Civilization and the evolution of short sighted agents

Basuchoudhary, Atin and Allen, Sam and Siemers, Troy

Virginia Military Institute

19 November 2008

Civilization and the Evolution of Short-Sighted Agents

Atin Basu
Department of Economics & Business
Virginia Military Institute

Troy Siemers
Department of Mathematics & Computer Science
Virginia Military Institute

Samuel K. Allen[†]
Department of Economics & Business
Virginia Military Institute

November 2008

Preliminary: Please do not cite.

Abstract:

We model an assurance game played within a population with two types of individuals – short-sighted and foresighted. Foresighted people have a lower discount rate than short sighted people. These phenotypes interact with each other. We define the persistent interaction of foresighted people with other foresighted people as a critical element of civilization while the interaction of short sighted people with other short sighted people as critical to the failure of civilization. We show that whether the short sighted phenotype will be an evolutionary stable strategy (and thus lead to the collapse of civilization) depends on the initial proportion of short sighted people relative to people with foresight as well as their relative discount rates. Further we explore some comparative static results that connect the probability of the game continuing and the relative size of the two discount rates to the likelihood that civilization will collapse.

[†] *Corresponding Author*: allensk@vmi.edu

1 Introduction

The institutions of civilization are designed to reduce uncertainty (North, 2005, p. 1-13). A reduction in uncertainty generates benefits for individuals residing in a civilized society. Whether civilization succeeds or not depends on how well these institutions are designed. This design, in turn, is based on how well the designers – humans – perceive reality. We postulate that human perception of reality depends on the ability to identify and compare the costs and benefits of decisions over time. However, even with this sort of rational decision making process, civilizations can fail if the proportion of people who make good decisions over a long planning horizon is small enough.

Perceptions of reality are driven by a biological cognitive process that finds expression in human culture and belief systems (Wilson, 1998, p. 66). Since the biological cognitive process is less well known than human cultures and belief systems, we look for the reasons for misperceiving reality within such belief systems.[1] North (2005, p. 13-22), for example, argues that agents who believe they live in an ergodic world can have serious misalignments with reality in non-ergodic situations. This happens because experience of the past cannot prepare one for the non-ergodic events of the future. This lack of preparedness against shocks can lead to social failure. Tuchman (1984) makes the case that sometimes people behave "irrationally" because they somehow hold on to value systems that are incapable of dealing with a new reality. Diamond (2005, p. 434) suggests that this "irrationality" could be based on differences in planning horizons. People may be so focused on the current benefits of certain civilizational values that they ignore the future costs of their belief systems. This

[1] Scientists currently debate the extent of biological determinism in human culture. This debate over the biological cause of human misperceptions of reality is important. However, this debate is still quite nascent. Like Plato's cave dwellers then we look for the causes of human misperceptions of reality in the cultural shadows cast by some "real" but unknown cognitive process. This of course leaves us wide open to the criticism that we may be ourselves misperceiving reality in this paper!

approach to understanding the collapse of society can therefore be interpreted as the efforts of rational individuals weighing costs and benefits of their particular civilizational belief system. The future costs of these belief systems increase as beliefs diverge from reality. Short sighted agents may ignore these future costs. Thus, even if the decision making processes of these individuals are rational, then the relative weights they place on costs and benefits can lead to socially sub-optimal decisions. We argue, that even in an ergodic world, decisions made by rational people can lead to a collapse of society simply because people may have shorter planning horizons; i.e., high discount rates.

"Successful" civilizations tend to adopt institutions or mechanisms to promote cooperation that leads to greater well-being. We postulate that long-run progress for a society depends on its members' ability both to cooperate and to select actions that are consistent with the goal of maintaining the sustainability of resources over the long-run. Diamond (2005, p. 434) for example cites the case of poor fishermen who dynamite or poison coral reef fish to sustain themselves even with the knowledge that their actions kill the very reefs that sustain their food. Thus, societies with a culture of cooperation over a distant planning horizon are more likely to be successful. Societies marked by a lack of cooperation over the long term make short sighted decisions that lead to societal collapse. In other words, we operationalize one strand of the epistemological discussion on the causes of societal collapse by focusing on differences in how different people view the future.

We model the actions of individuals using a simple "assurance" game to illustrate that in the long-run, short-sighted and foresighted individuals may coexist, but one group will always represent the overwhelming majority. However, if each type of individual behaves deterministically then whether a society ultimately succeeds or fails depends critically on the initial proportion of types in the population. We discuss the relevance of perceptions and valuations of the future for the two types and discuss how they will influence the equilibrium. Our innovations lie in two areas. First, we interpret a well known equilibrium phenomenon in evolutionary coordination games as social collapse. Second, we argue that differences in perceptions of the future and in perceptions of whether the game continues or not are important determinants of societal collapse, or success. We believe this approach yields important policy insights. Thus, our model may suggest a framework for anticipating the impacts of migration between societies and their resulting chances of success or even a framework for resolving conflict.

2 The Model

We begin by assuming the fundamental or defining trait among individuals is their degree of foresight. For simplicity, our initial model assumes that all individuals are either foresighted or short-sighted. Thus, foresight is a fixed, binary trait. Again this is common in the evolutionary game theory literature. One may think of society as having a strategy set while individuals can either inherit or choose between them (Gintis, 2000, p. 148). Our agents merely choose strategies with a higher payoff or fitness. Naturally, we then believe that those with greater foresight will care more about the future – and discount future outcomes less than their short-sighted counterparts – and therefore will choose to use resources more efficiently over the long-run.

