Natural Selection, Economics and Probability

by Armen A. Alchian

Introduction

1. In this paper we emulate Karl Fox in plunging boldly beyond the traditional boundaries of economics, by applying the principle of natural selection to social life generally, to economic theory in particular, and (to a limited extent) to the foundations of probability.

2. The natural selection concept achieved fame, of course, through Darwin's Origin of Species (1859), and was applied almost at once to the social sphere. This literature has continued to the present day when it is in fact enjoying something of a revival. So, what are we doing that hasn't already been done?

3. The answer is that the concept itself requires clarification. Natural selection, as used here, is a universal principle applying to any system having variability. It is important not to tie it down to the concrete biological-genetic-demographic matrix to which it was first applied.

Once clarified, it turns out to have a remarkably wide range of implications, albeit of a long-run character.

4. What relation does this approach have to sociobiology? Little or none. Sociobiology claims that the genetic limits of human nature are a (or the) major influence on the structure of social life. The extent to which this claim is true is interesting but not relevant to our analysis,

since the latter applies in any case. (The relative importance of different transmission mechanisms will be affected, as well as the length of time required for trends to work themselves out, but not the trends themselves.)

5. The present article is part of a much larger work in progress that we call FOCUS (for Friction, Organization, Capital, Uncertainty, and Selection) and itself focuses on the letter S. We cannot hope to give a full presentation in one article, even of the Selection part (the pieces fit together like a system of simultaneous equations). Nor is the time ripe to plunge into formal models. Instead we give a panoramic sketch of the major ideas, ~~developing them~~ showing how they eventuate in some concrete predictions (long-range, as we mentioned above), some of them surprising. The discussion is at the verbal level throughout, except at one point where a cameo sketch of some of the rather elegant formalism lurking just below the surface is presented.

The Natural Selection Concept

1 Fox outside traditional boundaries

2 we follow: in many cases solutions to basic problems come from combining basic ideas (like assembling the pieces of a simultaneous equation)

3 in this case from economics, probability and biology
(part of larger FOCUS project - we concentrate on the U (uncertainty) and the S (selection))

4 cannot hope to give a full presentation in 1 article; nor aim at full rigor panoramic sketch of ideas exploratory
nonetheless - remarkable what concrete conclusions we can reach with the simplest of arguments

5 foundational problems exist in both econ and probability - stat
is good ignoring the correct approach what is probability

6 "natural selection" is central idea in biology - but it is important to see it as a universal property of systems having variability, and not to tie it too narrowly to the genetic - demographic level (the mistake of sociobiology)

~~7 abstract concept of selection~~

Frequency distribution

8 start with the idea of a frequency distribution - a list of mutually exclusive attributes, and the number of entities in a population having these attributes at a given time e.g. people classified by age, sex, income, education, location, etc., animals classified by species ...

number of molecules of each type
number of molecules of each type). Now, going to a later time, the distribution in space will have changed

This redistribution can arise in many ways: by "immigration" of new entities to the system, by "migration" of old entities from the system; by transformation or "migration" of entities to another attribute-state (including complex transformations, in which a cluster of entities transforms collectively into a different cluster, as in economic production)
collectively
or by creation, "birth" of new entities, or destruction "deaths" of existing entities

The rates at which these various processes occur will in general be influenced by the distribution of these attributes themselves; a natural selection situation occurs when the growth in numbers having some attribute is positively associated with the number already having that attribute
situation
(a situation of "positive feedback", or "autocatalysis"; in economics this situation is usually associated with "economies of scale", except that the essential element of time is usually understressed)

& Some clarifying comments; first, positive association is generally only over some range of values, allowing the possibility of reversal, at least temporarily (eg congestion effects)
reversal

Second, the positive association may operate indirectly; the "size" of attribute 1 may encourage growth of attribute 2, which in turn encourages growth of attribute 1 - so that these attributes rise (or fall) together in "symbiotic" association; in general, a cluster of attributes may mutually aid each other's growth

nature of selection: one attribute's share can be enlarged only at the expense of some other

9. On the measure of size; this may be in absolute or relative terms (e.g. relative frequencies); because a relative measure brings out the competition.

also there may be several different natural measures of size – frequencies when there are natural units, but also total volume, or mass, or other weighting may be more appropriate.

e.g. chemical measures may be the terms of numbers (moles, using Avogadro's number as the counting unit) or mass; biological species in numbers or in biomass; ~~people~~ in social life, total wealth or money value is often a more appropriate measure than numbers (further discussion later)

10. The above ideas apply to biological systems – where indeed the concept of natural selection was first recognized – but they are framed to apply more generally.

