The Evolution of Inequality

Samuel Bowles

SFI WORKING PAPER: 2001-10-060

The Evolution of Inequality

Samuel Bowles
Santa Fe Institute and University of Massachusetts
July 7, 2001

Abstract

Under what conditions can class divisions characterized by high levels of inequality be designated *evolutionary universals*, using Talcott Parsons's term to refer to social arrangements which have emerged independently and persisted in a wide variety of environments? To explore this question, I represent economic institutions as bargaining conventions and then in order to better capture the historically observed processes of institutional evolution I extend recent models in stochastic evolutionary game theory in four ways: i) non-best response (idiosyncratic) play is modeled as intentional rather than accidental, ii) non best response play is coordinated through a process of collective action, iii) substantial rates of non-best response play are introduced, and iv) the sub-populations making up the classes may be of different sizes. In this model, contrary to the conventional formulation, highly unequal and economically inefficient institutions may be stochastically stable states in the implied dynamical system, while more egalitarian institutions may prove ephemeral.

1. Introduction[1]

Hernán Cortés' long letters to King Charles of Castile describe the exotic and unusual customs he and his armed band encountered as they advanced toward Temixtitan in 1519. But in light of the thirteen millennia which had passed since there had been any sustained contact between people of the Old World and the New, what is striking about his account of Mexico is how familiar it all was. Upon reaching Temixtitan, he wrote:

> *There are many chiefs, all of whom reside in this city, and the country towns contain peasants who are vassals of these lords and each of whom holds his land independently; some have more than others.. And there are many poor people who beg from the rich in the streets as the poor do in Spain and in other civilized places.* (Cortés, (1986):68, 75)

He continues, describing (105) the "many temples or houses for their idols" and their extensive markets with well developed regulation of weights and measures: "It seems like the silk market at Granada, except that there is a much greater quantity." (104) He remarks (68) that "the orderly manner which, until now, these people have been governed is almost like that of the states of Venice or Genoa or Pisa."[2]

Some types of social arrangements -- social ranking, markets, states, monogamy, private property, worshiping supernatural beings, and sharing the necessities of life among non-kin, for example -- have been ubiquitous over long periods of human history and have emerged and persisted in highly varied environments. Others have been of passing importance, generally occupying limited ecological niches. Speculation about the characteristics contributing to the evolutionary success of these prevalent institutional forms has occupied great minds since David Hume, including Alexis De Tocqueville, Karl Marx, Karl Menger, and Frederich Hayek.

Some, who like Cortés are impressed by the similarity of institutions in quite differing environments, have postulated a coherent set of "modern" social arrangements towards which most independent societal trajectories are said to be tending. Talcott Parsons (1964) termed

---

[1] For presentation to the Poverty Traps workshop at the Santa Fe Institute, July 20-22, 2001, and eventual publication in *Poverty Traps,* edited by S.Bowles, S. Durlauf, and K. Hoff (Russell Sage Foundation, 2002). Thanks to Ugo Pagano, Yong-jin Park, Elisabeth Wood , Jorgen Weibull and Peyton Young for contributions to this paper, and to the MacArthur Foundation for financial support. Http://www-unix.oit.umass.edu/~bowles/ Correspondence to bowles@econs.umass.edu or bowles@santafe.edu.

[2] Robert Boyd drew my attention to this example of evolutionary convergence. The colonization of the Americas from Asia took place about 12,000 B.C.

these *evolutionary universals*, namely those ways of ordering society which crop up with sufficient frequency in a variety of circumstances to suggest their general evolutionary viability. Parsons offered vision as a biological analogy and identified money and markets, bureauracacy, stratification, and democracy as human social examples. Frederich Hayek refers to the markets and private property nexus -- his “extended order” -- in a similar vein. Many attribute the evolutionary success of these institutions to their societal efficiency: Douglass North writes: .."competition in the face of ubiquitous scarcity dictates that the more efficient institutions will survive and the inefficient ones perish."(1981), see also Williamson (1985). Marx's conception of the historical succession of institutions under the influence of changing technology similarly posits a tendency -- albeit a very long term one-- for efficient institutions to prevail.

Others have stressed the fundamentally path-dependent evolution of social structure, with distinct societal histories emerging as the result of initially small differences or chance events. In contrast to the above accounts of institutional convergence, this view stresses the long term coexistence of distinct evolutionarily stable institutions. Mexico provides a telling example of this pattern of institutional divergence as well. Cortés’ service to the crown was rewarded with the title Marqués del Valle de Oaxaca. Modern-day researchers in the Valley are puzzled by the juxtaposition of villages with extraordinarily high homicide rates with others in which homicide is virtually unknown. These villages do not differ in any of the commonly ascribed causes of violence such as alcohol use, boundary quarrels, crowding and political competition; some however are distinguished by long established traditions of anti-violence coupled with the absence of social rank and the rotation of village offices. (Paddock (1991), (1975), Greenberg (1989)).

These institutional and behavioral differences among the villages in the Valley of Oaxaca along with the familiarity of the institutions that Cortés encountered in Temixtitan pose one of the questions to which this chapter is addressed: what are the general characteristics (if any) of evolutionarily successful institutions? To provide an answer we will need an understanding of the birth, diffusion and eclipse of institutions and the process by which one institution supplants another. This will require an account of how characteristics of institutions contribute to their evolutionary success. Remarkably, many treatments are silent on this crucial issue: North for example does not explain how institutions “compete” to weed out inefficient arrangements and Parsons provides no mechanism by which the institutions he identifies might become “universal.”

There are two fundamental processes which might provide an account of the differential success of distinct institutional arrangements. The first, which Hayek believes accounts for the success of his “extended order,” is competition among groups governed by differing institutions; successful institutions are those which contribute to the success of nations, bands, ethno-linguistic units and other groups in competition with other groups.[3] The second process, emphasized by Marx in his treatment of the relationship of technical progress to institutional change, is internal to groups, and can take the form (as with Marx) of conflict among interdependent actors who are differentially benefited by one or another institutional form or of the piecemeal accommodation of institutions to new demands and opportunities. Here I analyze within-group processes of institutional evolution.

This will require an account of how individual intentions and chance events combine with environmental constraints to make history. This is a demanding task, which I will attempt here by focusing on a class of institutions which take the form of conventions. Because of their historical importance I will focus on economic institutions which regulate the size of the social surplus and its distribution. I will ask if there is any reason to expect that the most persistent conventions – those resembling Parsons' evolutionary universals – would be characterized by relatively equal or unequal divisions of the surplus.

An institution may be represented as one of a number of possible conventional equilibria in which members of a population typically act in ways that are best responses to the actions taken by others and have formed expectations that support continued adherence to these conventional actions. Examples include simple principles of division such as "finders keepers" or "first come first served," as well as more complicated principles of allocation such as the variety of rules which have governed the exchange of goods or the division of the products of one's labor over the course of human evolution. Because a convention is one of many possible mutual best responses defined by the underlying game, institutions are not environmentally determined, but rather are of human construction (but not necessarily of deliberate design). Thus it is of some interest to ask which of these will obtain, and why one convention might in the course of time be eclipsed by another. Game theorists refer to this as the problem of equilibrium selection; historians call it the problem of structural change, institutional innovation, or (in some settings) social revolution.

Two quite distinct approaches to the within group processes bringing about institutional innovation may be identified.

The first, similar to Sewall Wright's use of drift to explain a movement from one fitness peak across a fitness valley to another peak, is that proposed by stochastic evolutionary game theory (Young (1993), Young (1998), Kandori, Mailath, and Rob (1993)). In this Darwin-inspired approach change occurs through the chance bunching of individuals' idiosyncratic play of non-best-responses sufficient to tip the underlying dynamic process from the basin of attraction of one conventional equilibrium to another. Changes in language use, contractual

[3] I investigate the evolutionary success of egalitarian institutions in a multi level selection framework in Bowles (2002) and Bowles and Hopfensitz (2001)

shares, market days, and etiquette have been modeled in this manner (Young (1995)).