Thus, both types behave according to their time preferences. This general framework could correspond to numerous, wide-ranging scenarios from the interactions of hunter-gatherers; decisions of business leaders in corporate settings; the actions of public policy-makers; collaborations among groups of students; decisions of agrarian societies; or the behavior of smokers and non-smokers, just to name a few.

While the implications can be generalized to other applications, we model foresight and shortsightedness in the context of a "stag hunt" game. The stag hunt game is a somewhat fanciful name used to describe the general class of "assurance" games. Our choice stems from the interpretation of an assurance game as representative of a societal dilemma described in Rousseau's *A Discourse on Inequality* (Poundstone, 1992, pp. 218-221). Maurice Cranston's (1984) translation describes the stag hunt dilemma in the following way:

> "If it was a matter of hunting a deer, everyone well realized that he must remain faithfully at his post; but if a hare happened to pass within the reach of one of them, we cannot doubt that he would have gone off in pursuit of it without scruple and, having caught his own prey, he would have cared very little about having caused his companions to lose theirs."[2]

A stag – the hunt for which requires intensive cooperation – can feed the entire village. Going after the hare would feed the individual. Thus, an assurance game may be used to understand the nature of cooperation in society. Cooperation provides a way to understand delayed gratification. Hunting a stag requires time and cooperation, but can feed the entire society. Going after the hare can feed the individual now.

We therefore model two kinds of people. One group of people can resist the temptation of going after the hare – that is they always stick to the cooperative plan of hunting the stag. In other words, these people are capable of committing to long term decisions. We will refer to this type as "Stag-types" because they possess the foresight to cooperate in endeavors that require patience to achieve an efficient outcome. Another group of people show their impulsiveness by

[2]Read *stag* for *deer* in the translated excerpt.

going after the hare. Thus, the people who go after the hare make short term decisions, i.e., they do not cooperate, and are referred to as "Hare-types."

This modeling approach allows us to give a great deal of importance to the idea that cooperation is the basis for civilization. In our model far sighted individuals will cooperate. This behavior is not necessarily altruistic. Far sighted individuals interacting with other far sighted individuals, get a bigger private benefit than short sighted individuals acting on their own. This lays out the strongest possible case for the cooperative basis for civilization. In fact, short sighted individuals who interact with far sighted individuals also receive a higher payoff than if they act alone. However, far sighted individuals are punished quite severely for their willingness to delay gratification when they interact with short sighted individuals since without cooperation the long term benefits are unattianable.[3]

In keeping with the general approach used in evolutionary game theory we proceed with a limited view of rationality. That is to say our actors are allowed to make mistakes whatever their type. Moreover, some people may be unwilling or unable to make the socially optimal choice even if it has consequences for their survival. Hence we have a distinction between those who cannot make long term decisions and those who can. The short-sighted, "Hare-types" place less value on the future, and given the option will prefer greater payoffs in the near-term even if it means overall inefficient outcomes for society.[4]

In our game there are two possible strategies that people can play. However, the strategies are entirely determined by type. Those who make short term decisions always choose Hare and those who make long term decisions always

[3] One may picture a lone Stag hunter waiting for his colleagues to show up for the hunt while they are off finding rabbits. The rabbit hunters feed their family while the village starves. The lone stag hunter is likely to decide to go after rabbits in the next period. Soon rabbits are scarce and everyone starves. Society collapses.

[4] One can similarly think of the smoker who – knowing the health risks involved – chooses to trade minutes or hours of satisfaction today for weeks or years of satisfaction in the future.

choose Stag. Let x denote the proportion of the population that always follows the Stag strategy while the proportion $(1-x)$ always follows the Hare strategy. Of course, actors playing each type of strategy interact with both other actors playing the same strategy or a different one. Keeping in mind the assurance game we can therefore specify the payoffs illustrated in Figures 1 and 2. We assume that outcomes stem from random, pair-wise matching of players.

Figure 1: Immediate Payoffs, in general

	Stag	*Hare*
Stag	$e/2, e/2$	$(1-\gamma)\alpha e, \gamma\alpha e$
Hare	$\gamma\alpha e, (1-\gamma)\alpha e$	$\alpha e/2, \alpha e/2$

Where the payoffs today depend on the following parameters[5]:

e defines the total payoff to efficient behavior

$\alpha \in (0,1)$ denotes the efficiency loss due to myopic behavior (i.e., less cooperation)

$\gamma \in [0,1]$ denotes the "greediness" of H-types, so S-types may be "punished" when interacting with H-types.

If we let $e = 6$, $\alpha = 1/3$, and $\gamma = 1$, then the payoffs will correspond to those in Figure 2 below.[6]

[5]We assume an equal division within pairs of the same type; however, relaxing this assumption should not alter the overall outcomes

[6]This framework is not a zero-sum game, so interactions do generate non-negative payoff for all pairs, and hence society will grow. There will be more treatment on the issue of growth later in the paper. Moreover, we also restrict α to be less than 1/2 else the game becomes a *prisoner's dilemma*.