Biologists concentrate on a particular mode by which natural selection operates – Mendelian inheritance and gene frequencies.

In the social world there are many other modes of transmission besides relative survival rates – there is learning, growth, migration, invention, mutual adaptation etc. ~~(Donkeys or mares, [illegible])~~

~~11. The application of ideas from biology to the analysis of social life itself has a long history (see Sorokin Contemp. Socio. Theories). including the ~~social Darwinists~~ pejoratively named social Darwinists, Spencer, Haeckel, Sumner & Keller, Carver, Ross(?), Bagehot, and to some extent Darwin himself in the Descent of Man. Of these the most eminent is Herbert Spencer whose writings deserve a most sympathetic revival (But this is another paper).~~

12. A related but distinct approach is to apply natural selection to the realm of ideas.
KEEP ideas. A set of exclusive hypotheses do in effect compete with each other

For example, rival scientific theories explaining the same set of phenomena. This notion has a long history

KGBP

; it must prominent modern representative is Karl Popper (see Campbell evolutionary epistemology). F. von Hayek takes a similar approach in connection with alternative lines of business enterprise. (See Donald T. Campbell, "evolutionary epistemology," 413-463, The Philosophy of Karl Popper (Open Court, 1974), for a brilliant survey

3. What about rival ideologies, as opposed to scientific theories? How do ideologies about rival ideologies — say in the sphere of religion, or nationalism, or economic or political organization — compete? There is certainly a struggle for adherents among them, but adherence seems to take the form of loyalties, attitudes and values rather than beliefs or disbeliefs. Nonetheless, every ideology has a cognitive substratum — it makes assertions about the world — and is therefore subject to the same kind of intellectual competition as scientific theories. Conversely, rival theories carry action programs with them — how to conduct research — and are themselves obtain subjects of fierce loyalties (cf. Kuhn's study of scientific revolutions and the more recent Sharon Traweek). We should therefore expect the same general principles to apply to selection among rival ideologies as among theories, and keep in mind the emotional-activist concomitants of this process.

Is the notion of natural selection among ideas merely a metaphorical analogy to natural selection among physical entities, or is there some underlying structural identity? The latter is the case, and the unifying concept is that of probability. (For the time being we take a simple "subjectivist" interpretation of probability). Let $P(H_i)$ be the "intensity of belief" in assertion H_i, on a scale from 0 (complete disbelief) to 1 (complete belief). Here $H_1, \ldots, H_n$ are competing hypotheses. New evidence E is observed, having likelihoods $P(E|H_i)$. The evidence leads to a

shift of belief intensities to $p'(H_i)$. The basic principle of Bayesian inference is that $p'(H_i) = p(H_i|E)$, which is proportional to $p(E|H_i)p(H_i)$. In other words each hypothesis gets knocked up or down in proportion to how well it predicted the observation E.

Now probability is, first of all, a measure in the technical sense (an additive set function) just as physical measures are. Second, probability mass ~~shift over time~~ gets redistributed over time (changes), just as physical mass does. Thirdly, there is a structural identity between the two. Consider for example, n genotypes with relative frequencies $m_1, \ldots m_n$ at one time; the relative survival of these is given by selective coefficients $s_1, \ldots s_n$ yielding "posterior" relative frequencies proportional to $m_1 s_1, \ldots m_n s_n$

15 More important than this structural similarity is the fact that probabilities and measures must be considered conjointly. The state of beliefs (cognitive structure) of a person is part of his characteristics, and beliefs state have a distribution over the population. State of belief, together with preferences and other characteristics determine actions, which in turn determine one's relative success in the world. So beliefs are themselves subject to selection. In general, people with belief systems that are "closer to the truth" in some sense will tend to prosper (though this association is itself only probable - some people are "lucky" than others).