The second approach, initiated by Marx, stresses asymmetries among the players and explains institutional innovation by the changing power balance between those who benefit from differing conventions.[4] In this framework, revolutionary change in institutions is likely when existing institutions facilitate the collective action of those who would benefit from a change in institutions, and when, because existing institutions are inefficient by comparison to an alternative, there are substantial potential gains to making a switch. This collective-action-based approach has been used to model conflicts among classes resulting in a basic transformation of social organization, such as the French, Russian, and Cuban revolutions (Soboul (1974), Lefebvre (1947), Trotsky (1932), Zeitlin (1967)) as well more gradual changes in institutional arrangements such as the centuries long erosion of European feudalism (Brenner (1976)).

Do these approaches allow us to say anything about the characteristics of evolutionarily successful institutions? Though the underlying causal mechanisms are different, the Marx-inspired approach shares with Darwin-inspired stochastic evolutionary game theory the result that institutional arrangements which are both inefficient and unequal will bear an evolutionary disability and will tend to be displaced in the long run by more efficient and more egalitarian institutions.[5] This is quite an arresting claim in light of the long term historical persistence of social arrangements which would appear to be neither egalitarian nor efficient. I will explore this proposition as a way of both introducing and extending the stochastic evolutionary game theoretic approach.

I begin in the next section with a simple non-stochastic population game in which the stage game exhibits two conventional equilibria, using this game to illustrate the determinants of institutional persistence and accessibility based on the familiar concept of risk dominance. Next I introduce the stochastic evolutionary approach and explain how it yields a rather strong characterization of successful institutions akin to Parsons' evolutionary universals: these are institutions which define *stochastically stable states*. I also give some reasons why the application of stochastic evolutionary game theory to real historical evolutions may be somewhat limited. I then augment the stochastic framework by introducing asymmetries between the players who intentionally pursue conflicting interests through collective action. I provide numerical examples to explore the status of equal and efficient conventions as stochastically stable states. I conclude that for an important class of institutional innovations,

---

[4] Marx ((1859)1976). For a modern exposition, Cohen (1978).

[5] Efficient conventions yield a larger joint surplus, while in a more equal convention the share of the least well off is larger.

the purely stochastic and symmetrical version of the model gives a misleading account. The reason is that the dynamics supported by accidental and intentional action are not the same, and that by introducing intentional action in pursuit of common interests, insights more consistent with the above empirical examples are generated.

2. The persistence and accessibility of historically contingent institutions

Because nothing of importance is lost in taking an especially simple case, I confine myself to the analysis of the evolutionary dynamics governing transitions between two conventions in a 2x2 game in a large population of individuals subdivided into two groups, the members of which are randomly paired to interact in a non-cooperative game with members of the other group. Individuals' best response play is based on a single period memory, and they maximize their expected payoffs based on the distribution of the population in the previous period.

The two population subgroups of equal size (normalized to unity) are termed A's and B's, and each when paired with a member of the other group may chose action 1 or 0, with the A's payoffs, a_{ij} representing the payoff to an A-person playing action i against a B-person playing action j, and analogously for the B's. Because these payoffs conform to the von Neumann-Morgernsern axioms we can assume that the players will act in ways that maximize their expected payoffs, given beliefs about what the other players will do.

Suppose that if the members of the pair choose the same action they get positive benefits, while if they chose different actions they get nothing; the sub-groups may be economic classes selecting a contract to regulate their joint production, which will only take place if they agree on a contract, with the no-production outcome normalized to zero for both. Both A's and B's may choose action 1 or action 0. The payoffs, with the A's as row and the B's as column player, are thus:

Figure 1: Payoffs in the Contract Game

	B offer contract 1	B offer contract 0
A offer contract 1	a_{11}, b_{11}	0,0
A offer contract 0	0,0	a_{00}, b_{00}

To capture the conflict of interest between the two groups, assume that $b_{00} > b_{11} = a_{11} > a_{00} > 0$ so the B's strictly prefer the outcome in which both play 0, the A's prefer the equal division

outcome which results when both play 1.[6] Both of these outcomes are strict Nash equilibria, and thus both represent conventions, which I will denote E_0
and E_1 (or {0,0} and {1,1}.)

We may describe the state of this population in any time period t by $\{\alpha_t,\beta_t\}$, where α is the fraction of the A's who played 1 in the previous period and β is the fraction of the B's who played 1, with the two conventions being defined by the states {1,1} and {0,0}. Both populations are normalized to unit size, so I refer indifferently to the numbers of players and fraction of the population, abstracting from integer problems. For any state of the population, expected payoffs a_i and b_i from A's and B's respectively playing strategy i, clearly depend on the distribution of play among the opposing group:

$$a_1 = \beta a_{11};\ a_0 = (1-\beta)a_{00};\ b_1 = \alpha b_{11};\ \text{and } b_0 = (1-\alpha)b_{00}.$$

The relationship between the population state and the expected payoffs to each action is illustrated in figure 2.

Individuals take a given action -- they are 1-players or 0-players -- and they continue doing so from period to period until they update their action, at which point they may switch. Suppose that at the beginning of every period some fraction ω of each sub population may update their actions (this might be due to the age structure of the population, with updating taking place only at a given period of life, in which case the "periods" in the model may be understood as "generations"; of course updating could be much more frequent).[7]

[6] I refer to {1,1} as the "equal" convention as a shorthand. The levels of well-being attained by the A's and B's cannot be determined without additional information (if A's interact with only one B, while B's interact with many A's, the "equal" convention would exhibit unequal incomes of the two groups, for example).

[7] Giving individuals a longer (than one period) memory, or a less naive updating rule, or a more limited knowledge of the distribution of types in the other subpopulation, would not yield substantially different insights about the questions explored here. The overlapping generations assumption concerning updating means that the stochastic shocks due to idiosyncratic play (to be introduced presently) are persistent as the realized distribution of play in the previous period reflects the shocks experienced over many past periods.

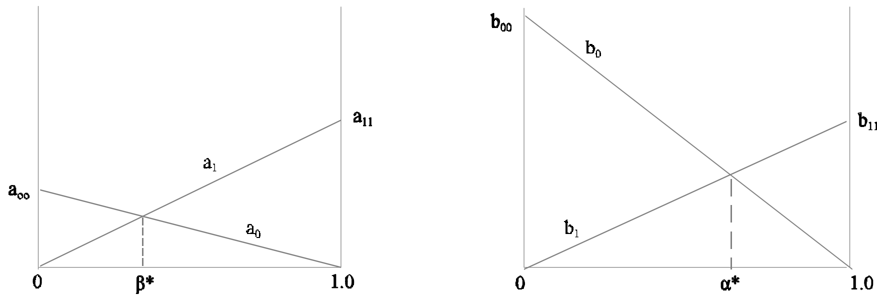

Figure 2 Expected payoffs. Note: A's payoffs depend on β the fraction of B's offering contract 1, while the B's payoffs depend on α the fraction of A's offering contract 1. Because $b_{00} > b_{11} = a_{11} > a_{00}$, the convention E_1 (that is, α=1=β) is preferred by the A's while E_0 is preferred by the B's.

The updating is based on the expected payoffs to the two actions; these expectations are simply the payoffs which would obtain if the previous period's state remained unchanged (the population composition in the previous period being common knowledge in the current period.) While the updating process is not very sophisticated, it may realistically reflect individuals' cognitive capacities and it assures that in equilibrium -- when the population state is stationary -- the beliefs of the actors formed in this naive process are confirmed in practice.