Figure 2: Immediate Payoffs, a specific example

	Stag	*Hare*
Stag	$3,3$	$0,2$
Hare	$2,0$	$1,1$

However, we can further differentiate between Stag and Hare types. The Stag types are patient and therefore have a lower discount rate (r_S) than that (r_H) of the impatient Hare types. In other words, the Stag types discount their future less heavily than the Hare types. Say further that there is some chance that a meteor will hit the earth (or some other environmental disaster) and all human interaction will cease with probability $(1-p)$.[7] Thus, Stag types discount the future with a discount factor $\delta_S = p/(1+r_S)$ and Hare types discount the future with a discount factor $\delta_H = p/(1+r_H)$. Note that $\delta_S > \delta_H$. After accounting for the future stream of payoffs, the game illustrated in Figure 2 may be represented by Figure 3.

Figure 3: Present Value of Future Payoffs

	Stag	*Hare*
Stag	$3\left(\frac{1+r_S}{1+r_S-p}\right), 3\left(\frac{1+r_S}{1+r_S-p}\right)$	$0, 2\left(\frac{1+r_H}{1+r_H-p}\right)$
Hare	$2\left(\frac{1+r_H}{1+r_H-p}\right), 0$	$\left(\frac{1+r_H}{1+r_H-p}\right), \left(\frac{1+r_H}{1+r_H-p}\right)$

Note that the payoffs are structured to make fore-sighted cooperation, i.e., when stag types coordinate with other stag types, the most efficient outcome. In fact, the payoff to shortsighted agents when they interact with foresighted agents is higher than if they interact with other shortsighted agents. This is in line with our fundamental assumption that civilization, based on the ability to plan for the future, generates benefits for individuals – whatever their planning horizon.

[7] In other words, p is the probability that life, and the game, continues.

3 Equilibrium Strategy

We are interested in explaining the success and failure of civilization as a function of the ability of people to plan for the future.

We note here that cooperation (Stag, Stag) is the payoff dominant equilibrium in this game. Thus players "should not have any trouble coordinating their expectations at the commonly preferred equilibrium point" (Harsanyi and Selten, 1988, p. 81). We interpret this cooperative equilibrium as civilization precisely because it is the payoff dominant equilibrium. However, the non-cooperative equilibrium (Hare, Hare) risk dominates the cooperative equilibrium in the sense that a deviation from the non-cooperative equilibrium will lead to a greater loss in payoff relative to a deviation from the cooperative equilibrium (see Aumann, 1987 for a discussion of the trade-off between efficiency and risk). Thus rational people may choose the less efficient equilibrium (Weibull, 1997, p. 31). We interpret this lack of cooperation as social breakdown and the fall of civilization. It seems to be a matter of interest, then, to find the conditions under which such breakdown is likely.

We show that whether one equilibrium or the other is a steady state is a function of the initial proportion of farsighted people to shortsighted people as well as the relative gap in discount rates of the two types. Thus, our model shows that the success or failure of a civilization could be an artifact of pure chance since the initial proportion of shortsighted people to farsighted people could be a matter of pure chance. Further we show cooperation, or civilization, also depends on the relative size of the discount rates through which the two types view the future.

Theorem 1 *The Hare-type behavior is an evolutionary-stable "strategy" (ESS) only if* $x < \frac{1}{3\left(\frac{1-\delta_H}{1-\delta_S}\right)-1}$. [8]

[8] See Appendix A for the formal proof. Throughout the paper we will use the payoffs

This theorem suggests that social failure (i.e., inefficient outcomes for society) depends crucially on the initial proportion of people who are Hare types. This in turn depends on the discount factor of people with a long term planning horizon relative to a short term planning horizon.

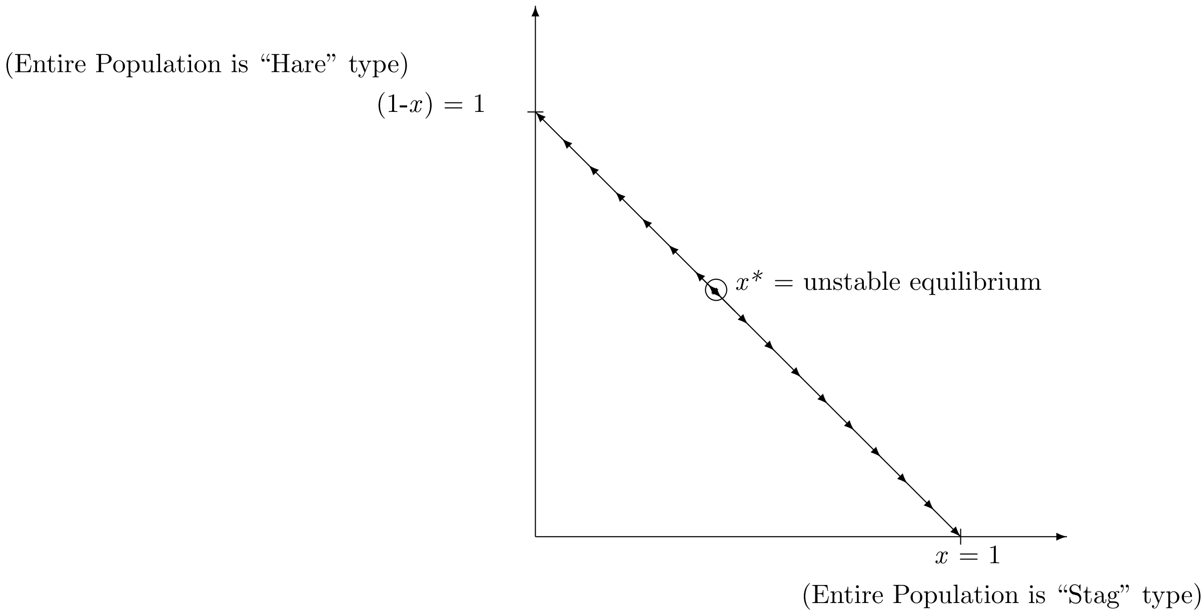
The expected payoff to the Hare types is higher on average than the Stag types under the conditions listed in Theorem 1. Stag type people benefit only when there are other Stag types around.[9] Thus, the expected payoff to Stag types is higher than that of Hare types if and only if the proportion of the Stag types is relatively large. We refer to the "tipping point," as x^*, which is derived in the proof of Theorem 1: $x^* = \frac{1}{3\left(\frac{1-\delta_H}{1-\delta_S}\right)-1}$. If the proportion of Stag-types is less than x^*, then Hare-types will eventually dominate society.

