the probability of holding money are independent, but this will only be

correct if the entire population is using the same strategy. A final alternative is to have them calculate the equivalent of $\pi_m(f)$ given the current distribution, but this would require complicated calculations and the difference between this and our assumption will be small for small χ .

The convergence of our learning algorithm is immediate since $\chi > 0$. 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. ■

For the reader unfamiliar with the Gauss-Seidel method this result may appear both surprising and complex. However it is actually rather simple. To understand this realize first of all that we can rewrite the update rule (equation 1) as

$$S_{a\omega}^{t+1} = \tau U_{a\omega} + (1 - \tau) S_{a\omega}^t + \tau \frac{1}{1 + \rho} P_{a\omega}^t S^t \quad (2)$$

and the maximum weight on any element of S^t is $(1 - \tau) + \tau \frac{1}{1 + \rho} < 1$. Thus if the strategy is fixed this is a contraction mapping for a given $\{\alpha, \omega\}$, and since $\chi > 0$ we can update the strength of every $\{\alpha, \omega\}$. This means that if a person does not change his strategy his estimates of the true continuation values will get close to the actual values in finite time. The only tricky part of the Gauss-Seidel method is to verify that this result will not be changed by the person updating their strategy. Since the objective function is concave after an initial period all changes will be towards the correct strategy and it will not confuse this result. Thus in finite time this person will learn the optimal strategy, and anything that can be learned in finite time will be learned sooner or later with probability one.

As we said above there is no guarantee that one person being able to learn implies that society is able to learn. However, with Gauss-Seidel algorithms Kushner [17] has shown that in general they can learn pure population equilibria. The independent proof below also tells us exactly when society may converge to a certain equilibrium, and this information is critical for finding the radii of the equilibria.

Proposition 2 *Assume that I is large enough that the best response does not change when $i \in \{I - 3, I - 1\}$ people are using the same strategy. Then society will converge to an equilibrium almost surely.*

Furthermore, if the best response is independent of any two player's strategies it will converge to the equilibrium strategy that is currently a best response with positive probability.

Proof. See the appendix. ■

5 Evolution

The way evolution occurs in our model is that every period some individuals decide to try out a “grand experiment”—they mutate to some new strategy. Usually, as time goes on, these individuals learn their experiment was not a good idea and they revert to the current accepted practice. Occasionally, however, a large enough spike of mutations will occur and a revolution will happen, overturning the accepted practice and establishing a new standard.

How large are these spikes? This is the key question on which our results will rest. If the spike that it takes to exit the barter equilibrium is smaller than the spike it takes to leave the monetary exchange equilibrium then monetary exchange will be stochastically stable—it will be the evolutionarily dominant strategy in the long run.

On the face of it this is a complex question because there are many different spikes that could cause us to leave—say—the barter equilibrium. However if one spike takes more mutations than another it will occur infinitely less frequently as $\varepsilon \rightarrow 0$. To be precise say that spike a takes I_a people and spike b takes I_b where $I_b > I_a$ then the relative likelihood of a spike of type b to a spike of type a is $\frac{\varepsilon^{I_b}}{\varepsilon^{I_a}} = \varepsilon^{I_b - I_a}$, and as $\varepsilon \rightarrow 0$ $\varepsilon^{I_b - I_a} \rightarrow 0$. Furthermore *nothing else about a spike matters*. If, for example, spike a requires mutations that happen with probability $p_a > 0$ and spike b requires mutations that happen with probability $p_b > 0$ the relative likelihood is $\frac{p_b^{I_b} \varepsilon^{I_b}}{p_a^{I_a} \varepsilon^{I_a}} = \frac{p_b^{I_b}}{p_a^{I_a}} \varepsilon^{I_b - I_a}$ and as one can see the ratio $\frac{p_b^{I_b}}{p_a^{I_a}}$ doesn't affect the previous statement. The elegance of this method is that all we need to find is the smallest number of mutations that it will take to leave each of our equilibria.

Notice that, at least to a degree, one thing that doesn't matter is the order in which mutations happen. We could, for example, have a small number of people mutate. Then people who interact with them could mutate, and this could cause a chain reaction which results in a transition. All that matters is that this happens quickly relative to the speed that people learn.

One should recognize that even the initial mutants are not acting irrationally in this model. In the analytic model of Kiyotaki and Wright [15] there are multiple equilibria. In any such model it is possible for a large group of players to switch strategy at any time, resulting in a new equilibrium. The only difference between this model and that one is that we allow for the possibility that people may be wrong about when the transition will occur. Furthermore this comment is doubly strong for people who have interacted with a lot of mutants lately. All these people have to decide is that their beliefs are out of date and that they should pay more attention to local information than global information. There are many reasons that people might mutate, and given that we analyze the model when this probability is extremely small perhaps the question should be how mutations can be completely ruled out. In general it is when these mutations almost never happen that the results are the clearest, because it is then that only the size of the spike matters.

In our model the analysis will be more complicated than simply finding the smallest necessary spike, because we have to keep track of the strengths people associate with each strategy. Thus we have to prove that for any smaller spike society will not mistakenly converge to the other equilibrium. This will be true as long as people do not experiment too much, since as $\varepsilon \rightarrow 0$ society will be in a given equilibrium for a long time, thus their strengths will nearly be the true values. After this occurs only the number of mutations will matter.

Lemma 5 $\exists \bar{\chi} > 0$ such that if $\chi \leq \bar{\chi}$ then $r^*(F) = (1 - \pi_c(F))(1 - \mu)$ and $r^*(B) = (1 - \mu)\pi_c(\tilde{F}) + \mu$, and if π^* is a mixed population equilibrium $r^*(\pi^*) \leq \lim_{I \rightarrow \infty} \frac{2}{I}$.

Proof. See the appendix. ■

We would like to mention that in this lemma our decisions about how beliefs are formed have stacked the deck against us. We decided that agents believe that all people holding money are people who will trade for money, thus the fraction that are holding a consumption good and will trade for money is as small as possible. In fact if barter is the current mode of exchange then only mutants will be using fiat money, and they will end up holding money if enough of them aren't around. This makes $r^*(B)$ as large as possible, and this is bad for money.

How is a large $r^*(B)$ bad for monetary exchange? To answer that we need to use Lemma 5 to find out what strategy is most likely in the long run. The first conclusion we can reach is that we can essentially ignore mixed population equilibria. In large populations their radius is trivially small and society will never be in one of these equilibria for a meaningful length of time.

Now let us focus only on the barter and monetary exchange equilibria, specifically let us focus on finding the probability of F in the long run, denoted $P(F)$. One key result from Freidlin and Wentzell [8] is that if $P(E|E')$ is the probability of transitioning to $E \in \{B, F\}$ from $E' = \{B, F\} \setminus E$ then:

$$P(F) = \frac{P(F|B)}{P(F|B) + P(B|F)}$$

and as explained above as $\varepsilon \rightarrow 0$ $P(F|B) \rightarrow \varepsilon^{\lceil r^*(B)I \rceil}$. For simplicity write $\varepsilon_s = \varepsilon^I$ then $P(F|B) \rightarrow \varepsilon_s^{r^*(B)}$ and in large economies:

$$P(F) = \frac{\varepsilon_s^{r^*(B)}}{\varepsilon_s^{r^*(B)} + \varepsilon_s^{r^*(F)}} = \frac{1}{1 + \varepsilon_s^{r^*(F) - r^*(B)}} .$$

Thus if $r^*(F) > r^*(B)$ as $\varepsilon_s \rightarrow 0$ $P(F) \rightarrow 1$.