Individuals are are represented simply bearers of the strategies they have adopted, their number is unchanging, while the distribution of strategies among them varies. I will analyze the single period change in the population state $(\Delta\alpha,\Delta\beta)$ under the assumption that individual updating of strategies is monotonic in average payoffs so that $\Delta\alpha$ and $\Delta\beta$ have the signs respectively, of $(a_1 - a_0)$ and $(b_1 - b_0)$. The resulting population dynamics are illustrated in figure 3, where signs of $\Delta\alpha$ and $\Delta\beta$ in the relevant regions are defined by:

(1) $\alpha^* = b_{00}/(b_{11}+b_{00})$

$\beta^* = a_{00}/(a_{11}+a_{00})$

these population distributions equating the expected payoffs to the two strategies for the two sub-populations, respectively. These values of α and β define best response functions: for $\alpha<\alpha^*$ B's best response is to play 0, and for $\alpha \geq \alpha^*$ B's best response is to play 1, with β^* interpreted analogously.

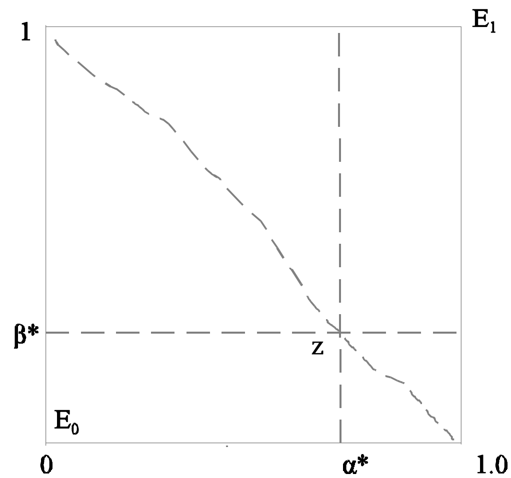

Figure 3. The state space. Note: E_o and E_1 are absorbing states in the non stochastic dynamic given by (2) and (3); z is a saddle point.

For states α<α* and β<β* (in the southwest region of figure 3) it is obvious that Δα and Δβ are both negative and the state will move to {0,0}. Analogous reasoning holds for the north east region. In the north west and south east regions of the state space we may define a locus of states from which the system will transit to the interior equilibrium α*,β* with states below that locus transiting to {0,0}, and above the locus to {1,1}. The basin of attraction of {0,0}, is the area below the the dotted downward sloping line in figure 3; its size will vary with α*β*.

While the interior equilibrium {α*,β*} is an unstable Nash equilibrium (a saddle, as can be confirmed by inspecting the vector field), the outcomes {0,0} and {1,1} are absorbing states of the dynamic process, meaning that if the population is ever at either of these outcomes it will never leave. There being more than one such absorbing state, the dynamic process is non-ergodic and the institutional outcome is thus historically contingent; the system is deterministic once the initial conditions are given.

Because best response play renders both conventions absorbing states, it is clear that in order to understand institutional change, some kind of non-best -response play must be introduced. The basic idea is that if the status quo convention is {0,0} but a sufficiently large number of A's play 1 for some reason not captured by the model thus far (hence called idiosyncratic) then in the next period the best response of the B's, having encountering these 1-playing A's will be to play 1 as well, the best response of the A's encountering these 1-playing B's will be to play 1, and so on, leading to the "tipping" of the population from the {0.0} to the {1,1} convention. I will presently introduce two distinct approaches to modeling idiosyncratic responses -- chance and collective action. But before doing this, notice that we already have enough information to say something about which convention might be more persistent, namely the one which requires relatively more idiosyncratic play to dislodge.

To understand the extent of vulnerability or robustness of a convention, define the *risk factor of a conventional equilibrium* , E_j , as the smallest probability γ_j, such that if an individual believes that his game partner will play j with probability greater than γ_j, then j is the unique best response to adopt. Suppose the unequal convention {0,0} obtains -- both A's and B's playing 0. If, for whatever reason, some number β^* or greater of the B's were to play 1 rather than 0, then in the next period, playing 0 would no longer maximize the A's expected payoffs, and those who were updating would shift to playing 1. Similarly if α^* or more of the A's were to play 1 rather than 0, some of the best responding B's would switch from 0 to 1. Reading $\wedge$ as "the minimum of," the risk factor of the conventions {0,0} and {1,1} are thus

(2) $\gamma_0 = (1-\alpha^*) \wedge (1-\beta^*)$

$\gamma_1 = \alpha^* \wedge \beta^*$

Using (4) and $b_{00} > b_{11} = a_{11} > a_{00}$ we have

(3) $\gamma_0 = (1-\alpha^*) = 1-\{b_{00}/(b_{11}+b_{00})\} = \{b_{11}/(b_{11}+b_{00})$

$\gamma_1 = \beta^* = a_{00}/(a_{11}+a_{00})$

The equilibrium with the lowest risk factor is the *risk dominant equilibrium*.[8] Rearranging the above expressions, the convention {0,0} will be risk dominant if

(4) $a_{00}b_{00} > a_{11}b_{11}$

or, what is equivalent $\alpha^*\beta^* > (1-\alpha^*)(1-\beta^*)$.

A related measure of persistence of an equilibrium E_j is r_{jk} the *reduced resistance* on the path from E_j to E_k defined as the minimal number of individuals in a population adhering to the convention E_j who, should they idiosyncratically switch their strategy to k, would induce their best responding partners to switch theirs and thus induce a subsequent movement to the other convention {k,k} without further idiosyncratic play. Thus r_{01} the resistance on the path from E_0 to E_1 is $\beta^* = \gamma_1$ and in general in a 2x2 game $r_{jk} = \gamma_k$.

[8] This is true of 2x2 games. Risk dominance may be undefinable in 3x3 and other games.

Considered in terms of historical processes, a risk dominant equilibrium has two characteristics. First, it is *resistent* to perturbation: once at or near the equilibrium it takes a substantial amount of non-best-response play to displace it. Resistence to perturbation is analogous to the simpler concept of evolutionary stability or non-invadeability introduced by Maynard Smith and Price (1973), with r_{jk} representing what Weibull (1995) termed *the invasion barrier* or the minimum number of mutant k-players who would proliferate if introduced into a population of j players. The process of intentional institutional change can be modeled by interpreting r_{jk} as a "collective action barrier" namely the extent of coordinated non-best-response play suficient to induce a movement to the other convention without further non-best-response play. Second, a risk dominant equilibrium is *accessible*: the *other* convention is *not* resistant, and it does not require much bunching of non-best-response play to displace the population state into the basin of attraction of the risk dominant equilibrium (accessibility is analogous to the simpler concept of capacity to invade -- called *viability* by Axelrod and Hamilton (1981).) For the equilibrium {0,0}, β^* measures its resistance to perturbation, and α^* measures its accessibility. The fact that in the 2x2 coordination game structure, the accessibility of a convention is just one minus the resistence of the other will be important below.

3. Chance and change

How, then, might institutional change occur? The answer suggested by stochastic evolutionary game theory is that the accumulation of chance actions that are not best responses by individuals may displace the population state from one convention, moving it into the basin of attraction of the other convention. Suppose there is a probability ε that when individuals are in the process of updating, each may switch type for idiosyncratic reasons with $(1-\varepsilon)$ thus representing the probability that the individual pursues the best response updating process described above The idiosyncratic play accounting for non-best-responses need not be irrational or odd; it simply represents actions whose reasons are not explicitly modeled in the payoff matrix. Included is experimentation, whim, error, and as we will see in the extension of this approach in the next section, intentional acts seeking to affect game outcomes but whose motivations are not captured by the above game.