Since there are no à priori reasons to believe that the proportion of one type of actor relative to the other in a civilization will be large or small, the success or failure of a civilization may be determined by pure chance! In other words, inefficient outcomes may persist (David, 1985).

assumed in Figure 2.

[9]This occurs because $\gamma = 1$, an assumption that could be relaxed.

Figure 4: Proportions of Types



Corollary 1 *There is no stable polymorphic ESS.*

The expected payoff to both types monotonically increases with respect to the proportion of Stag types. If $x < \frac{1}{3(\frac{1-\delta_H}{1-\delta_S})-1}$ then the expected payoffs favor the Hare-type (i.e., $E(H) > E(S)$). In this case the population will be taken over by the Hare types. This occurs because we further assume that population growth of each type is tied to their relative payoffs. If $x > x^*$ then $E(H) < E(S)$. In this case the population will be taken over by Stag types. Thus, very small mutations of behavior around x^* will move the population toward a monomorphic equilibrium at Hare or Stag depending on whether the mutation favors Hare or Stag. In other words, an equilibrium where $x = x^*$ is not stable.

This result has important policy ramifications. A shock to the actual proportion of Hare types relative to Stag types could alter the very nature of a society in a very deterministic way. This could occur due to migration or a random event. Let x^* denote the proportion (of Stag-types in the population) corresponding to this unstable equilibrium. Then, at any given point in time either (a) the proportion of foresighted individuals is greater than x^* is moving toward long-run success, or (b) the proportion of foresighted individuals is less than x^* is moving toward long-run inefficiency (failure). In either case, the importance – of a shock – for society lies almost entirely in whether or not the proportion x shifts to the other side of x^*. On a smaller scale this intuition could be used to manage conflict at a tactical level. Eliminating short sighted agents at a given point in time could increase the proportion of foresighted people – thus increasing cooperation and reducing conflict.[10]

Moreover, small mutations – if a Hare type mistakenly cooperates or vice versa – will unravel the polymorphic equilibrium toward one where everyone learns to be foresighted or short-sighted.[11]

4 Discounting, Expectations & Equilibria

We note that whether a society moves toward the efficient equilibrium or not depends on the actual proportion of foresighted people relative to the proportion of foresighted required to make society move toward the efficient equilibrium. Thus, if the actual proportion of foresighted people is less than this requirement, then society moves toward collapse. However, this "required" proportion of foresighted people depends on two things in our model. First, it depends on the probability the game continues. Second, it depends on the relative discount

[10] This result also has ominous overtones. Abimael Guzman, a philosophy professor who leads the terrorist Shining Path outfit in Peru, famously advocates the death of 10 percent of the Peruvian population as a means of tipping Peru toward a Maoist "utopia" (Washington Post, 2006)

[11] Note that an ESS is always perfect but not vice versa. In this case the polymorphic equilibrium is perfect but it is not ESS because small mistakes unravel the equilibrium (Selten, 1975).

rates of each type of individual. In this section, we look at this relationship with a view toward understanding how these parameters affect the likelihood of societal collapse.

The unstable equilibrium corresponds to the populations where the proportion of Stag-types is equal to x^*. However, this, unstable, tipping point for society depends upon the likelihood that life continues, p, and the relative discount rates that each type places on the future, r_S and r_H. Without loss of generality, we assume that $r_S = 0$. The comparative statics relating to x^* reveal how this unstable equilibrium changes with respect to p and r_H.

Theorem 2 *The unstable equilibrium, x^*, is decreasing in p, the probability that life continues.*

This essentially means that the bleaker the future becomes, the greater the proportion of foresighted, patient individuals required to achieve an efficient societal outcome.[12] Here is the intuition underlying this result: with a less certain future (i.e., as p moves closer to zero than to one), the present value of payoffs to both types diminish, as illustrated in Figure 2. However, the payoff decreases relatively less for the Hare types since they placed less value on the future anyway. (Mathematically, this is resulting from r_H in the denominator of the payoff for Hare types.) Thus, with a bleaker future, or lower p, the payoff and hence growth of Hare types is relatively larger. This means that achieving the efficient outcome (the equilibrium with all Stag types) requires more foresighted individuals.