Theorem 1 proves that the above intuition is correct and characterizes when $r^*(F) > r^*(B)$. The most interesting case is when χ is nearly zero and the society is complex— x is nearly zero—or it is hard to find someone to trade with. In this case the only important parameter is μ , and if $\mu < \frac{1}{2}$ then monetary exchange will be evolutionarily dominant or stochastically stable.

Theorem 1 *If $\chi \leq \bar{\chi}$ and there is less money than goods ($\mu < \frac{1}{2}$) then in complex societies monetary exchange is stochastically stable when χ is nearly zero. More generally this is true when:*

$$\frac{1 - \mu}{\mu} > \frac{1}{1 - 2x\bar{L}_c(\chi) + \bar{O}_c(\chi)} \quad (3)$$

where $\bar{L}_c(\chi) = \frac{L_c(\chi|F) + L_c(\chi|\bar{F})}{2}$ and $\bar{O}_c(\chi) = O_c(\chi|F) + O_c(\chi|\bar{F})$.

Proof. See the appendix. ■

Notice that agents using the monetary exchange strategy do barter if the opportunity arises. Like ourselves, if there is a situation where barter can be used we won't ignore this possibility. However, again like in the real economy, most trades will use money if x is small (the economy is complex).

Interpreting the condition in Theorem 1 is somewhat complicated because $\frac{\mu}{1-\mu}$ is the ratio of money to goods available for trading in a *trading period*. In order to interpret this condition we must first determine how long a trading period is. There are two ways to estimate this. First in our model we assume all goods are worthless at the end of the trading period. However we made this assumption merely to be sure that no one could use a good as money, thus this is not the right method to estimate the length of a trading period. Second we could base it on the average time between trades for an average consumer. This is the correct method, and this clearly suggests a trading period is less than a day long in modern economies, and must be less than a week. A further difficulty in estimating μ is that it depends on the stock of goods, not the amount traded. In each period only $x^2(1-\mu) + x\mu$ of the available goods will be traded.

It is clear from this condition that if $x \geq \frac{1}{2}$ or the market is flooded with money then money will not be stochastically stable. The latter requirement is appealing because it suggests that during periods of hyperinflation a currency could collapse. As well we can see the more complex is the economy (the smaller is x) then money is more likely to be stochastically stable. Below we show how we can relate our condition to the velocity of money.

5.1 The relationship between the evolutionary stability of money and the velocity of money

In each trading period the quantity traded will be $(1 - \mu)^2 x^2 + (1 - \mu) \mu x$. Thus in T periods the total value of goods traded will be $(1 - \mu) T x ((1 - \mu) x + \mu)$ and the velocity of money is

$$V(T) = \frac{(1 - \mu)}{\mu} T x ((1 - \mu) x + \mu) .$$

Notice that this depends on the time period—like it does in reality. Empirically one can estimate the velocity per month, per quarter, per year, or even per decade and will develop different estimates. However in theoretic models it is standard that $V(T) \geq 1$, and in modern economies $V(T)$ is generally large, so we can rewrite the condition in Theorem 1 under this assumption.

Unfortunately this introduces a new unknown variable, T . One solution to this problem is to write the condition in terms of $\frac{V(T)}{T}$, since in our model this will be independent of T . Another solution is to make some reasonable assumption about $T(x)$. Clearly this assumption should be that $T(x)$ is decreasing in x . In simple economies people rarely engage in trade, almost all goods they consume come from their farm or from hunting and gathering. In contrast people who live in cities have to trade for nearly every good they consume, trading several times a day. Thus a reasonable approximation is $T(x) = \gamma \frac{1}{x}$, then

$$V(\gamma) = \frac{(1 - \mu)}{\mu} \gamma ((1 - \mu) x + \mu) .$$

Lemma 6 *If V is the velocity of money then money will be stochastically stable if*

$$x \leq \frac{\mu + \frac{V(T)}{T} 2\bar{L}_c(\chi)}{2(1 - \mu)} \left(\sqrt{1 + \frac{\frac{V(T)}{T} 4(1 - \mu)(1 + \bar{O}_c(\chi))}{\left(\mu + \frac{V(T)}{T} 2\bar{L}_c(\chi)\right)^2}} - 1 \right) . \quad (4)$$

if $T(\gamma) = \gamma \frac{1}{x}$ then the condition is:

$$x \leq \frac{V(\gamma)(1 + \bar{O}_c(\chi)) - \gamma\mu}{2V(\gamma)\bar{L}_c(\chi) + \gamma(1 - \mu)} \quad (5)$$

Condition 4 only requires that given $V(T)$ x is small enough, or the economy is sufficiently complex. Condition 5 also requires that $V(\gamma) > \gamma\mu$, thus we can be certain that money will be stochastically stable if $V(\gamma)$ is large relative to μ and the economy is complex.

6 Discussion and Extensions

6.1 The Coincidence of the Model and the History of Money

We should mention that as the first paper to merge to disparate literatures—stochastic evolution and money—we can not hope to explain the history of money. Any such attempt would be ridiculous and outside the scope of this analysis. However we do feel that our model at least matches a significant moment of the history of money, and would like to clarify this point. If one considers our model as one of "barter" against "any type of money" then our model predicts there will be long periods where society will use money, followed by brief breakdowns of the medium of exchange and the rise of some (possibly new) method of monetary exchange.

As an example of how far our paper is from the true history of money we note that there has never been a direct transition from a barter economy to a fiat money economy. As Dowd [7] points out fiat money 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.

Why is commodity money an intermediate step between fiat money and barter? This is clearly a subject for future analysis but perhaps we can gain some insight by asking a different question here: in this model are we analyzing true fiat money? One of the critical differences between fiat money and commodity money is trust, and trust in two dimensions. First can I trust others to accept the money? Our model focuses on this dimension. But second, can I trust the issuers of the fiat money to keep the money stock constant? In almost all models of money if there is too much money issued it becomes worthless. While this is captured in our model by money becoming stochastically unstable when μ is too large we really have bypassed this issue by assumption.

Thus, perhaps, it is better to think of our money as either fiat or commodity money (we consider this alternative further in section 6.3) or to think of our model as one of “barter” versus “any type of money.” By this criterion we believe our model matches the history of money to a first order of approximation. To investigate this further let us look at the relative likelihood of the two equilibria:

$$\frac{P(F)}{P(B)} = \varepsilon_s^{r^*(B) - r^*(F)} .$$

As $\varepsilon_s \rightarrow 0$ if monetary exchange is stochastically stable then this diverges to infinity. Now we agree with Sandholm [22] that the model is not logically consistent if $\varepsilon = 0$. In our conclusions we drive ε to zero to develop precise results, but we really want to think about a situation where mutations or grand experiments are uncommon but not unheard of. This means that $\varepsilon_s = \varepsilon^I$ is small but strictly positive.