The important result is that the presence of idiosyncratic play transforms the dynamical system described above from a non-ergodic one (in which both {0,0} and {1,1} are absorbing states, and which is the eventual outcome is determined by the initial conditions) to an ergodic process with no absorbing states. The simplest case arises when $\omega = 1$ (everyone updates in every period). Then the Markov process described by the model yields a strictly positive transition matrix, indicating that from any state the system may transit to every other state with positive probability. To see that this is true, suppose all members of both subpopulations are "selected" for idiosyncratic play and note that any distribution of their responses is possible, thus giving positive weight to the probability of moving to any state, irrespective of the originating state.[9]

Two implications for the history of social interactions follow. First, the population state is perpetually in motion, or at least susceptible to movement. And as a result of this, second, after a sufficiently long period the state of the population is determined independently of the initial conditions, but the population state is path dependent (where it was in the recent past influences where it will be at any moment). History matters, and it never ends.

The fact that the population state is perpetually changing does not mean, of course, that all states are equally likely; the long run average behavior of the system can be studied. Stochastic evolutionary game theory investigates which conventions are relatively susceptible to disruption by idiosyncratic play, and which are robust. It does this by describing a dynamical system of the above type as a discrete time Markov process specifying the probability in any period of transiting from the current state to each of a finite number of other states. The system, of course, will spend more time in some states than others; this is because, as the discussion of accessibility and persistence in the previous section suggests, some states are easier to get to and harder to leave, and these states therefore will occur more frequently than others. These are the *stochastically stable states,* defined as those will occur with non-negligible probability even when the rate of idiosyncratic play is arbitrarily small.

A major result of stochastic evolutionary game theory is that for 2x2 coordination games of the type studied here, and under some (innocuous) restrictions on the updating process, the stochastically stable state is the risk dominant convention (Young, (1998)) An equivalent statement is that E_j is the stochastically stable state if the resistance along the path leading to this equilibrium is less than the resistance on the path from it to the other equilibrium or $r_{ij} < r_{ji}$. Thus over a sufficiently long historical period the population will most frequently be at or near the convention with the largest basin of attraction, the reason being that chance events are more likely to tip the population into this basin of attraction, and less likely to tip it out, making the convention in question both accessible and resistant to perturbations. If a conclusion is to be drawn it is that among the institutions we observe historically we would

[9] Where ω<1 transition between states will take longer, but the above intuition is nonetheless correct, because if in every period any distribution of play among the potential innovators is possible, then in a sufficiently long period of time any distribution of play among the entire population is also possible.

expect the preponderance of them to be stochastically stable states.[10]

What do we know about the characteristics of stochastically stable states? A striking theorem proved by Young (1998) is that the institutions supporting these stochastically stable risk dominant equilibria are characterized by both relative efficiency and equality (for the 2x2 games studied here the stochastically stable states are egalitarian in the maximin sense: they maximize the payoff of the least well of group.[11] Analysis of the 2x2 contract game will be facilitated if we write $a_{00}+b_{00} = \rho$, so recalling that in the {1,1} convention two individual payoffs are normalized to unity their sum is 2 so $\rho/2$ is a measure of the relative efficiency of the {0,0} convention; when ρ takes the value of 2 the two conventions produce the same the joint surplus. Further let the A player's share of joint surplus in the B-favoring {0,0} equilibria be $\sigma<½$, with $(1-\sigma)$ the share of gained by B.

Figure 4: Modified Payoffs in the Contract Game

$a_{11}=1$ $b_{11}=1$	0,0
0,0	$a_{00}=\sigma\rho$ $b_{00}=(1-\sigma)\rho$

Note: $\rho/2$ and σ respectively measure the efficiency and equality of {0,0} relative to {1,1}.

The convention {0,0} will, as we have seen, be the risk dominant equilibrium and hence stochastically stable state if $\alpha^*\beta^* > (1-\alpha^*)(1-\beta^*)$ which using the payoffs in figure 4 requires that

(5) $\sigma(1-\sigma)\rho^2 > 1$

It is clear from this condition that both relative efficiency and equality of shares contribute to stochastic stability of a convention (the term $\sigma(1-\sigma)$ is maximized for $\sigma = ½$.) Figure 5 illustrates the relationship between efficiency and equality as determinants of stochastic stability: YY is the locus of combinations of ρ and σ such that $\sigma(1-\sigma)\rho^2 =1$ and which thus

[10] Young, ((1998), theorem 4.1). In the updating model on which this theorem is based (and the one mentioned in the next footnote) agents have a memory of m periods, and sample (s<m) from their memory to form expectations. Results concerning stochastic stability generalize beyond the 2x2 coordination games and apply even where risk dominance cannot be defined.

[11] Young ((1998), theorem 9.1).

equate the risk factor of {0,0}to the risk factor of the the egalitarian convention {1,1} (for which ρ=2 and σ=½).

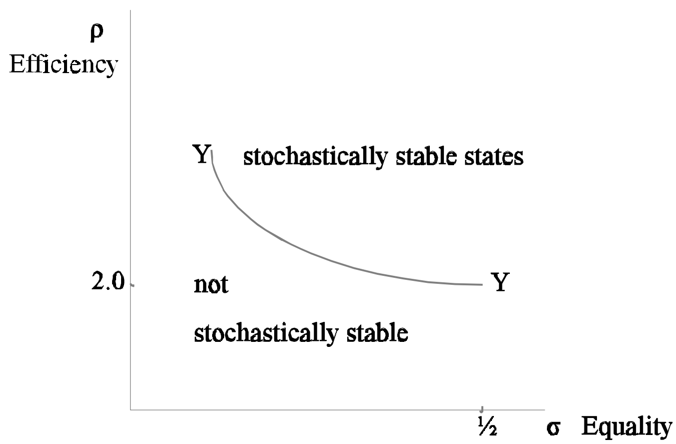

Figure 5 An efficiency equality tradeoff. Note: σ and ρ measure the equality and efficiency, respectively of E_0 relative to E_1 (for which ρ=2, σ=½). Conventions with contracts above the YY locus are stochastically stable when the alternative is convention E_1.

It is easy to see why efficient conventions would be favored in this setup: for at least one group the expected payoffs will be maximized by acting in accordance with the contract the efficient convention prescribes even if one believes that the probability that ones partner will doing likewise is less than one half. Inefficient conventions are neither persistent nor accessible because it takes a large amount of non-best-response play to induce best responders to shift from an efficient to an inefficient convention. Note that this is not because best responders anticipate the consequences of their switching for the population dynamics. Rather, their response is purely individual and based on past (not anticipated future) population states; no individual is attempting to implement the more efficient convention.

Less transparent is the result that highly unequal conventions are not good candidates for stochastic stability. This is a consequence of the fact that they are easily unravelled, because as Young ((1998):137) puts it: "it does not take many stochastic shocks to create an environment in which members of the dissatisfied group prefer to try something different." Notice in this example it is the idiosyncratic play of the privileged group that unravels the unequal convention. To see why the processes of transition between the two conventions depends on the share of the less well off in the unequal convention we can use (3) and figure

4 to get the following expressions for the risk factors fo the two equilibria.

$$\gamma_0 = (1-\alpha^*) = 1-\{b_{00}/(b_{11}+b_{00})\} = \{b_{11}/(b_{11}+b_{00}) = 1/\{1+(1-\sigma)\rho)$$

$$\gamma_1 = \beta^* = a_{00}/(a_{11}+a_{00}) = \sigma\rho/(1+\sigma\rho)$$

The effect of greater equality of shares in the unequal state on the risk factors of the two equilibria is given by

$$d\gamma_0/d\sigma = \rho/(1+\rho-\sigma\rho)^2 > 0$$

$$d\gamma_1/d\sigma = \rho/(1+\sigma\rho)^2 > 0$$

from which it is clear that greater equality of the unequal convention raises the risk factor of both conventions . But a comparison of the above expression shows that $d\gamma_1/d\sigma > d\gamma_0/d\sigma$, so that while increased *in*equality of {0,0} stabilizes both conventions, it reduces the risk factor of the more equal convention by more than the unequal convention.