Corollary 2 *The rate at which x^* decreases in p depends on r_H.*

The higher the discount rate of the short sighted type, the greater the difference between the short sighted and the farsighted types. Thus, even as p moves

[12] See Appendix C for the derivative of x^* with respect to p.

closer to zero and the present value of both types diminish, the payoff divergence between Hare types and Stag types rises with r_H. Thus, with a high r_H (given our payoffs this means an $r_H > 2$) x^*decreases in p at an increasing rate. Further, with a low r_H (given our payoffs this means an $r_H < 2$) x^*decreases in p at a decreasing rate.[13]

Theorem 3 *The unstable equilibrium, x^*, is decreasing in r_H, the discount rate of the short-sighted type of people.*

This implies that the less the short-sighted individuals care about the future (i.e., the greater r_H), the lower the proportion of foresighted, patient individuals required to achieve an efficient societal outcome.[14] The intuition here is as follows: As short-sighted individuals receive less benefit from their future payoffs (this happens increasingly as r_H rises), then their relative expected payoffs diminish compared with those of foresighted people. Thus, higher discounting of the future means that the sub-population of short-sighted individuals will grow less quickly and, all else equal, society will need fewer foresighted people to eventually achieve the efficient outcome.

Corollary 3 *As r_H rises, x^* falls at an increasing rate.*

The higher the discount rate of the short sighted type, the greater the difference between the short sighted and the farsighted types. The divergence between the short sighted individual's future payoffs and the foresighted individual's payoff increases as r_H rises. This drives the increased rate with which x^* falls as r_H rises.[15]

[13]See Appendix C for the second derivative of x^* with respect to p.
[14]See Appendix C for the derivative of x^* with respect to r_H.
[15]See Appendix C for the second derivative of x^* with respect to r_H.

5 Population Growth

The sub-populations of each type grow based on the relative expected payoffs. In other words, if the expected payoff to Stag-types is greater than for Hare-types, then the population of Stag-types will grow more (in percentage terms) in the next time period. Here we view time discretely, but it can also be viewed as a continuous measure. The growth pattern is defined in Appendix B and examples are provided in Tables 1 and 2.

Generally, we allow our agents to follow a simple replicator dynamic process.[16] In our example this implies that the total population grows at a rate proportional to $x_tE(S_t)+(1-x_t)E(H_t)$ in each period t. Where $E(S_t)$ and $E(H_t)$ are the expected payoffs to the sub-populations of S-types and H-types, respectively. Moreover, in each period t the S-types grow at a rate $\frac{E(S_t)}{x_tE(S_t)+(1-x_t)E(H_t)}$ while the H-types grow at a rate $\frac{E(H_t)}{x_tE(S_t)+(1-x_t)E(H_t)}$.

Theorem 1 implies that if $x < x^*$ then, $\frac{E(S_t)}{x_tE(S_t)+(1-x_t)E(H_t)} < \frac{E(H_t)}{x_tE(S_t)+(1-x_t)E(H_t)}$. We note that in this process no one dies and the population grows at a steady pace. Within this growing population the proportion of each type of individual grows at a differential rate each period. Thus, it is possible to specify the number of periods it will take to reach the equilibrium predicted in Theorem 1 as a function of certain parameters by using this replicator dynamic process. A representative simulation illustrates this point.

[16] There are many ways to define evolutionary dynamics. Replicator dynamics however are generally accepted as standard (Gintis, 2000, p. 190)

Table 1: Example of Population Growth, if $x_0 > x^*$

Here we assume the payoffs are those found in Figure 3 and, the discount rate for the H-types is 0.5 (i.e., $r_H = 0.5$), the discount rate for S-types is normalized to zero (i.e., $r_S = 0$), and if the probability that the game continues is 80% (i.e., $p = 0.8$), then $x^* = 1/6$, and the path of x_t will follow the pattern in the table below. Under these conditions, H-types will grow at a slower pace than the S-types relative to the average growth rate of the population. Remember, if $x_0 > x^*$, then the efficient, foresighted equilibrium will result, where $x = 1$. If 30% of the population is foresighted to begin with (i.e., $x_0 = 0.3$), then the follow pattern will emerge:

t	x_t	**Expected Payoff for S-types**	**Expected Payoff for H-types**	**Relative Growth of S-types**	**Relative Growth of H-types**
0	0.3	4.5	2.78	1.364	0.844
1	0.35	5.32	2.90	1.415	0.772
2	0.38	6.42	3.06	1.427	0.680
3	0.43	7.79	3.26	1.388	0.580
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
40	0.99	15	4.286	1	0.286
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
∞	1	15	4.286	1.000	0.286

In this example, the short-sighted individuals always receive a payoff. Thus, even though their growth is positive, it will be dominated by the growth of foresighted people, and thus the efficient outcome will result, and the population will be virtually all foresighted.