If there is a positive probability of mutations then our model predicts that we will see long periods in which monetary exchange will be used, interspersed with brief episodes where monetary exchange breaks down and barter is the only type of trading that takes place, much like the empirical history of money. Of course in the empirical history usually the government adapts a bad policy before the currency breaks down, but this model suggests that the bad policy may be the result of bad economic conditions, or that the fundamental cause is the inherent instability of money.

If anything, one problem in our model is how slowly the transition takes place. Usually episodes where currency fails are followed by the government reestablishing a money standard fairly quickly. To respond to this point note that in our model the strategy of a large trader (like the government) could have a large impact on the evolutionary process. If such a trader refused to barter this could quickly affect the expectations of other traders, and in our model this would cause people to begin to trust money quickly.

Notice that if $\varepsilon_s = \varepsilon^I$ then ε_s will be decreasing as the population grows. This would suggest that monetary exchange will be more stable in larger economies. Similarly the radius of the monetary exchange strategy is decreasing in μ , and the radius of barter is increasing in μ . Thus the transition to monetary exchange will be quicker (and money more stable) if there is a small stock of money. The intuition behind this insight is that the more money there is, the harder it will be to get a stable group of people that can trade money among themselves. This suggests that a government that wants to introduce (or reintroduce) fiat money should first introduce a small stock of money, and then, once people start using the currency, should introduce a larger stock of money. However the short run benefit of a small stock of money should be balanced against the long run welfare of society, and this is the issue we shall address next.

6.2 Long Run Welfare and the Optimal Stock of Money

When the probability of mutations is positive the stock of money has a new welfare impact. In Kiyotaki and Wright [15] there is a benefit to increasing money because it increases the number of trades ($x > x^2$) but there is a cost because it decreases the quantity of consumption goods ($1 - \mu$). In an evolutionary model a second cost is added, because increasing the stock of money also decreases $r^*(F) - r^*(B)$, thus decreasing the fraction of time spent in the monetary exchange equilibrium. While this impact will be small if the probability of mutations (ε) is reasonably small and the population (I) is reasonably large, we would like to briefly discuss it.

In the standard model Kiyotaki and Wright [15] define a priori welfare in the equilibrium E ($E \in \{F, B\}$) as:

$$W(E) = \mu \Pi(m) + (1 - \mu) \Pi(c) .$$

After simplification these are:

$$\begin{aligned} W(B) &= \frac{1 + \rho}{\rho} (1 - \mu) U(c) \\ W(F) &= \frac{1 + \rho}{\rho} [\mu U(m) + (1 - \mu) U(c)] \end{aligned}$$

where $U(z)$ is the expected flow utility of a person holding good $z \in \{c, m\}$, explicitly $U(c) = (1 - \mu)x^2U$ and $U(m) = (1 - \mu)xU$. It is easy to immediately see that using money Pareto dominates barter as long as $U(m) > 0$. Thus the optimal money stock maximizes $W(F)$, and the first order condition is:

$$1 = x \frac{2 - 2\mu}{1 - 2\mu}$$

and the optimum is when $\frac{1 - \mu}{\mu} = \frac{1}{1 - 2x}$. However this is exactly the stock of money where the monetary exchange equilibrium is not stochastically stable.

To explain this more clearly, if this level of money stock is selected then society will only spend half of it's time in the Pareto efficient monetary exchange equilibrium. When we choose the money stock we assumed that society would always use a monetary exchange equilibrium, thus this is no longer optimal.

This suggests we need to construct a welfare function that takes into consideration the stability of money. On the face of it this is intractable because we need to calculate welfare during transitional stages—when people are using different strategies. However as the probability of mutations becomes very small the amount of time spent in transitional phases becomes minuscule, thus it is more important to evaluate welfare in the two pure population equilibria. In the same spirit we will assume that the probability of experiments (χ) is zero. Learning will still occur as if there were experiments but we will ignore the impact of experiments on welfare.

The most instructive welfare function to calculate will be the *long run* welfare function. This welfare function is equivalent to a government maximizing welfare when it does not know the current equilibrium—thus the probability of the monetary exchange equilibrium is $P(F)$. This is:

$$W = \frac{1 + \rho}{\rho} [P(F) \mu U(m) + (1 - \mu) U(c)]$$

and we can write the first order condition as:

$$\frac{\partial P(F)}{\partial \mu} \mu \frac{1 - \mu}{1 - 2\mu} + P(F) = x \frac{2 - 2\mu}{1 - 2\mu} .$$

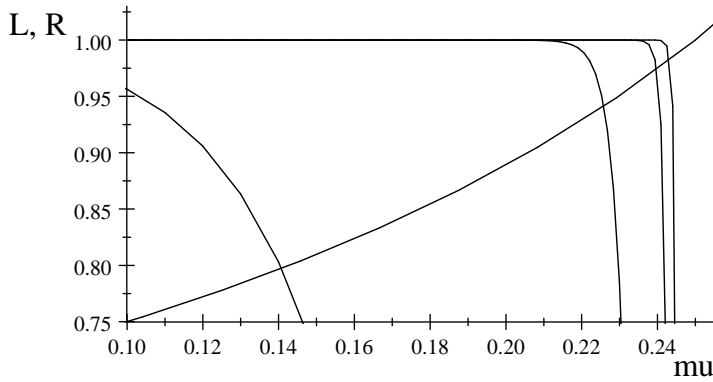
Since $\frac{\partial P(F)}{\partial \mu} < 0$, $\frac{\partial P(F)}{\partial \mu} \mu \frac{1-\mu}{1-2\mu} + P(F) < 1$ and this implies a lower stock of money. However it will not be much lower for reasonable values of ε_s . To understand why we can rewrite this condition as:

$$(1 - P(F)) P(F) \left((1-x) \mu \frac{2-2\mu}{1-2\mu} \ln \varepsilon_s \right) + P(F) = x \frac{2-2\mu}{1-2\mu}$$

and $P(F)$ as:

$$P(F) = \frac{1}{1 + \varepsilon_s^{(1-2\mu)(1-x\frac{2-2\mu}{1-2\mu})}}.$$

If $1 - x \frac{2-2\mu}{1-2\mu} > 0$ (or $\frac{\mu}{1-\mu} < \frac{1}{1-2x}$) and the value is large relative to $\varepsilon_s^{(1-2\mu)}$ then $\varepsilon_s^{(1-2\mu)(1-x\frac{2-2\mu}{1-2\mu})} \sim 0$ and $P(F) \sim 1$, thus the entire left hand side is basically $P(F)$ and close to 1. Only when $\frac{1-2x}{2-2x} - \mu$ is nearly zero does $P(F)$ drop significantly below 1 and then the left hand side will decrease precipitously. At this point it will intersect with $x \frac{2-2\mu}{1-2\mu}$, and we will have our optimal money stock. While $\frac{\mu}{1-\mu} < \frac{1}{1-2x}$ will be optimal the difference will not be large if ε_s is reasonably small. In the graph below we illustrate a case where the effect is strongest, when x and ε_s are both large.