16 To put the paragraphs above together: a double selection process is going on simultaneously. within individuals, beliefs are being modified by experience. This is learning. At the "migration" of belief systems to states more in accordance with the real world. At the same time, individuals are themselves undergoing changes in their fortunes and numbers, in part on account of

circumstances. A famous remark by Planck is that ~~progress in science~~ the outlook of science changes by ~~progress~~ of the older systems dying off.) If so, the external forces are dominant over the internal. (Cf. Joseph Priestley and the phlogiston theory).

17. To sit this out significantly, let P_{it} be the belief system of person i at time t (represented say by a probability distribution over certain random variables) Let w_{it} be the "weight" of person i at time t in some appropriate measure (say number, influence or wealth). Then, ~~if~~ $\sum_i w_{it} = 1$, $P_t = \sum w_{it} P_{it}$ is "average social opinion". After some time, this changes to $P' = \sum w_i' P_i'$. Here the change $P_i \to P_i'$ is the internal one yielded by experience and learning of person i, while $w_i \to w_i'$ is the external change in relative fortune.

18. P and P' above are derived from physical measures. An analogous process occurs within Bayesian inference itself. Let there be alternative hypotheses $H_1, \dots, H_n$ with probabilities $w_1, \dots, w_n$; let P_i be the probability over some random variable conditioned on H_i; then $P = \sum w_i P_i$ is my overall probability. Let evidence E be observed; then things change to $P' = \sum w_i' P_i'$, where $P_i'(F) = P_i(F|E)$ as expected but also w_i' equals $w_i P(H_i|E) \neq w_i$ in general.

18. (The paragraph above should not be taken too literally. It presupposes that beliefs are, or, represented by probabilities, all over the same set of random variables (otherwise $\sum w_i P_i$ is not well defined). This latter is false - energy thought of by one person are not thought of - given access to - another. And even the representation by probabilities must be discussed further.

18. Critique of Economic Theory

19. We now propose the following program: to examine how the fundamental concepts and principles of economic theory apply in a world in which natural selection operates. Which principles are implied by natural selection, which are compatible with it, which (if any) actually contradict it? We shall find entries in all three categories.

20. At first glance one might expect few conclusions to flow from the mere presence of natural selection. It is, after all, a regulative rather than constitutive principle, regulative rather than constitutive, as Kant might say, with no concrete predictions per se. And this is true if one looks to it to make "short-run" predictions. Here a dynamical mechanism — captured say by an econometric model — simply plays itself out. But even mechanisms operating in the short-run arise from long-run processes.

21. Natural selection comes into its own in the "long-run". For its influence while weak, is persistent. A weak force pulling constantly in one direction will ultimately assert itself against strong forces that cancel each other out over time. The decades and millennia of social evolution — and the eons of biological evolution behind them — have made us the kind of beings we are today and shaped the institutions we participate in. (This kind of analysis harks back to the style of the great classical economists.)

22. As a preliminary we must choose the appropriate measurement scale - shall it be numbers of people, monetary values, physical mass, or weighted modifications of these, etc.? (The question is a deep one, and we shall discuss it further below in connection with selection rules). The appropriate scale

economics th. appropriate scale is monetary value: all dollars have an equal vote, and equal monetary values exchange for (or are referring to) each other (we are referring to cross-sectional comparisons, not to value comparisons over time). cross-sectional comparisons