Table 2: Example of Population Growth, if $x_0 < x^*$

Again, the payoffs are identical to those found in Figure 3 *and* $r_S = 0$, $r_H = 0.5$, and $p = 0.8$, thus $x^* = 1/6$, and the path of x_t will follow the pattern in the table below. However, if $x_0 < x^*$, then the *inefficient*, short-sighted equilibrium will result, where $x = 0$. If 10% of the population is foresighted to begin with (i.e., $x_0 = 0.1$), then the follow pattern will emerge. In this example, foresighted people receive rapidly diminishing payoffs and eventually stop growing all together. Society rapidly moves toward failure! It takes roughly 10 periods to reach the short-sighted equilibrium and the collapse of civilization.

t	x_t	**Expected Payoff for S-types**	**Expected Payoff for H-types**	**Relative Growth of S-types**	**Relative Growth of H-types**
0	0.100	1.5	2.36	0.660	1.038
1	0.083	1.24	2.32	0.558	1.040
2	0.064	0.97	2.281	0.442	1.039
3	0.047	0.70	2.243	0.322	1.033
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
10	0.0008	0.011	2.144	0.0053	1.0008
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
∞	0	0	2.143	0	1

Note that the replicator dynamic process provides the speed with which a particular population will converge to an equilibrium. This "speed" of convergence, however, depends on certain parameters. So, the initial proportion of S to H types in the population, the discount rates, and the probability the game ends all drive the speed with which an equilibrium is reached. These are issues we will address in future iterations of this paper. However, even in this preliminary version of our paper we sense how fast people decide on a particular strategy. In other words, we get a sense of how fast a civilization may collapse. Distinguishing between the effect of the parameters that drive this process may allow us to compare the speeds at which different societies may collapse. Arguably this approach can be used to predict how soon *any* social organization may collapse relative to others. Thus, for example one might use this insight to predict which hedge funds are more likely to go bankrupt in an economic downturn given their short positions.[17]

Mailath (1998) points out evolutionary models mimic models of learning. This happens for two reasons. On one hand, evolutionary models cannot really model learning since in these models the unit of analysis is not really the individual. On the other hand, the notion that successful behavior in one period increases the proportion of agents following that behavior in the next period suggests a certain naive learning. At any rate, the fact that populations converge on a Nash Equilibrium (in the sense that no one individual has an incentive to deviate from this equilibrium) suggests an awareness of at least what constitutes success and failure and the ability to mimic successful behavior. However, one can think of a richer model with adaptive individuals who not only know that they are being successful, but also why they are being successful. In this case there would indeed be reason for those with the benefit of foresight to act to increase their numbers in the population through extensive education programs that emphasize the benefit of civic engagement. In fact, one could argue that the role of public education in the US is designed for precisely this purpose. The failure of the public education system without any plausible alternate avenue for teaching and reinforcing the value of foresighted behavior may then lead to

[17]The short positions would give the observer a sense of the relative discount rates for each firm. Then, sorting these funds as foresighted or shortsighted and given the probability of the game continuing (determined by the economic environment, but same for all the firms) one could possibly decide if these firms would cooperate or not. A replicator dynamic process might then tell the observer how soon the firms would go bankrupt – or recover.

dire consequences. Again, if agents did have the ability to alter their behavior, then in a society where there is a large proportion of people with short planning horizons those with a cooperative bent would gradually change their behavior to the society's detriment until the entire population has a short planning horizon. However, in the present model, the proportion of types in the population changes one generation at a time depending on the relative payoffs, but this will yield the same result (Mailath, 1998 p. 1354-1356).

6 Areas for Expansion and Model Enrichment

In future iterations of this paper we plan on exploring the replicator dynamic process fully – particularly with respect to the effect of relative discount rates and the probability that the game might end. We also hope to flesh out the links of our model with institutional systems designed to move society toward a more efficient outcome. In this context we wish to address the relative importance of external shocks to the system. The critical shocks will alter the proportion of types in society. Such a shift would cause a succeeding society to unravel toward inefficiency or short-sighted behavior, or vice versa. The relative possibility of this will depend on the speed at which the prevailing type propagates.

A focus on the relative importance of risk dominance and payoff dominance may yield interesting results as well. For example if the payoff from the Hare type increases when they are around Stag types to the point where there may be no individual benefit to the Hare type then the game will resemble a prisoner's dilemma. Thus as these relative payoffs change the relative importance of risk dominant equilibrium to the payoff dominant equilibrium changes (see for example, Bearden, 2001). Ultimately as these relative payoffs change the game morphs into a prisoner's dilemma. In this case polymorphic ESS is possible.

Thus, exploring the social conditions specified by a prisoner's dilemma class of games relative to the assurance class of games would increase the applicability of evolutionary modeling in understanding social change.

7 Conclusion

Our innovations lie in two areas. First, we interpret a well known equilibrium phenomenon in evolutionary coordination games as social collapse. We show that the initial proportions of people with foresight, relative to those without, matter critically in determining whether society collapses or not. Second, we argue that differences in perceptions of the future and in perceptions of whether the game continues or not are important determinants of societal collapse, or success. Thus, as discount rates between the foresighted and shortsighted types diverge, the evolution of the efficient equilibrium becomes more likely. Civilization is more likely to survive as the discount rates diverge. Moreover, as the probability that life continues falls it becomes harder to achieve the efficient equilibrium. That is civilization is unlikely to survive extinction level threats even when the imminence of the threat is probabilistically determined.

References

[1] Aumann, R. (1987) "Correlated Equilibrium as an Expression of Bayesian Rationality." *Econometrica*, 55: 1-18.

[2] Bearden, J.N. (2001) "The Evolution of Inefficiency in a simulated Stag Hunt." *Behavior Research Methods, Instruments, and Computers*, 33(2): 124-129.

[3] David, Paul. (1985) "Clio and the Economics of QWERTY." *American Economic Review*, 75: 332-337.