Graph of $L(\mu) = \frac{\partial P(F)}{\partial \mu} \mu \frac{1-\mu}{1-2\mu} + P(F)$ and $R(\mu) = 2x \frac{1-\mu}{1-2\mu}$ for $x = \frac{1}{3}$ and $\varepsilon = 10^{-k}$, $k \in \{10, 100, 300, 500\}$

$R(\mu)$ is monotonically increasing, and $L(\mu)$ is monotonically decreasing, and as $\varepsilon_s \rightarrow 0$ the graph of $L(\mu)$ shifts to the right. With the $L(\mu)$ on the far left, when $\varepsilon_s = 10^{-10}$, the optimal money stock is about half its unconstrained value, but this is much too high even if the population is as small as 100, even $\varepsilon_s = 10^{-100}$ implies a person mutates an average of once per ten periods. If $\varepsilon_s = 10^{-1000}$ the optimal money stock is around 24%, or 96% of its unconstrained optimum.

One can do the same analysis conditional on the current equilibrium. The welfare function in both equilibria looks similar to the above:

$$W(E) = \frac{1+\rho}{\rho} [\beta(F|E, \rho) \mu U(m) + (1-\mu) U(c)]$$

where $\beta(F|F, \rho) = \frac{\rho + \varepsilon_s^{r^*(B)}}{\rho + \varepsilon_s^{r^*(F)} + \varepsilon_s^{r^*(B)}}$ and $\beta(F|B, \rho) = \frac{\varepsilon_s^{r^*(B)}}{\rho + \varepsilon_s^{r^*(F)} + \varepsilon_s^{r^*(B)}}$. Notice that if $\rho = 0$ then $\beta(F|F, \rho) = \beta(F|B, \rho) = P(F)$, but that since $\rho > 0$ as $\varepsilon_s \rightarrow 0$ $\beta(F|F, \rho) \rightarrow 1$, and $\beta(F|B, \rho) \rightarrow 0$. This suggests that when people are bartering the optimal money stock will be very small, and society will transition relatively quickly to using fiat money. When society is using fiat money the money stock will be increased, increasing the chance of switching back to barter. However notice that having an endogenous money stock is not something our model is designed to address, a proper analysis of this prediction would require a richer model.

6.3 A Generalized Model

In this section we wish to present a generalized model of our interaction. We will generalize it in three different ways. First is to allow for a more general covariance between the preferences of individuals. Second to better capture the possibility that our money is a commodity we will allow for a possible flow value for money. Third we will consider a model where, instead of κ being a transaction cost, κ is a production cost. One can consider any of these models in isolation by setting other parameters to their default values.

6.3.1 Generalizations

A generalized probability of trades: To generalize the probability of trades we let the probability of a single coincidence of wants be x and the probability of a double coincidence of wants be xy . This results in a much larger set of preferences that can be covered, many of which represent reasonable models. Our base model is the *independent preferences* model, the preferences of i and j are independent so $y = x$.

One extreme of this class of models is found in Kiyotaki and Wright [14], which we refer to as the *consumption chain* model. In this model if there are $M > 2$ types of goods then a person who produces good m must consume good $(m - 1) \bmod M$. One can imagine that the goods are durable and good $m - 1$ is a necessary input for the production of good m .⁵ In this model $y = 0$. This is a better description of the modern economy than the independent preferences model. For example, a simplified model would have consumers buying from retail stores. Retail stores would consume the goods of producers, and producers would consume the goods of raw material producers.

At the other extreme is the *village preferences* model. In this model agents usually want goods from their own “village” and if they do then they always know who to trade with. Otherwise they have to wait for a wandering trader, who is equally likely to have any of the M goods in the economy. If γ is the probability an agent wants a good from her village then $x = (1 - \gamma) \frac{1}{M}$ and $xy = \gamma + (1 - \gamma) \frac{1}{M^2}$ so $y = \frac{\gamma}{1 - \gamma} M + \frac{1}{M}$. Notice that if $\gamma > \frac{1}{M+1}$ then $y > 1$. Notice that this is also a crude model of partially directed search (Corbae, Temzilides and Wright [6]). In this interpretation γ is the probability that you will find the right person to barter with.

A flow value for money: As we mentioned in our model money could really be either fiat or commodity money. On the side of those who would argue for commodity money is that the quantity of money is constant. Without an issuing body and a varying money supply it is easier to think of our money as a commodity in fixed supply. A counter argument is that in a growing economy we assume the stock of money is a fixed ratio of total output, and this requires a gradual introduction of more currency. Fundamentally, within our model the difference is really one of semantics.

In this alternative model we will capture a key difference between the two interpretations. Generally someone holding commodity money will get some benefit. For example it might be a random probability that he can use it for consumption—with the special condition that they will be holding money after consuming money. We can represent this as a flow benefit from holding money of ϕ . This also has value within a model of fiat money. If we had a banking sector then perhaps $\phi > 0$ could represent a real rate of return. Alternatively $\phi < 0$ could represent an inflation cost or $\phi > 0$ a deflation benefit (as suggested by the Friedman Rule).

⁵Note that if consumption goods were formally durable then we would be considering a model with both commodity and fiat money, which is beyond the scope of this analysis.

However within the model interpreting ϕ causes some difficulties. If $\phi < 0$ then this means that the stock of money is growing at a faster rate than the economy, which in our analysis would mean money was becoming less and less stochastically stable. If $\phi > 0$ then there is either deflation or a banking system with associated borrowing and lending. Within our model we do not allow for inflation, deflation, borrowing or lending, thus this can only be considered a reduced form analysis.

Including a Cost of Production: In our base model there is a possibility of confusion between the two benefits of money. First one can trade it without a transaction cost, second it can be a store of value. A simple change can remove this by making the transaction cost from using money and bartering the same. This change also removes the arbitrary assumption of the base model that production is costless. In a general model clearly this should not be true, thus we simply treat κ as a cost of production. In this alternative model we assume that $U > \kappa$ to avoid cases where autarky is an equilibrium.

6.3.2 Equilibrium and Stochastic Stability in the General Model

These changes in total do produce a significantly different condition for the mixed population equilibrium, and thus potentially can change our results. First of all the flow values of holding money and a consumption good are now:

$$\begin{aligned} U(m) &= P_{c_i}(m)U - P_{c_{-i}}(m)\kappa + \phi \\ U(c) &= P_{c_i}(c)U - (P_{c_i}(c) + P_{c_{-i}}(c) + P_m(c))\kappa. \end{aligned}$$

Where $P_{z'}(z)$ is the probability of trading for a unit of good $z' \in \{c_i, c_{-i}, m\}$ when holding a unit of $z \in \{c, m\}$.