The intuition behind this result can be seen in figure 6: as σ goes to zero (the poor get nothing in the unequal convention) the risk factor of the equal convention also goes to zero because in a population near the {1,1} convention even if the A's (the poor) believed that virtually all of the B's would play 0, their best response would still be to play 1 (as they would not benefit at all from concluding a contract with a 0-playing B. More unequal shares in the {0,0} convention also reduce the risk factor of the unequal convention (it takes more non best responding A's to induce the B's to forsake their highly beneficial 0,0 contracts); but the risk factor remains positive even when the B's get all of the joint surplus in {0,0} for in this case $\gamma_0 = (1-\alpha^*) = 1/(1+\rho)$.

The introduction of idiosyncratic play removes the deterministic dependence of outcomes on initial conditions which characterizes the non-stochastic approach. Rather, the stochastic approach allows predictions of the average population state over a sufficiently long historical period along with a rather strong characterization of the nature of these stochastically stable states. The approach thus provides one account of how the institutions which Parsons termed "evolutionary universals" might come to be recurrent historically and ubiquitous at any given point in time: institutions supporting stochastically stable states would have been, as he put it, "likely to be 'hit upon' by various systems operating under different conditions."

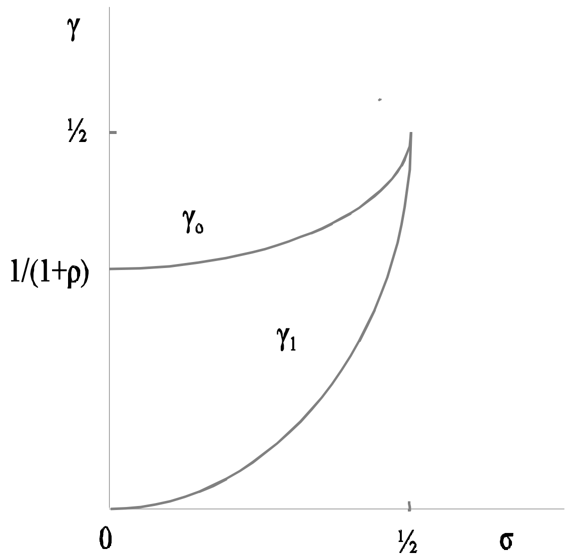

Figure 6. Risk Factors and the Degree of Inequality. Note: greater inequality of shares in the unequal convention (lower σ) reduce the risk factors of both conventions, but affect the equal convention more than the unequal one.

4. Collective action

The major shortcoming of the stochastic evolutionary approach is that it abstracts from the important role of collective action in the process of institutional innovation and transformation. It was not a fortuitous piling up of accidents that doomed apartheid or Communism, but rather the deliberate and coordinated actions taken by those seeking to live under other institutions. What difference does it make if we take account of the fact that non-best-response play is often intentional and bunched to some extent by design rather than purely by chance? Once intentional non-best-response play is introduced we will see that the egalitarian characterization of stochastically stable states is not robust if the rich are few and the poor many. The reason is that when non-best-response play is intentional there is just one way (rather than two) that a convention can be overturned (by the actions of those who would benefit more at the other convention) and the larger numbers of the poor militate against a sufficient fraction of them adopting a non best response to displace the equilibrium under which they do poorly.

By collective action I mean the intentional joint action towards common ends by members of a large group of people who do not have the capacity to commit to binding agreements prior to acting (that is, they act non-cooperatively). Examples are strikes, ethnic violence, insurrections, demonstrations, and boycotts. Of course an individual's participation in a collective action may be modeled as an idiosyncratic non-best-response which does not take the form of stochastically generated "errors" but instead represents an intentional action motivated by the desire to improve one's wellbeing and perhaps the wellbeing of others. For this reason it is likely that the extent of non-best response play will covary among individuals and depend on the payoff structure and other aspects of the pattern of social interaction defining the underlying game.[12]

To clarify the underlying processes, I will first analyze a degenerate case in which individuals participate in a non-best-response collective action when it is in their individual interest that the action take place (postponing a more realistic account of collective action as an n-person public goods game). Suppose that everyone may update in each period, and assume that there is a probability ε that each person is "called to a meeting" at which those attending consider undertaking a non-best-response action. For example assume the B-favorable convention {0,0} obtains and some fraction of B's (resulting from the draw) are considering switching to offer a 1-contract instead. But they cannot benefit from switching because they prefer the status quo convention, and destabilizing it -- should sufficiently many of the other possible B-innovators also switch -- could propel them to the alternate convention under which they would be worse off. These potentially idiosyncratic players would thus decline the opportunity to innovate.[13]

By contrast, imagine that the same number of A's were randomly drawn for deliberation of the merits of a switch away from the governing convention {0,0}, and suppose that should some of them adopt a non-best-response, this will be common knowledge. Each then might reason as follows. If they number $\varepsilon>\alpha^*$ and if all of them switched, the best response for the

[12] Bergin and Lipman (1996) analyse state dependent mutations. The proviso that play is non-cooperative excludes the degenerate case (with which I begin for purposes of illustration) of groups whose structure allows the assignment of obligatory actions to each of its members. Of course most successful collective actions include a wide range of selective incentives and sanctions to deter free riding, but few if any groups have the capacity to simply mandate group-beneficial behaviors by individual members.

[13] Favored groups, like the B's in convention {0,0} may deploy informal sanctions or state power to minimize idiosyncratic play of their own members.

B's would be to switch as well. Knowing this, should the fraction ε of the A's offer 1-contracts this period, they would anticipate the B's response and so would persist in offering 1-contracts in the next period and the A-unfavorable convention {0,0} would be displaced.

Because if $\varepsilon<\alpha^*$ there could be no benefit to collective action even if it were uniformly successful, let us analyze the case for which $\varepsilon \geq \alpha^*$. To lend some concreteness to the case let us say that switching means to engage with other A's in a strike, refusing to accept any outcome less than a_{11} (all this means is to offer a 1-contract, so the strategy set is unchanged.) We can explore the long run behavior of the system by calculating τ_0, the expected waiting time before a strike by the A's induces a transition from convention {0,0} to {1,1}. This is the inverse of the probability μ_0 that in any period a transition from {0,0} will be induced or $\tau_0 = 1/\mu_0$ To determine this probability, count the subsets of A's sufficiently numerous to induce a transition, then determine the probability (given $\in$) that each subset will be drawn; summing the probability that any transition inducing event occurs gives us μ_0. Thus if there are n A's, any subset with $n\alpha^*$ or more members will induce a transition, and each member of each subset is "called" with probability $\in$, and using $C_{n,m}$ to indicate the number of subsets of m members in a population of n individuals we have

$$\mu_0 = \sum C_{n,n\alpha+i} \in^{n\alpha+i} \quad \text{for } i = 0 \ldots n(1-\alpha^*)$$

An example will clarify the calculation. Suppose $\in = 0.1$, four individuals (W,X,Y,and Z) make up the A sub population, $\alpha^*=3/4$; then the A-unfavorable convention E_0 will be displaced by any of the following combinations: WXY, XYZ, YZW, ZWX, and WXYZ the first four of which will each occur with probability 0.001 and the last with probability .0001, so $\mu_o = .0041$ and $\tau_o = 244$ periods. As we want to know the long run average behavior of the system, we calculate τ_1 in a manner analogous to the above and express the average time at or near E_0, λ_0 as

$$\lambda_o \equiv \tau_o/(\tau_0 + \tau_1)$$

with $\lambda_1 \equiv 1-\lambda_o$. If there are three B's and $1-\beta^*$ (the critical fraction required to displace the B-unfavorable convention E_1) is 2/3 then $\mu_1 = .031$ and $\tau_1 = 32.2$ periods, so $\lambda_o = 0.88$.