23. It is important to note that th. appropriate scale itself depends on th. institution and th. overall social structure. In the realm of politics — in democracies — it makes more sense to count each voting citizen equally. equally. We have then, two different ratios of exchange in the economic and political spheres, th. poor have a comparative advantage in politics, th. rich in th. marketplace. (Furthermore, "arbitrage" operations arise as a result, th. economic sphere invades th. political via campaign contributions, corruption, etc., th. political sphere invading the economic via redistributive laws). Elsewhere, other appropriate scales would be military power, religious-ideological influence, and official prerogative. (Arbitrage operates among these as well: How many divisions does th. Pope have? asked Napoleon) OMIT

24. One common jibe against the social Darwinists is that th. poor are fitter than th. rich since they produce more offspring. The fallacy involved here should be clear from th. discussion above. It is to take a scale of measurement appropriate to the biological world — biomass — and plunk it down into th. social where it does not belong. Similar fallacies abound — e.g. Bernhardi's equation of citizens with raw military power.

25. We are now ready to begin th. critique, using money value as the measurement scale and taking a long-run viewpoint, consider th. range of possible human behavior of a certain type. Is there a systematic relation between the values of this variable and changes of wealth? If so,

those which do not, and the latter will tend to be selected relative to those values yielding an actual decline in wealth. Here "selection" refers to the redistribution of total wealth from people having unfavorable traits to those having favorable values. If we imagine the range of possible traits distributed from left to right, from least to most conducive to wealth accumulation, then wealth will gradually redistribute itself toward the right.

26. Several clarifications must be made. First, we can extend the range of traits beyond known human experience. If far enough, nobody falls into the uppermost third of the range, then of course no selection will occur there: variety must already exist for selection to work on.

Second, the process is relative in two senses. First of all, we are talking about the function of total wealth accruing to various traits, not to the size of the total pie (whether the latter is growing or shrinking comes out as a by-product of this inquiry.) Next, where you are relative to people in general is what counts. Suppose everyone exhibited trait unfavorable to wealth accumulation. Those, those, who are least unfavorable would gain.

Third, actual wealth accumulation may not go as expected if the traits in question is correlated with other traits pulling in the opposite direction. The argument is a ceteris paribus one. (We may also add that correlation among traits is itself subject to natural selection. Traits that play well together get correlated by being selected in tandem).

27. We have individuals, families, regions, nations, etc. At which level does the selection process operate? All of them simultaneously! The logic of the natural selection argument makes sense at all levels. A trait may be favorable to an individual but unfavorable to the group he belongs to. Then the individual gains

imposing sanctions internalizing
imposing sanctions on their members to avoid such conflicts (i.e. "internalizing" of such
"externalities" occurs at all levels of the social hierarchy).
externalities hierarchy

25. Finally, note that redistribution of wealth occurs via transfer of own traits
can occur by individuals themselves changing traits ("migration" via learning, as
discussed above) as well as by passively Keeping a trait and observing
passively Keeping
good or bad results. A more complete discussion would also include the
intergenerational
mechanisms of intergenerational transmission of wealth within families and the
turnover of power in organizations.

~~25. We begin with the discussion of profit maximization~~

25. We now examine a series of economic principles concerning human behavior
(and market structure). It is remarkable what far-ranging conclusions follow
from the mere introduction of natural selection. The whole fabric of economic
theory becomes strange, yet still familiar: the angle of vision has changed.

26. We begin with the discussion of profit-maximizing behavior. This is the one
area where natural selection type arguments have long been recognized. To be exact, the emphasis
emphasis has been on the fact that even with ~~quite drastic~~
random behavior on the part of firms
Firms the results will be as if they were
maximizing profits. Those which accidentally do the right thing will
prosper, just as seed accidentally scattered on fertile ground will sprout.
(c.f. Hoover, Location of Economic Activity) S. Enke, "On maximizing profits"
Amer. Ec. Rev 41: 566-578, 1951
A. Alchian, "Uncertainty, evolution and economic theory" J Pol. Ec. 58: 211-222, 1950
M. Friedman "The methodology of positive economics", in Essays in Positive
Economics (University of Chicago Press, 1953)
this is an important observation, and accords with the spirit of

natural selection in producing teleological results via non-teleological mechanisms. Indeed this line of reasoning has been applied to social life generally, which produces a spontaneous order without conscious planning (Think of the growth of language or money). This tradition dates back to Mandeville, Hume, Smith, and Burke, down to Hayek in recent years (FA Hayek, Law Legislation and Liberty (University of Chicago Press, 1973) vols I-III). Another application in economics is to consumer behavior, where downward-sloping demand arises even from random acts. GS Becker, "Irrational behavior and economic theory" JPE 1962 J. Pol. Ec.