The condition which we use to find the mixed population equilibrium has also changed. Like before it basically makes someone holding a consumption good indifferent between trading for money and not trading. However now the value of holding money has to compensate them for the production cost. The mixed population equilibrium now is characterized by:

$$\begin{aligned} \frac{1}{1+\rho}\Pi(m) - \kappa &= \frac{1}{1+\rho}\Pi(c) \\ U(m) - \Delta\frac{(1+\rho)^2}{\rho}\kappa &= U(c). \end{aligned}$$

Thus the mixed population equilibrium is:

$$\begin{aligned} \pi_c(f) &= L_c(\chi|f)y - L_{c1}(\chi|f)\frac{\phi - \rho\kappa}{(1-\mu)x(U-\kappa)} - O_c(\chi|f) \\ \pi_m(f) &= L_m(\chi|f) \end{aligned} \tag{6}$$

where:

$$\begin{aligned} L_c(\chi|f) &= \frac{(1-\varepsilon)^2 + \frac{\varepsilon}{1-\mu}\frac{\kappa}{U-\kappa}}{(1-2\varepsilon)\left(1-\varepsilon-2\varepsilon\delta(f)\frac{1-x}{x}\frac{\kappa}{U-\kappa}\right)} > 1 \\ O_c(\chi|f) &= \varepsilon\frac{\kappa}{(1-\mu)(U-\kappa)}\frac{1-\varepsilon-\frac{\varepsilon}{x}}{(1-2\varepsilon)\left(1-\varepsilon-2\varepsilon\delta(f)\frac{1-x}{x}\frac{\kappa}{U-\kappa}\right)} \\ L_{c1}(\chi|f) &= \frac{1}{(1-2\varepsilon)\left(1-\varepsilon-2\varepsilon\delta(f)\frac{1-x}{x}\frac{\kappa}{U-\kappa}\right)} > 1 \end{aligned}$$

If we set $U - \kappa = \tilde{U}$ and $\phi = 0$ then we have the model where we only generalize the probabilities of trades. Notice that in this case even the experiment functions are the same as before, the only difference is that x has been replaced with y . However this change illustrates an important fact about the value of money. Money only has a benefit if a double coincidence of wants is less likely than a single coincidence of wants ($y < 1$). And instead of $x \rightarrow 0$ representing a complex society the important fact is that $y \rightarrow 0$. This is very reassuring since the probability of x in our society (with directed search and other elements not included in this model) is very high, but the probability of barter is extremely low. Most of us are part of a complex production process, and our marginal product per unit of output is very low, much lower than the value of the goods we purchase. For example as professors we probably have not benefited any one student (or family) enough to be able to barter for an automobile, but the impact on the whole class is far from trivial.

The relationship between ϕ and κ is intuitive and worth discussing. The production cost functions as a one shot transaction cost, the flow value is received every period. Thus this ratio is essentially the difference between receiving ϕ forever and paying κ once. The denominator is easiest to understand if we set $\chi = \kappa = 0$. Then the equilibrium condition is:

$$\phi + (1 - \mu) x \pi_c(f) U = (1 - \mu) xyU$$

and one can clearly see that increasing ϕ means one can decrease $\pi_c(f)$ by $(1 - \mu) xU$. Notice that Wright [29] also derives a similar mixed population equilibrium, but in his analysis x does not appear because of a normalization.

We also can now find a general condition when the mixed population equilibrium does not exist. For purposes of analysis it is best to set $\chi = 0$, the general conditions can easily be derived from the mixed population equilibrium. If

$$(1 - \mu) xy(U - \kappa) > (1 - \mu) x(U - \kappa) + \phi - \rho\kappa \quad (7)$$

then there is no money equilibrium and if

$$(1 - \mu) xy(U - \kappa) < \phi - \rho\kappa \quad (8)$$

then there is no barter equilibrium. If there is no barter equilibrium then we must also have

$$(\rho + \mu x) U + (1 - \mu) xyU \geq \phi - \rho\kappa \quad (9)$$

or the only equilibrium is autarky because no one holding money wants to trade for a consumption good.

Here we can see the general importance of the sign of $\phi - \rho\kappa$. If $\phi - \rho\kappa > 0$ then as $x \rightarrow 0$ there is only the money equilibrium, if $\phi - \rho\kappa < 0$ then as $x \rightarrow 0$ there is only the barter equilibrium. If $y \rightarrow 0$ then we only have both equilibria when $\phi - \rho\kappa < 0$.

The condition for stochastic stability has also been modified by these changes, it is now:

$$\frac{1 - \mu}{\mu} > \frac{1 - 2 \frac{\phi - \rho\kappa}{x(U - \kappa)} \bar{L}_{c1}(\chi)}{1 - 2y \bar{L}_c(\chi) + 2 \frac{\phi - \rho\kappa}{x(U - \kappa)} \bar{L}_{c1}(\chi) - \bar{O}_c(\chi)} \quad (10)$$

and money is always stochastically stable (again let $\chi = 0$) when:

$$\phi - \rho\kappa > \frac{1}{2} x(U - \kappa)$$

never when:

$$\frac{1}{2}x(U - \kappa) + \phi - \rho\kappa < 2yx(U - \kappa) .$$

One can easily verify that there are cases where both equilibria exist and yet one or the other is always stochastically stable.

If we consider ϕ as a choice variable then we can see that within the model it should always be set at the highest feasible level. For example welfare in the barter equilibrium is:

$$W(B) = \frac{1 + \rho}{\rho} \left[\beta(F|B, \rho) [\mu(1 - \mu)x(U - \kappa)] + \mu\phi + (1 - \mu)^2 xy(U - \kappa) \right] .$$

We can not write this as a simple function of $U(c)$ and $U(m)$ since both give different returns in the different equilibria,

$$\begin{aligned} U(m) &= 1_F(1 - \mu)xU + \phi \\ U(c) &= (1 - \mu)xy(U - \kappa) - 1_F\mu x\kappa \end{aligned}$$

where $1_F = 1$ if the economy is in the monetary exchange trading equilibrium and zero if it is in the barter equilibrium. One can easily establish that $\beta(F|B, \rho)$ (and $\beta(F|F, \rho)$, and $P(F)$) is increasing in ϕ , thus within the model increasing ϕ is always welfare increasing. In order to close the model one either must impose a social cost of increasing ϕ or model ϕ as a transfer from people who consume.⁶ In both models one finds that the optimal ϕ is greater than zero, and furthermore that as $y \rightarrow 0$ that barter ceases to be an equilibrium.

Notice that if we consider ϕ as the rate of return on loans then this could cause a uni-directional evolution. As shown in Berentsen, Camera, and Waller [5] such loans can be Pareto improving in search models of money. Thus before the monetary exchange becomes common $\phi = 0$, after monetary exchange becomes common $\phi > 0$ and barter might cease to be an equilibrium. Of course ϕ would depend on the day-to-day demand and supply of money in such a model, but our learning algorithm implies people will learn the long run value of ϕ , and not be strongly affected by day-to-day fluctuations. Thus a barter economy could evolve to a monetary exchange economy and never look back.

7 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 an intrinsically 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.

While our result shows that fiat money can arise, it does not show how fiat money did arise historically. As Dowd [7] argues a direct transition to fiat money has never happened in history. Perhaps in this respect our model would be better analyzed as a model of commodity money, and an important insight from our model is that the value of this commodity can be zero and yet commodity money will still arise.

In fact the most surprising thing, given our analysis, is that money has taken so long to arise. A critical part of understanding the answer to this is noting that the supply of money in our model is fixed. There

⁶We investigated this when utility was based on a divisible consumption good and agents had constant elasticity of substitution utility functions.

are two dimensions to trust in a model of money, and we have avoided the one of trusting the agency that issues the money. As Ritter [21] suggests this is not a simple problem, though Berentsen [4] shows that it is possible for private monopolies to issue a trustworthy currency. Notice that one result in Berentsen is the flow value of money is negative (there is inflation) in our general model we realize this can also make money less stable. These issues should be investigated.