Figure 7 gives the results of this calculation for $\in = 0.1$, where the two sub-populations each have 12 members and for various values of σ and ρ. Where E_0 is identical to E_1 ($\rho = 2$ and $\sigma = ½$) the average time spent at the two conventions is equal. One can see a band of conventions (similar to the locus YY in figure 5) with which like ($\rho = 2$ and $\sigma = ½$) generate equal average waiting times (for example, $\rho = 2.5$ and $\sigma = 0.2$ generates this result, as does $\rho = 2.25$ and $\sigma = 0.3$). The population will spend virtually all of the time at conventions more efficient or more equal than these and virtually none of the time at conventions less efficient or less equal. The restriction of non-best-response play to group beneficial actions doubles the waiting time between transitions, but it does not alter the long term average behavior of the population.[14]

Figure 8 illustrates the effects of modifying the conventional stochastic evolutionary model by extending the analysis to sub-populations of different size along with the restriction of idiosyncratic play to group-beneficial non-best responses. By contrast to Figure 7, when population sizes differ the intentional nature of non-best-response behavior makes a difference: unequal and quite inefficient conventions may be highly persistent. Thus while a convention with $\sigma = 0.3$ needed a ρ of 2.25 to be equally persistent to E_1 in the equal population size case, it has this status when $\rho = 1.25$ if the A's number 18 and the B's only 6. Where there are 21 A's the population will spend most of the time in the unequal convention even if its level of efficiency is half that of the equal convention.

So far I have abstracted from the problem of collective action by assuming that whenever a sufficient fraction of a sub-population is "called" they will adopt a non best response if they (and their group) would benefit from all of those called adopting the non best response. Collective action may be more adequately modeled by imposing a particular social structure on the process generating non-best-response play, one sufficient to explain why actions which are non-best-responses in the initial game may nonetheless be the result of intentional action when the game is amended to include the possibility of collective action. Thus what is needed is a model of the coordination problem posed by collective action, nested in the larger population game representing institutional evolution. Taking account of both the intentional nature of collective action and the coordination problem peculiar to it will augment the stochastic approach in illuminating ways.

Because collective actions generically take the form of n-person public goods games in which the dominant strategy is non-participation if preferences are wholly self-regarding, the extended model must address incentives for each to free ride on the activities of others in pursuit of commonly shared objectives. A second desideratum is that the model should reflect

[14] This counter-intuitive result does not hold for the non-degenerate model of collective action presented below. The reason why it is true for the degenerate case is as follows. Recall that the risk factor γ_i of the convention {i,i} is the minimum number of players whose idiosyncratic play may induce a movement from the other convention to {i,i} without further errors, and the convention with the minimum risk factor is the stochastically stable state. Thus in the case of stochastic (non-intentional) best responses, $\gamma_1 = \alpha^* \wedge \beta^* = \beta^*$ and $\gamma_0 = 1-\alpha^* \wedge 1-\beta^* = 1-\alpha^*$. The intentional collective action framework of course makes β^* and $(1-\alpha^*)$ irrelevant, as these are germane only to transitions from which those switching do not benefit. So define the (collective action) modified risk factors $\kappa_1 = \alpha^*$ and $\kappa_o = 1-\beta^*$ (non best responses by A's are the only way that E_0 can be displaced, and analogously for E_1.) The stochastically stable state is then $\kappa_1 \wedge \kappa_o$. Because $\kappa_1 \wedge \kappa_o = \gamma_1 \wedge \gamma_o$, the two approaches must identify the same convention as the stochastically stable state.

the fact that opportunities for collective action often arise by chance, or at least in ways too complex to tractably model, examples being economic depressions, wars, price shocks, booms, and natural disasters. Finally, unlike idiosyncratic play, participation in collective action is not only intentional rather than accidental but is also conditional on one's beliefs about the likelihood and consequences of a substantial number of one's kind changing behaviors. For this reason facts about global rather than simply local payoffs (that is, payoffs both in the present convention and in the alternative, rather than those in the neighborhood of the current population state alone) may have a bearing on the outcomes.[15]

Engaging in this collective activity yields in-process benefits of two types. First, irrespective of the consequences of the action, conformism (or punishment of non-conformists) imposes a cost on those not adopting the most common action. So, let c be the cost of being a sole non-conformist, and the conformism costs to those striking being (1-s)c where s is the fraction of those "called" who strike, and the costs to the non strikers is sc. Further, there are net benefits or costs associated with the action which may be independent of the numbers participating, including both the time, resources and possibly risk of harm associated with the collective action as well as the positive value of participating, or what Wood (2002) terms the "pleasure of agency".[16]

It is reasonable to suppose that these subjective benefits depend on the magnitude of the gains to be had if the action is successful, not primarily because these gains are a likely consequence of one's individual participation (which is very unlikely in large groups) but because the magnitude of the gains to be had is plausibly related to the strength of the norms motivating the action. So let the net subjective benefits for an A engaging in a collective action to displace convention {0,0} be

(6) $\delta=\delta(a_{11}-a_{00})$

[15] This means that individuals are forward looking to the extent that they can anticipate the consequences of successful collective action.

[16] Compelling evidence from the histories of collective action (e.g. Moore (1978)) anthropology (Boehm (1993)) and experimental economics suggests that individuals knowingly engage in costly actions to punish violations of norms, even when these actions cannot otherwise benefit the individual. Examples of the latter include ultimatum game respondents' rejection of what are deemed unfair offers even when these are quite substantial, and a willingness to engage in costly punishment of low contributors in public goods experiments, even when the punishment cannot possibly affect one's own payoffs (as for example on the last round of the game) Fehr and Gaechter (2000). See Bowles and Gintis (2000) and the works cited there.

where δ is a positive constant, reflecting the fact that joining a collective action in pursuit of an institutional change from which one and one's peers will not benefit confers no benefits.[17] If the strike fails (because two few participate in it) the status quo convention will persist, and all A's will get a_{00} in subsequent periods independently of whether they participated in the strike or not, so the relevant comparison is between the single period net benefits to striking (insisting on contract 1, refusing contract 0) or not are:

(7) $u_1 = \delta(a_{11}-a_{00}) -(1-s)c$

(8) $u_0 = a_{00} -sc$

These payoff functions are illustrated in figure 9, from which it is clear that potential innovators' expected payoffs will exceed those of non-participants, and they will hence elect to strike, if they believe that at least s* of their fellows will join in, where (finding the value of s that equates u_0 and u_1):

(9) $s^* = ½ - [\delta(a_{11} - a_{oo}) - a_{oo}]/2c$

How might A's beliefs be formed? The simplest supposition consistent with the above model is that having no information about what the others will do, each believes that the likelihood of each of the others participating is ½, so the expected fraction participating is ½, and all will participate if s* less than one-half.[18]

Thus unanimous participation (of those "called") will occur if striking is the risk dominant equilibrium of the collective action game, requiring that the numerator of the second term in (9) be positive, or that the "pleasure of agency" outweighs the loss of a single period's

---

[17] Conventions typically not only allocate gains but also influence the cultural and political conditions relevant to the net costs and benefits of engaging in collective action. But I here abstract from this (the δ's are not subscripted to indicate the convention defining the status quo ante.)