31. But the argument carries beyond this point, to the distribution of entrepreneurs' motives themselves. Given such a distribution, the more profit-minded firms will tend to be those that find the more profitable niches, the seeds that are more eager to grow and the more fertile sites. Furthermore, conscious selection of this sort is more efficient than selection arising from random behavior, for two distinct reasons. First, the latter involves wasted motion in making trial-and-error bankruptcy seeds fall on stony ground. Second, motivation persists over time, while luck does not, so we will find less squandering of gains under conscious motivation than under a random regime. It follows that conscious motivation makes an independent contribution over and above "spontaneous order". E. Penrose, "Biological analogies in the theory of the firm" Amer. Ec. Rev. 42:804-819, 1952, S. Winter, "Economic 'natural selection' and the theory of the firm," Yale Ec. Essays 4:225-272, 1964

32. Imagine then a distribution of motives among the managers of firms in an industry, from the single-minded pursuit of profit at one end, to concern with leisure, tradition and receiving a living at the other. Wealth will tend to redistribute from the latter end to the former. Note that the redistribution is relative: All firms may have positive income, but the strongly motivated ones have a higher rate of return. (The concept of economic — as opposed to accounting — profits captures this nuance. Firms with below normal but positive rates of return are

making negative economic profits but not going bankrupt.) Note also the world "tends" to "find" tendency the short run dumb luck may win out, but these breaks cancel breaks cancel out over time, and the association of wealth with profit motivation accounts itself. (The concept of "tendency" may be formalized by stating $\Delta E(r|m)/\Delta m > 0$, where m is the strength of profit motivation, r the rate of return, and E the conditional expectation operator). Finally, notice that

33. Imagine the foregoing tendencies operating over a long period of time. It is then reasonable to attribute high profit motivation to the surviving firms. The others will have dropped out, or been pushed out, or perhaps persist as small toys on the fringes of the economy — or will themselves have adopted profit-making behavior in order to survive. (Incidentally, the most profitable behavior of all may be to proclaim disinterest in profits, contribute to local civic enterprise, etc. Scrooge-like behavior may be self-defeating: the principle of enlightened self-interest). This argument appears to be the major rationale for the standard assumption of profit maximization in economic theory.

Are people behaviorists?

34. Now generalize from business motives to motives in general. Can anything be said about utility functions — indeed, if, indeed, behavior can be represented via utility functions? What about behavior in specific areas — family and children, attitudes toward work, saving and consumption, attitudes toward risk, ... It will turn out that quite far-reaching statements can be made in all these areas. Toward the future?

35. As a preliminary, consider the changing concepts of natural selection in biology. Before After the Origin of Species appeared in 1859, the popular concept was of physical combat, the survival of the ferocious. Tennyson's 'nature red in her, prosaically

reproduction, the leaving of a relatively large progeny to reproduce in the next generation (see Singham, The Meaning of Evolution) the latter concept is superior. Why? Because it closes a cycle: by going from adult in one generation to adult in the next it gets the total effect of selection pressures over the life cycle. It is this overall effect that is selected for.

36. To close a cycle in the social world, one must not restrict attention to any portion of it, not even to the entire business world. In particular one must not neglect the household-reproductive-leisure sector. In terms of time, effort and resources invested, this sector dominates all others.

37. Natural selection operates simultaneously in all sectors, including the household sector. Again, we consider a range of attributes — qualities, behavior patterns, motives, etc. — that have a differential effect on the growth of the number of entities having these attributes, leading to a redistribution of size (in appropriate units) among these attributes.