Another issue that should be explored is the evolution of the price level. A fundamental problem in this regard is how to handle the distribution of money. Every tractable model has used either a continuum of agents (Shi [25] and Green and Zhou [9]) or a centralized market (Lagos and Wright [18]) to make sure that the distribution of money is predictable. We intend to use the alternative specification of assuming quantity is divisible and money is not like the Shi-Trejos-Wright model ([24] and [26]). In this model we will analyze both the method and terms of trade.

There is no end to other directions this literature could be taken. An important direction would be to investigate a simple analytic model which has a barter equilibrium, a commodity money equilibrium, a fractionally based commodity money equilibrium, and a pure fiat monetary exchange equilibrium. Unfortunately we do not know of a tractable model which has all four equilibria. Another interesting direction would be partially directed search (Corbae, Temzildes and Wright [6]).

In the end we have established the core result. Not only is money Pareto superior to barter but it will also arise naturally in barter economies. It is left to future research to discover when this insight holds in more general models.

8 Appendix

Proof of Lemma 1. If i trades at $\{c_{-j}, c_{-i}\}$ and j trades with probability $\eta_j > 0$ the trade will take place with probability $(1 - \chi)\eta_j$. If i does not trade then the trade will take place with probability $\chi\eta_j < (1 - \chi)\eta_j$. 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 $\Pi(m|\sigma)$, and $\Pi(c|\sigma)$ 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 he gets $U + \frac{1}{1+\rho}\Pi(c|\sigma)$ if he does not trade then he gets $0 + \frac{1}{1+\rho}\Pi(m|\sigma)$. The proof will be done when we establish upper bounds for $\Pi(m|\sigma)$.

If the strategy calls for trading at $\{m, c_{-i}\}$ then by optimality we know that $-\kappa + \frac{1}{1+\rho}\Pi(c|\sigma) \geq 0 + \frac{1}{1+\rho}\Pi(m|\sigma)$ and we are done. Thus the strategy must require no trading at $\{m, c_{-i}\}$ and $\{m, c_i\}$. If $\chi = 0$ then this person will never trade when holding money thus $\Pi(m|\sigma) = 0 < U + \frac{1}{1+\rho}\Pi(c|\sigma)$ since $\Pi(c|\sigma) \geq 0$ by optimality. Thus if there is a χ such that $0 + \frac{1}{1+\rho}\Pi(m|\sigma) > U + \frac{1}{1+\rho}\Pi(c|\sigma)$ $\Pi(m|\sigma)$ must be increasing in χ . But

$$\begin{aligned} \Pi(m|\sigma) &= (1 - (1 - \mu)\chi) \left(0 + \frac{1}{1+\rho}\Pi(m|\sigma)\right) + (1 - \mu)\chi \left(xU - (1 - x)\kappa + \frac{1}{1+\rho}\Pi(c|\sigma)\right) \quad (11) \\ \frac{\partial \Pi(m|\sigma)}{\partial \chi} &= (1 - \mu) \left(\left(xU - (1 - x)\kappa + \frac{1}{1+\rho}\Pi(c|\sigma)\right) - \left(0 + \frac{1}{1+\rho}\Pi(m|\sigma)\right) \right) \end{aligned}$$

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

Proof of Lemma 3.

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

$$\Pi(m|F) \geq \Pi(c|F) \quad (12)$$

or

$$(P_{c_i}(m) - P_{c_i}(c))U \geq (P_{c_{-i}}(m) - P_{c_{-i}}(c))\kappa \quad (13)$$

and this requires that:

$$\begin{aligned} (P_{c_i}(m) - P_{c_i}(c))U &\geq (P_{c_{-i}}(m) - P_{c_{-i}}(c))\kappa \quad (14) \\ (1 - \mu)(1 - \chi)x(1 - x)(1 - 2\chi)U &\geq -(1 - \mu)(1 - x)x\chi(1 - 2\chi)\kappa \end{aligned}$$

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

$$\begin{aligned} (P_{c_i}(c) - P_{c_i}(m))U &\geq (P_{c_{-i}}(c) - P_{c_{-i}}(m))\kappa \quad (15) \\ (1 - \mu)(1 - \chi)x^2(1 - 2\chi)U &\geq (1 - \mu)(1 - x)x\chi(1 - 2\chi)\kappa \end{aligned}$$

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

$$\frac{1}{1+\rho}\Pi(m) \leq -\kappa + \frac{1}{1+\rho}\Pi(c) \quad (16)$$

or

$$(P_{c_{-i}}(c) + \rho + P_{c_i}(m) + P_m(c))\kappa \leq (P_{c_i}(c) - P_{c_i}(m))U \quad (17)$$

If $\chi = 0$ and $\rho < (1 - \mu)x^2\frac{U}{\kappa}$ one can easily show this is true. One can also show that the left hand side is increasing in χ and the right hand side is decreasing and that when $\chi = \frac{1}{2}$ the right hand side is zero and

the left hand side positive. Thus there is a maximal χ for each ρ such that this holds, call this $\chi(\rho)$, and $\chi^* = \min \left\{ \frac{1}{2}, \frac{xU}{xU+(1-x)\kappa}, \chi(\rho) \right\}$. ■

Proof of Lemma 4. In a mixed population equilibrium an observer must be indifferent between trading for a unit of money or not. This means that:

$$\frac{1}{1+\rho}\Pi(m) = \frac{1}{1+\rho}\Pi(c) \quad (18)$$

or

$$U(m) = U(c) \quad (19)$$

$$P_{c_i}(m)U - P_{c_{-i}}(m)\kappa = P_{c_i}(c)U - P_{c_{-i}}(c)\kappa \quad (20)$$

Notice that in the mixed strategy equilibrium only $P_{c_i}(m)$ is strategic—depending on the individual one is interacting with—and this is only a function of $\pi_c(f)$, thus $\pi_c(f)$ is independent of $\pi_m(f)$ and after some calculation can be found to be the value above. Further notice that:

$$O_{c3}(\chi|f) = \chi \frac{3 - 2\chi + \delta(f)(1 - 2\chi) \frac{\kappa}{U} \frac{1-x}{x}}{(1 - 2\chi)(1 - \chi) + \delta(f)\chi(1 - 2\chi) \frac{\kappa}{U} \frac{1-x}{x}} \quad (21)$$

is increasing in $\delta(f)$. Thus $\pi_c(F) > \pi_c(\tilde{F})$. The equation equalizing the inflow and outflow of agents in the money state is:

$$\begin{aligned} & (1 - \mu)\mu\pi_m(f)(1 - \pi_c(f))(x(1 - \chi)\chi + (1 - x)\chi^2) \\ & = (1 - \mu)\mu\pi_c(f)(1 - \pi_m(f))\left(x(1 - \chi)^2 + (1 - x)\chi(\delta(f)(1 - \chi) + (1 - \delta(f))\chi)\right) \\ & \frac{\pi_m(f)}{1 - \pi_m(f)}O_{io}(\chi|f) = \frac{\pi_c(f)}{1 - \pi_c(f)} \end{aligned} \quad (22)$$

and after manipulation this is equivalent to the condition 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 $\omega \in \Omega$, being sure that the last action pair is $\{T, T\}$.