[18] The choice of ½ is conventional but arbitrary; individuals may have prior beliefs of the fraction likely to participate based on previous similar situation and the like. If individuals then apply their reasoning to each of the others (each, supposing that half will participate, will also participate) they would then correctly predict that s=1; but while this second round of induction may determine whether the individual expects the collective action to be successful in displacing the convention, is not relevant to the individual's behavior, as the relative payoffs of participating or not are independent of the success of the action.

income. (Note that while payoff dominance of the alternative (a_{11}-a_{00} > 0) is a necessary condition for participation, it is not sufficient, as it does not insure that $\delta(a_{11}-a_{00})-a_{00} > 0$.)

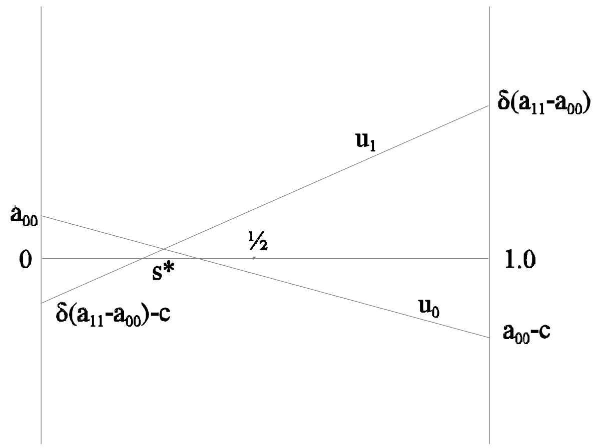

Figure 9. The Collective Action Problem. Note if $s^*<½$ the risk dominant equilibrium is universal participation in the non-best-response action (from Wood, 2002)

The properties of the dynamical system are substantially altered by the representation of idiosyncratic play as intentional collective action. Notice that if $\delta(a_{11}-a_{00})-a_{00} < 0$ collective action will not take place (irrespective of the numbers of randomly drawn potential innovators), so the A-unfavorable convention {0,0} is an absorbing state. Thus the dynamical system with collective action as the form of non-best-response play is non-ergodic, and institutional lock-ins are possible, with initial conditions determining which of the two conventions will emerge, and then persist forever. To see that this must be the case for a finite pleasure of agency parameter δ, consider an unequal convention with $a_{11}-a_{00} = \eta$ letting η become arbitrarily small must make $\delta(a_{11}-a_{00})-a_{00} < 0$ so collective action by A's will not occur and E_0 if it ever occurs will persist forever. Thus there must exist a set of conventions, less equal than E_1 and no more efficient, that are absorbing states.

However, if technical change made the E_1 progressively more efficient by comparison to E_0, $\delta(a_{11}-a_{00})$ would eventually exceed a_{00}, the dynamics would become ergodic and a transition from E_0 to E_0 would eventually take place, with transitions in reverse direction becoming more unlikely over time as the increase in b_{11} raises the resistance along the path from E_0 to E_1 (or what is the same thing, reduces μ_1). Thus if technical progress were for a period of time stimulated by a set of institutions, they might become predominant until the institutional demands of new technologies required different conventions. This is very roughly Marx's account of history as a progressive succession of "modes of production," each liberating the "forces of production" for a period, then becoming a "fetter" on further technological advance and being replaced through the collective action of the class which would benefit by a shift to a newer convention more consistent with the new technologies.

5. Conclusion: the institutional ecology of inequality

The integration of chance and collective action developed here is far from the first proposed marriage of Darwin to Marx. Writing to Engels in 1869, Marx saw parallels between *The Origin of Species* and their just completed first volume of *Capital:* "Although it is developed in the crude English style, this is the book which contains the basis in natural history for our view." Darwin, embattled by Creationists, declined Marx's subsequent offer that *Volume II* would be dedicated to him, perhaps sharing the view that Voltaire is said to have expressed when on his deathbed he was asked by the attending priest to condemn the devil: "This is no time to make new enemies."

Stochastic evolutionary game theory has recently made available powerful analytical tools of Darwinian inspiration, providing an illuminating framework in which to consider the problem of institutional change and "evolutionary universals." A particularly important contribution is to provide a causal mechanism – missing from the Parsonian, Marxian and neo-institutionalist accounts -- accounting for the evolutionary success of efficient and egalitarian institutions.

However, taking account of differences in group size and the intentional nature of collective action suggests that the standard stochastic evolutionary game theory model may need extension to be relevant to the historical evolution of institutions. I have suggested two reasons why evolutionarily successful institutions may be neither efficient nor egalitarian. First, independently of group size, moderate levels of inequality may deter collective action (where the degree of inequality is insufficient to motivate participation); thus unequal conventions may persist indefinitely. Second, independently of the problem of motivating collective action, where the poor are many and the rich few the population may spend most of its time governed by conventions which are both highly unequal and inefficient.

In this second case, the reason why the system will spend most of the time in the unequal shares convention is that the B's, who prefer this convention are relatively few in number, so that the likelihood that a random draw will yield a number of them sufficient to displace the convention which they do not prefer is greater than for the A's. This advantage of small numbers is unrelated to conventional reasoning as to why collective action in large groups is difficult to sustain.. The conclusion is that societal inequality of the type described is capable of sustaining highly unequal and inefficient conventions over long periods.

A related concern about the stochastic evolutionary game framework is that it applies only to the very long run. For reasonable updating processes, group sizes, and rates of idiosyncratic play, the average waiting times for transitions from one basin of attraction to another are extraordinarily long, certainly surpassing historically relevant time spans, and for some not unrealistic cases exceeding the time elapsed since the emergence of anatomically modern human life.[19] While the biological processes underlying Sewall Wright's shifting balance theory of phenotypic change may work over hundreds of thousands of generations, an analogous approach in the social sciences must be relevant to vastly shorter time horizons.

However, a number of plausible modifications in the updating process can dramatically accelerate the dynamic process, yielding more historically relevant predictions. Common to all of the modifications is that they take explicit account of some distinctive human characteristics which may help to explain a rate of institutional change considerably more rapid than the rates of phenotypic change governed by biological evolution. These may be summarized by saying that human interactions tend to be local, intentional, and guided by conformism. More so than most other animals, humans maintain well defined group boundaries among non-kin, most interactions taking place within rather than across these boundaries. And within-group behavioral differences are dampened by the well documented human tendency towards conformist learning, which gives a privileged status in the updating process of actions which are prevalent among one's peers. Small group membership increases the relative importance of unlikely random events and hence the likelihood that stochastic variations in play will induce transition times among conventions; this induces rapid transition times for the population as a whole. Conformism gives rise to positively correlated non-best-response play -- each member of the population is more likely to adopt a non-best response the more others are doing the same. This produces greater bunching of idiosyncratic play and hence under plausible conditions accelerates the process of transition.

We do not know, of course, whether conformism and social (group) segmentation can provide a plausible account of historically observed processes of institutional change, as this will depend on empirical knowledge of these processes which at present is quite inadequate. But given the apparent importance of both segmentation and conformism, the unreasonably long waiting times implied by models with uncorrelated errors and random matching and learning in the population could be addressed by taking a step in the direction of realism about the nature of human interactions.

Institutions differ in evolutionarily relevant ways not captured by measures of efficiency, distributional shares, and relative group size, of course. Some institutions may facilitate collective action of the disadvantaged, while others make it more difficult to coordinate. Marx, and many since, have believed that the social conditions of industrial capitalism constituted a schoolhouse of revolution, by contrast with earlier institutions of sharecropping, tax farming

[19] See also Axtell, Epstein, and Young (2001).

in societies of independent peasants, and slavery, for example. These differences may be represented in the differing net benefits of collective action, δ, subscripted by the conventions to which they apply.