[4] Diamond, Jared. (2005) *Collapse: How Societies Choose to Fail or Succeed*. Viking Press, New York.

[5] Dixit, Avinash; Skeath, Susan. (1999) *Games of Strategy*. W.W. Norton & Company, New York.

[6] Emmott, R. (2006) "Museum Explores Peru's Shining Path." *Washington Post*, http://www.washingtonpost.com/wp-dyn/content/article/2006/10/21/AR2006102100652.html (accessed 14 November 2008).

[7] Gardner, R. (2003) *Games for Business and Economics*. John Wiley and Sons, Hoboken, N.J.

[8] Gintis, H. (2000) *Game Theory Evolving - A Problem Centered Introduction to Modeling Strategic Interaction*. Princeton University Press, Princeton, N.J.

[9] Harsanyi, J.C.; Selten, R. (1988). *A General Theory of Equilibrium in Games*. MIT Press, Cambridge, MA.

[10] Mailath, G.J. (1998) "Do People Play Nash Equilibrium? Lessons From Evolutionary Game Theory." *Journal of Economic Literature*, 36: 1347-1374.

[11] North, Douglas. (2005) *Understanding the process of economic change*. Princeton University Press, Princeton, NJ.

[12] Poundstone, W. (1992). *Prisoner's Dilemma*. Doubleday, New York.

[13] Rousseau, J.J. (1984) *A Discourse on Inequality*. Translated by Maurice Cranston. Penguin, London.

[14] Selten, R. (1975) "Reexamination of the Perfectness Concept For Equilibrium Points in Extensive Games." *International Journal of Game Theory*, 4: 25-55.

[15] Tuchman, B. (1984) *The March of Folly: From Troy to Vietnam*. Ballantine: New York.

[16] Weibull, J. W. (1997). *Evolutionary Game Theory*. MIT Press, Cambridge, MA.

[17] Wilson, E. (1998). *Consilience*. Knopf: New York.

Appendix A: *Proof of Theorem 1*

The expected payoff from the Stag (S) strategy is

$$E(S) = 3x\frac{1+r_S}{1+r_S-p} \tag{1}$$

and from Hare (H) strategy is

$$E(H) = 2x\frac{1+r_H}{1+r_H-p} + (1-x)\frac{1+r_H}{1+r_H-p} \tag{2}$$

The H and S strategy provide the same expected payoff when $E(H) = E(S)$ i.e.,

$$2x\frac{1+r_H}{1+r_H-p} + (1-x)\frac{1+r_H}{1+r_H-p} = 3x\frac{1+r_S}{1+r_S-p}$$

Or

$$3\left(\frac{1+r_S}{1+r_H}\right)\left(\frac{1+r_H-p}{1+r_S-p}\right) = 1+\frac{1}{x} \tag{3}$$

Now we have defined $\delta_S = \frac{p}{1+r_S}$ and $\delta_H = \frac{p}{1+r_H}$. Thus,

$$\frac{1+r_S}{1+r_H} = \frac{\delta_H}{\delta_S} \tag{4}$$

And

$$\frac{1+r_H-p}{1+r_S-p} = \frac{\frac{1}{\delta_H}-1}{\frac{1}{\delta_S}-1} \tag{5}$$

Substituting (4) and (5) into (3) gives us

$$3\left(\frac{\delta_H}{\delta_S}\right)\left(\frac{\frac{1}{\delta_H}-1}{\frac{1}{\delta_S}-1}\right) = 1+\frac{1}{x}$$

Or

$$x = \frac{1}{3\left(\frac{1-\delta_H}{1-\delta_S}\right)-1} \tag{6}$$

Thus for (H) to be preferred over (S)

$$x < \frac{1}{3\left(\frac{1-\delta_H}{1-\delta_S}\right)-1} \tag{7}$$

Appendix B: *Population Growth*

The success or failure of society depends on the initial proportion of foresighted people: $x_0 = S_t/H_t$ when $t = 0$. If $x_0 < x^*$, then the growth of Hare types will dominate, and society will eventually reach the inefficient equilibrium, $x = 0$, where short-sighted people account for approximately 100% of the population. If $x_0 > x^*$, then the growth of Stag types will dominate, and society will eventually reach the efficient equilibrium, $x = 1$, where foresighted people account for approximately 100% of the population.

Growth in x proceeds as follows:

$$x_t = \frac{S_t}{S_t + H_t}, \quad S_{t+1} = S_t \cdot (1 + \text{growth}\, S_t), \quad H_{t+1} = H_t \cdot (1 + \text{growth}\, H_t)$$

and

$$\text{growth}\, S_t = \frac{E(S_t)}{x_t E(S_t) + (1 - x_t)E(H_t)} \quad \text{growth}\, H_t = \frac{E(H_t)}{x_t E(S_t) + (1 - x_t)E(H_t)}$$

where the expected values are based on the payoffs given in Figure 3:

$$\begin{aligned} E(S_t) &= 3x_t \left(\frac{1}{1-p}\right) \\ E(H_t) &= 2x_t \left(\frac{1 + r_H}{1 - p + r_H}\right) + (1 - x_t)\left(\frac{1 + r_H}{1 - p + r_H}\right) = (1 + x_t)\left(\frac{1 + r_H}{1 - p + r_H}\right) \end{aligned}$$