Since $\chi > 0$ this sequence has strictly positive probability. Further note that since others are not changing their strategy $P^t = P(\sigma_i)$ where σ_i is this agent's strategy. Then we can rewrite our formula for $S_{\alpha\omega}^{t+1}$ as:

$$\begin{aligned} S_{a\omega}^t & = (1 - \tau)S_{a\omega}^{t-1} + \tau \left(U_{\alpha\omega} + \frac{1}{1+\rho}P_{\alpha\omega}(\sigma_i)S^{t-1} \right) \\ S_{a\omega}^t & = \tau U_{\alpha\omega} + D_{\alpha\omega}(\sigma_i)S^{t-1} \end{aligned} \quad (23)$$

where

$$D(\sigma_i) = (1 - \tau)I_{|A|^*|\Omega|} + \tau \frac{1}{1+\rho}P(\sigma_i) \quad (24)$$

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

$$a_i(\omega) \in \arg \max_{\alpha_i \in \{T, N\}} E_{\alpha_j} \left[S_{\alpha_i \alpha_j \omega}^t \right] \quad (25)$$

combining these two facts we can write:

$$S_{\omega}^k = \max_{\alpha_i \in \{T, N\}} E_{\alpha_j} \left[\tau U_{\alpha_i \alpha_j \omega} + D(\sigma_i(\alpha_i))S^{t-1} \right] \quad (26)$$

where $k = \left\lceil \frac{t}{|A|^*|\Omega|} \right\rceil$ and $\sigma_i(\alpha_i)$ reflects the dependence of σ_i on the agent choosing α_i at state ω . As Kushner [16] shows since $\sum_j D_{ij} = (1 - \tau) + \tau \frac{1}{1+\rho} < 1$ the vector $\{S_{\omega}^k\}_{\omega \in \Omega}$ converges to the optimal value function.

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 agent will be using the optimal strategy.

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

Proof of Proposition 2. To show that society will always converge to an equilibrium we need to show that from any initial distribution of strategies they will converge to an equilibrium with positive probability in finite time. In order to do this we need to specify a path of finite length which results in society being in one of the equilibria. Notice that the equilibrium strategies depend only on whether people trade at three states $\{c_j, m\}$, $\{c_{-j}, m\}$, and $\{m, c_{-i}\}$, thus people will only learn the optimal strategy by trading a consumption good for money. Furthermore if trading at $\{m, c_{-i}\}$ is optimal trading at $\{c_j, m\}$ is not, and trading at $\{c_{-j}, m\}$ is optimal if and only if trading at $\{c_j, m\}$ is optimal, thus players must only learn whether to trade at either $\{m, c_{-i}\}$ or $\{c_j, m\}$ or neither.

Let us assume that the (unique) best response to the current distribution is independent of any two player's strategies, and choose one player who is not best responding to the current distribution. If this player is currently holding a consumption good choose another player who is holding money (and preferably not best responding) if this player is holding money choose another who is holding a consumption good. Have these players trade money back and forth until they are both best responding to the original distribution. While they are doing this have all other players either trade money for money or consumption goods. Since the best response at these states either does not matter or is a dominant strategy this will not affect the best response to the current distribution. Thus these players can learn without affecting the best response.

Now we simply repeat this process until all players who are not best responding learn to best respond. We then repeat the process one final time to converge all players to the neighborhood of the equilibrium values given the current distribution and we are done. Since $\chi > 0$ this finite path has strictly positive probability, thus in every time period it can be entered with positive probability, and in the fullness of time it will occur with probability one. Please note that we could easily have a larger group learn at the same time, but we just choose the smallest group for simplicity of exposition. In short there are many other finite paths, we have just shown that there is certainly one.

If the (unique) best response is not independent of any two player's strategies we must be certain that both of the people in our learning pair are not best responding to the current distribution. If one player who is not best responding is holding money and another one holding a consumption good then we can use the above path with these two players, since if they switch to a best response this will not change the best response. If all of them are holding money then we can have one of them trade at the state $\{c_j, m\}$, and since this is a dominant strategy for the opponent the opponent's strategy will not change. We can then use the path we found above. If all of them are holding a consumption good then have one of them trade for money. This may result in a new distribution, but in this new distribution there is at least one player who was and was not best responding to the old distribution holding money. Thus we can use the path above without loss of generality.

If the best responses are not unique then this means society is already in a mixed strategy equilibrium and we are done. ■

Proof of Lemma 5. From Proposition 2 we know that for given I an upper bound on the radius is:

$$\begin{aligned} r^*(F) &\leq \frac{[(1 - \pi_c(F))(1 - \mu)I]}{I} \\ r^*(B) &\leq \frac{[(1 - \mu)\pi_c(\tilde{F}) + \mu]I}{I} \end{aligned}$$

and if π^* is a mixed population equilibrium $r^*(\pi^*) \leq \frac{2}{I}$. To see the first two equations notice we can have two people trade for money before the mutations occur and then continue trading for money until a sufficient number of mutations has happened. Then we can use the path in Proposition 2 to converge to the other equilibrium with positive probability. Since $1 > \chi > 0$ the possibility that people will trade for money is always strictly positive and thus this path will lead to leaving on of the pure strategy equilibria. For the mixed strategy equilibria this upper bound is sufficient. For the pure strategy equilibria we must also find the lower bound. In other words we must find a distribution such that with probability one we will not exit either the barter or the fiat money equilibrium. We will first prove this starting at the B equilibrium.

This distribution will actually be $\frac{[(1 - \mu)\pi_c(\tilde{F}) + \mu]I - 1}{I}$, or if just one too few people mutate then we will return to the B equilibrium. First notice that we can assume that people's beliefs are as close as necessary to the true values in the B equilibrium, or that for all $\lambda > 0$ and $v > 0$ there is a $\varepsilon > 0$ such that if the probability of a mutation is less than this ε $P([S^t - \Pi(B)] < \nu) > 1 - \lambda$ for all people in the society. This is an immediate implication of Proposition 1 and the fact that mutations only occur with probability ε . Thus

we can assume all people are using the strategy B and that $[S^t - \Pi(B)] < \nu$, where the value of ν will depend on the population size and be selected below.

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} .

Now assume that $\left[I \left((1 - \mu) \pi_c \left(\tilde{F} \right) + \mu \right) \right] - 1$ people mutate to the \tilde{F} strategy. Then we know that since I is discrete $\Pi(c|\tilde{B}) - \Pi(m|\tilde{F}) \geq 2\nu > 0$ for all χ . Let this be the ν 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 $\{\alpha, \omega\}$ $S_{a\omega}^t - S_{a\omega}^*$ is negative, thus every time $\{\alpha, \omega\}$ is updated $S_{a\omega}^t - S_{a\omega}^*$ 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 \tilde{B} strategy the initial value of the strengths $\left\{ S_{\{N,T\}\{c_j,m\}}^t, S_{\{N,N\}\{c_j,m\}}^t, S_{\{N,T\}\{c_j,m\}}^t \right\}$ are all at least $(1 - \chi) (1 - \mu) \frac{x^2}{\rho} U + \chi \frac{1}{1+\rho} Z_1(\chi) - \nu$ where $Z_1(\chi)$ is a function such that $\lim_{\chi \rightarrow 0} Z_1(\chi) = 0$. As well given the new distribution of strategies $\Pi(c|\tilde{B}) = (1 - \chi) x^2 (1 - \mu) \frac{1+\rho}{\rho} U + \chi Z_2(\chi)$ where $\lim_{\chi \rightarrow 0} Z_2(\chi) = Z_2 > 0$.