Rather than pursuing these extensions of models depicting within group processes of institutional change, we turn now to the manner in which between group interactions may induce institutional evolution. By contrast to the within group models, the multi-level selection approach gives rather strong predictions of the evolutionary success of institutions that are both egalitarian and efficient. The reasons why this is so, as we will see, are quite different than those advanced for similar conclusions by the Marxian or the stochastic evolutionary game theory approaches.

Works Cited

Axelrod, R. and W. D. Hamilton (1981). "The Evolution of Cooperation." *Science* **211**: 1390-1396.

Axtell, R. L., J. M. Epstein, et al. (2001). The Emergence of Classes in a Multi Agent Bargaining Model. *Social Dynamics*. S. Durlauf and H. P. Young. Cambridge, MIT Press.

Bergin, J. and B. L. Lipman (1996). "Evolution With State-Dependent Mutations." *Econometrica* **64**(4): 943-956.

Boehm, C. (1993). "Egalitarian Behavior and Reverse Dominance Hierarchy." *Current Anthropology* **34**(3): 227-254.

Bowles, S. (2000). "Economic Institutions as Ecological Niches." *Behavior and Brain Sciences* **23**.

Bowles, S. (2002). *Economic Behavior and Institutions: An Evolutionary Approach to Microeconomics*, Princeton.

Bowles, S. and A. Hopfensitz (2001). *The Co-Evolution of Individual Behaviors and Social Institutions*. Workshop on Groups, Multi-Level Selection, and Evolutionary Dynamics, Santa Fe Institute.

Brenner, R. (1976). "Agrarian Class Structure and Economic Development in Pre-Industrial Europe." *Past and Present* **70**: 30-70.

Cohen, G. A. (1978). *Karl Marx's Theory of History: a Defence*. Princeton, Princeton University Press.

Cortes, H., translated and edited by Anthony Pagden (1986). *Letters From Mexico*. New Haven, Yale University Press.

Fehr, E. and S. Gaechter (2000). “Fairness and Retaliation: The Economics of Reciprocity.” *Journal of Economic Perspectives*.

Greenberg, J. B. (1989). *Blood Ties: Life and Violence in Rural Mexico*. Tucson, The University of Arizona Press.

Kandori, M. G., G. Mailath, et al. (1993). “Learning, Mutation, and Long Run Equilibria in Games.” *Econometrica* **61**: 29-56.

Lefebvre, G. (1947). *The Coming of the French Revolution*. Princeton, Princeton University Press.

Marx, K. ((1859)1976). *Capital I*. Harmondsworth, Penguin.

Moore, B. J. (1978). *Injustice: The Social Bases of Obedience and Revolt*. White Plains, M.E.Sharpe.

North, D. C. (1981). *Structure and Change in Economic History*. New York, W. W. Norton & Co.

Paddock, J. (1975). “Studies on Anti-Violent and 'Normal' Communities.” *Aggressive Behavior* **1**: 217-233.

Paddock, J. (1991). “Violence and Sociality in Human Evolution.” *Current Anthropology* **32**(4): 391-428.

Parsons, T. (1964). “Evolutionary Universals in Society.” *American Sociological Review* **29**(3): 339-357.

Smith, J. M. and G. R. Price (1973). “The Logic of Animal Conflict.” *Nature* **246**: 15-18.

Soboul, A. (1974). *The French Revolution*. London, NLB.

Trotsky, L. (1932). *The History of the Russian Revolution*. London, V. Gollancz.

Weibull, J. W. (1995). *Evolutionary Game Theory*. Cambridge, MA, MIT Press.

Williamson, O. E. (1985). *The Economic Institutions of Capitalism*. New York, Free Press.

Wood, E. (2002). *Insurgent Collective Action and Civil War In Rural El Salvador*. Cambridge, Cambridge University Press.

Young, H. P. (1993). "An Evolutionary Model of Bargaining." *Journal of Economic Theory* **59**(1): 145-168.

Young, H. P. (1998). *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton, NJ, Princeton University Press.

Young, P. (1995). "Conventions." *Journal of Economic Perspectives* **10**: 105-122.

Zeitlin, M. (1967). *Revolutionary Politics and the Cuban Working Class*. Princeton, Princeton University Press.

Figure 7. Average fraction of time spent at E_0 (Equal Sub-population Size Case). Note: the other convention is E_1 for which $\rho = 2$ and $\sigma = ½$).

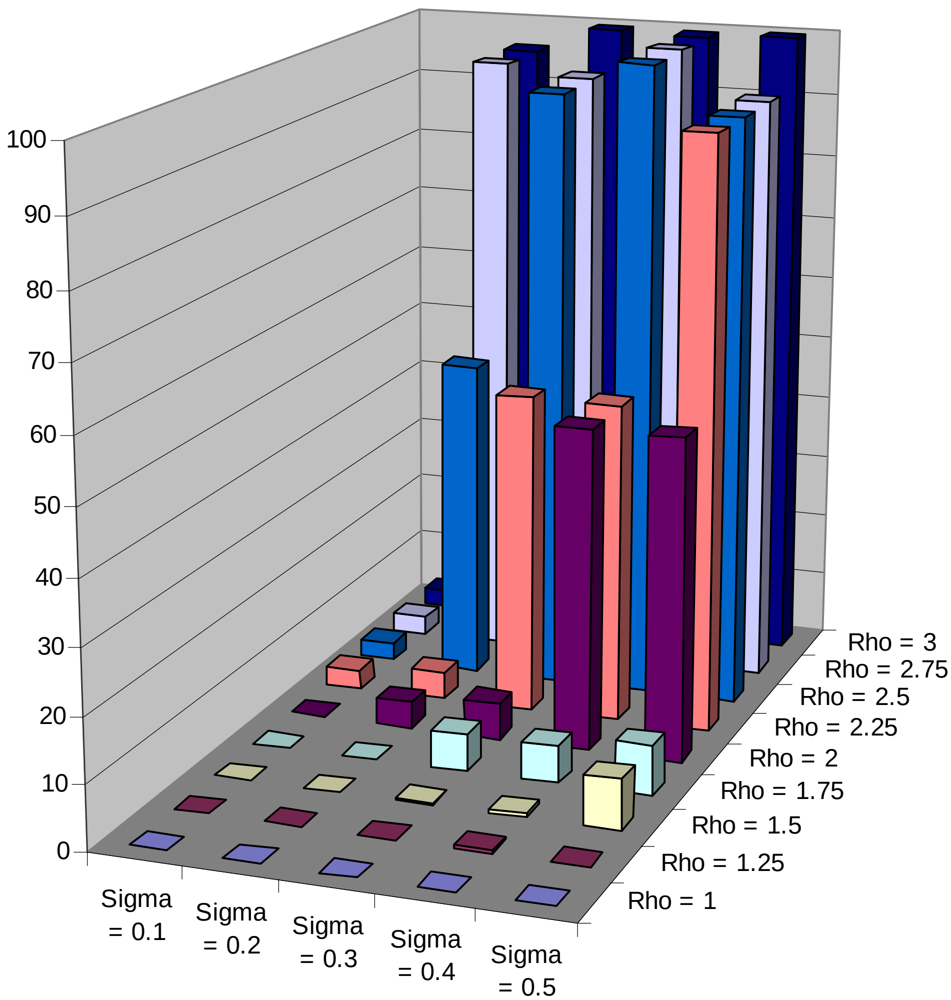

Figure 8. Average fraction of time spent at E_0 (Unequal Sub-population Size Case). Note: the other convention is E_1 for which $\rho = 2$ and $\sigma = ½$).