To simplify, set $\beta_S = \left(\frac{1}{1-p}\right)$ and $\beta_H = \left(\frac{1 + r_H}{1 - p + r_H}\right)$.

$$\begin{aligned} E(S_t) &= 3\beta_S x_t \\ E(H_t) &= \beta_H (1 + x_t) \end{aligned}$$

Thus, growth in x can be written as a system of finite difference equations:

$$\begin{aligned} x_{t+1} &= \frac{S_t \left(1 + \frac{3\beta_S x_t}{3\beta_S (x_t)^2 + \beta_H (1 - (x_t)^2)}\right)}{S_t \left(1 + \frac{3\beta_S x_t}{3\beta_S (x_t)^2 + \beta_H (1 - (x_t)^2)}\right) + H_t \left(1 + \frac{\beta_H (1 + x_t)}{3\beta_S (x_t)^2 + \beta_H (1 - (x_t)^2)}\right)} \\ S_{t+1} &= S_t \left(1 + \frac{3\beta_S x_t}{3\beta_S (x_t)^2 + \beta_H (1 - (x_t)^2)}\right) \\ H_{t+1} &= H_t \left(1 + \frac{\beta_H (1 + x_t)}{3\beta_S (x_t)^2 + \beta_H (1 - (x_t)^2)}\right) \end{aligned}$$

Appendix C: *Proofs of Theorems 2 & 3 and their Corollaries.*

Theorem 2 states that x^* is decreasing in r_H. Theorem 3 states that x^* is decreasing in p. To demonstrate these, we must show that the derivatives of x^* with respect to r_H and p are alway negative. We can re-write x^* such that

$$\begin{aligned}
x^* &= \left(\left(\frac{1-\dfrac{p}{1+r_H}}{1-\dfrac{p}{1+r_S}}\right)-1\right)^{-1} \\
&= \left(3\left(\frac{\dfrac{1+r_H-p}{1+r_H}}{\dfrac{1+r_S-p}{1+r_S}}\right)-1\right)^{-1} \\
&= \left(3\left(\frac{(1+r_H-p)(1-r_S)}{(1+r_H)(1+r_S-p)}\right)-1\right)^{-1} \\
&= \frac{(1+r_H)(1+r_S-p)}{3(1+r_H-p)(1-r_S)-(1+r_H)(1+r_S-p)}
\end{aligned}$$

or

$$x^* = \frac{(1+r_H)(1+r_S-p)}{3(1+r_H-p)(1-r_S)-(1+r_H)(1+r_S-p)}$$

Equivalently,

$$x^* = \frac{1+r_H+r_S+r_Sr_H-p-pr_H}{2-2p+r_S(2-3p)+r_H(2+2r_S+p)}$$

Next, separately differentiate with respect to r_H and p:

$$\begin{aligned}
\frac{dx^*}{dr_H} &= \frac{-3p(1+r_S)(1+r_S-p)}{(2-2p+r_S(2-3p)+r_H(2+2r_S+p))^2} \\
\frac{dx^*}{dp} &= \frac{-3(1+r_H)(r_H-r_S)(1+r_S)}{(2-2p+r_S(2-3p)+r_H(2+2r_S+p))^2}
\end{aligned}$$

Now, we assume *without loss of generality*, that $r_S = 0$, thus:

$$\begin{aligned}
\frac{dx^*}{dp} &= \frac{-3r_H(1+r_H)}{(2-2p+r_H(2+p))^2} \\
\frac{dx^*}{dr_H} &= \frac{-3p(1-p)}{(2-2p+r_H(2+p))^2}
\end{aligned}$$

This indicates that whenever either p or r_H increases, x^* decreases since p is strictly between zero and one, and r_H is strictly positive.

Corollaries 2 and 3 are derived from the second derivatives of x^*. These are given below; however, for simplicity, we continue to analyze the case where $r_S = 0$. We begin by rewriting the first derivatives as

$$\begin{aligned}
\frac{dx^*}{dp} &= -3r_H(1+r_H)(2-2p+r_H(2+p))^{-2} \\
\frac{dx^*}{dr_H} &= -3p(1-p)(2-2p+r_H(2+p))^{-2}
\end{aligned}$$

This gives

$$\begin{aligned}
\frac{d^2x^*}{dp^2} &= -3r_H(1+r_H)[-2(2-2p+r_H(2+p))^{-3}(-2+r_H)] \\
&= \frac{-3r_H(1+r_H)(-2)(-2+r_H)}{(2-2p+r_H(2+p))^3} \\
&= \frac{6r_H(1+r_H)(r_H-2)}{(2-2p+r_H(2+p))^3} \\
\frac{d^2x^*}{d(r_H)^2} &= -3p(1-p)[-2(2-2p+r_H(2+p))^{-3}(2+p)] \\
&= \frac{-3p(1-p)(-2)(2+p)}{(2-2p+r_H(2+p))^3} \\
&= \frac{6p(1-p)(2+p)}{(2-2p+r_H(2+p))^3}
\end{aligned}$$

These indicate that $\dfrac{d^2x^*}{d(r_H)^2}$ is always positive and

$$\begin{aligned}
\frac{d^2x^*}{dp^2} &< 0, \qquad r_H < 2 \\
\frac{d^2x^*}{dp^2} &= 0, \qquad r_H = 2 \\
\frac{d^2x^*}{dp^2} &> 0, \qquad r_H > 2
\end{aligned}$$