Since $\Pi(m|\tilde{F}) \leq \Pi(c|\tilde{B}) - 2\nu \leq (1 - \chi) x^2 (1 - \mu) \frac{1+\rho}{\rho} U + \chi Z_2(\chi) - \nu$ and the limiting value of $S_{\{T,T\}\{c_j,m\}}^t$ is $\frac{1}{1+\rho} \Pi(m|\tilde{F})$ there clearly is a $\bar{\chi}$ such that

$$\begin{aligned} (1 - \bar{\chi}) (1 - \mu) \frac{x^2}{\rho} U + \bar{\chi} \frac{1}{1+\rho} Z_2(\bar{\chi}) - \frac{1}{1+\rho} \nu &< (1 - \bar{\chi}) (1 - \mu) \frac{x^2}{\rho} U + \bar{\chi} \frac{1}{1+\rho} Z_1(\bar{\chi}) \\ \bar{\chi} (Z_2(\bar{\chi}) - Z_1(\bar{\chi})) &< \nu \end{aligned} \quad (27)$$

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

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 one trades 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 χ , thus for small enough χ the statement is correct. ■

Proof of Theorem 1. From Young [30] we know that a limit set (equilibrium in this game) is stochastically stable if it has the minimum stochastic potential. From Hasker [10] we know that the stochastic potential of an equilibrium σ^* is:

$$c_{\sigma^*}^* = c^*(E) - r^*(\sigma^*) + ca(\sigma^*) \quad (28)$$

where E is the emergent seed, $c^*(\cdot)$ is the cost of going between limit sets, and $ca(\cdot)$ is the cost of going from the emergent seed to the limit set. 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 $ca(\sigma^*) \leq \frac{1}{I}$.

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:

$$\begin{aligned} c_F^* &\leq c_{\pi^*}^* + \frac{1}{I} - r^*(F) \\ &\leq c_{\pi^*}^* + \frac{1}{I} - \frac{[I(1 - \pi_c(F))(1 - \mu)]}{I} \end{aligned} \quad (29)$$

and clearly this is negative for $I \geq \frac{1}{1-\mu} + 1$.

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:

$$-\frac{[I(1 - \pi_c(F))(1 - \mu)]}{I} + \frac{1}{I} < -\frac{[I((1 - \mu) \pi_c(\tilde{F}) + \mu)]}{I} \quad (30)$$

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

References

- [1] Barto, A., Bradtke, S., Singh, S. (1995) “Learning to act using real-time dynamic programming.” *Artificial Intelligence* 72.
- [2] Basci, Erdem. (1999) “Learning by Imitation”; *Journal of Economic Dynamics and Control*, 23:1569-85.
- [3] Benveniste, Albert; Michel Metivier; and Pierre Priouret. (1987). *Adaptive Algorithms and Stochastic Approximations*. Springer-Verlag; Berlin.
- [4] Berentsen, Aleksander. (2006) “On the Private Provision of Fiat Currency.” *European Economic Review*, Vol. 50(7), pages 1683-1698.
- [5] Berentsen, Aleksander; Gabriele Camera, and Christopher Waller. (2005) “The Distribution of Money Holdings and the Non-Neutrality of Money.” *International Economic Review*, Vol. 46, No. 2, pp 465-487.
- [6] Corbae, Dean; Ted Temzilides; and Randall Wright. (2003) “Directed Matching and Monetary Exchange.” *Econometrica*. Vol 71:731-56.
- [7] Dowd, Kevin (2001) “The Emergence of Fiat Money: A Reconsideration” *Cato Journal* 20: 467-76.
- [8] Freidlin, Mark and Alexander Wentzell. (1984) *Random Perturbation of Dynamical Systems*. Springer-Verlag, New York.
- [9] Green, Edward; and Ruilin Zhou. (2002) “Dynamic Monetary Equilibrium in a Random Matching Economy ” *Econometrica*, 70:929-969.
- [10] Hasker, Kevin. (2004) “The Emergent Seed: Simplifying the Analysis of Dynamic Evolution.” Bilkent Economics Department Discussion Paper 04-06. Available at: <http://www.bilkent.edu.tr/~hasker/Research/Hasker-Emergent-Seed-05-08-26.pdf>.
- [11] Iwai, Katsuhito. (2001). “Evolution of Money.” in *The Evolution of Economic Diversity*. Antonio Nicita and Ugo Pagano, eds. Routledge, London and New York.
- [12] Jevons, W. Stanley. (1875). *Money and the Medium of Exchange*. King, London.
- [13] Kandori, Michihiro; George Mailath, and Rafael Rob. (1993): “Learning Mutation, and Long Run Equilibria in Games.” *Econometrica*, 61:9-56
- [14] Kiyotaki, Nobuhiro and Randall Wright. (1989) “On Money as a Medium of Exchange.” *Journal of Political Economy*. Vol. 97:927-54.
- [15] Kiyotaki, Nobuhiro and Randall Wright. (1993). “A Search-Theoretic Approach to Monetary Economics. *American Economic Review*. 83:63-77.
- [16] Kushner, Harold (1972) *Introduction to Stochastic Control Theory*. Holt, Reinhart, and Winston; New York.
- [17] Kushner, Harold (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>

- [18] Lagos, Richard and Randall Wright. "A Unified Framework for Monetary Theory and Policy Analysis." *Journal of Political Economy*. Vol. 113:463-84.
- [19] Lettau, Martin and Harald Uhlig. (1999). "Rules of Thumb versus Dynamic Programming." *American Economic Review*. 89:148-174.
- [20] Marimon, Ramon; Ellen McGrattan; and Thomas Sargent. (1990). "Money as a Medium of Exchange in an Economy with Artificially Intelligent Agents." *Journal of Economic Dynamics and Control*. 14:329-373.
- [21] Ritter, Joseph. (1995). "The Transition from Barter to Fiat Money." *American Economic Review*. 85:134-149
- [22] Sandholm, William. "Simple Formulas for Stationary Distributions and Stochastically Stable States." Forthcoming in *Games and Economic Behavior*.
- [23] Selgin, George. (2003). "Adaptive Learning and the Transition to Fiat Money". *Economic Journal*, 113:147-65.
- [24] Shi, Shouyong. (1995) "Money and Prices: A Model of Search and Bargaining." *Journal of Economic Theory*. 67, 467-496.
- [25] Shi, Shouyong. (1997) "A Divisible Search Model of Fiat Money." *Econometrica*, 65, 75-102.
- [26] Trejos, Alberto and Randall Wright. (1995) "Search, Bargaining, Money and Prices." *Journal of Political Economy*. Vol. 103:118-41.
- [27] Watkins, Chris (1989). "Learning with Delayed Rewards." Ph.D. thesis, Cambridge University, Psychology Department.
- [28] Wright, Randal. (1995) "Search, Evolution, and Money." *Journal of Economic Dynamics and Control*, 19:181-206.
- [29] Wright, Randall. (1999) "A Note on Asymmetric and Mixed Strategy Equilibria in the Search-Theoretic Model of Fiat Money." *Economic Theory*, 14:463-471.
- [30] Young, Peyton. (1993): "The Evolution of Conventions." *Econometrica*, 61:29-56.