38. To take some single examples, consider the range of propensities to save. Thus, individuals, families or nations that save a large proportion of their incomes will — other things being equal — find their wealth growing at a more rapid growth rate than those saving little. Overall wealth will gradually redistribute away from the latter groups toward the former.

39. The same argument applies to industriousness. Those who devote much time, care and effort to productive activity will — ceteris paribus — find themselves accumulating a larger share of the world's wealth, at the relative expense of the leisure-minded.

36.

39. The rise of the bourgeoisie in Europe is a classic historical illustration, and of the national level, the rise of England and the Netherlands vis-a-vis Spain and Portugal. Japan and the "newly industrializing countries" of Asia are a contemporary illustration.

40. Attributes correlated with industriousness and thrift will increasingly share in the redistributional process. According to Max Weber's famous (debatable) thesis, the source of these bourgeois attributes was in religious ~~outlook~~. We don't have to decide the question of causality. If in whatever reason Protestantism is correlated with these attributes, then its adherents will find to acquire a larger share of the world's wealth.

41. ~~Human beings~~ Labor is a factor of production, and it can be enhanced by investments in education, health care, etc. This insight leads to the analysis of investment in human beings as one would ~~analyze~~ evaluate investment in capital equipment - in terms of income streams, rates of return etc (Becker Human Capital). For the natural selection approach, however, the human capital concept has a much deeper significance.

Consider a range of behaviors, some of which enhance human capital, (e.g. studiousness, healthy eating habits, etc.) some of which do the opposite. Then - assuming as always that other things are equal - wealth will tend to get redistributed from ~~the latter~~ people in the latter category to people in the former.

42. It is important to note that the redistribution depends on the behavior, not on the motives behind it. One may have a health-enhancing diet, for example, as a by product of following some religious

or, it might just be habit.

start

40.

45. The general principle involved here is the same as that discussed before in connection with profit maximization: Profit-~~maximization~~-enhancing behavior will be rewarded whether or not it is consciously motivated. When we step out of the business world, however, this principle becomes far more important. The reason is that it is far more difficult to behave "rationally" in this respect outside the economic realm: Causal connections are less clear and less well understood, and there is no system of public prices to guide one's conduct. As Schumpeter put it, the economic is the realm of logic and rationality (J A Schumpeter, "Marx" in Ten Great Economists)

Nonetheless, natural selection still operates as a kind of non-Smithian invisible hand.

44. The principle above is analogous to Darwin's contrast between artificial and natural selection. After a long discussion of the strong effects man has achieved by artificial selection in varying species — especially pigeons — he notes that the effects of natural selection must be stronger still, for several reasons. First, nature selects continuously while man selects sporadically. Second, nature acts simultaneously on all features of the organism, overt and covert, while man selects only from among those visible features that catch his attention and fancy. (Darwin Origin of Species)

Artificial selection would correspond to that situation where a person selects his own conduct from a consciously motivated goal. Natural selection (in Darwin's sense) corresponds to the general fact that conduct is differentially rewarded or punished regardless of the motivation leading to that conduct. (We, ~~personally~~, are using the term "natural selection" in a wider sense.) The same arguments that Darwin adduces apply here: Conscious motivation is sporadic, and applies only to those

features of the world that we can perceive and that catch our attention. (We, of course, are using the term "natural selection" in a broader sense than Darwin's.)

45. Keep in mind that our use of the term "natural selection" is broader than Darwin's. Conscious motivation is itself differentially rewarded (hence "naturally" selected) if it is of the right sort. We argued above that profit-seeking motivation tends to be "stamped in" by this mechanism. Outside the business world this happens much less directly. "Reason uses its cunning" (Hegel) to reward in part differentially the various adherents of the different ranges of ideologies, nationalisms, religions and political programs, and styles of living that are extant at any time. In this way the overall pattern of social life gets shaped into a semblance of order.

46. In the business world, prices and rental rates equilibrate to form a network of relations studied in capital theory: the price of an asset equals the discounted value of its future stream of earnings (quasi-rents), and its rental rate equals the value of its marginal product. These values provide the guideposts within which profit-seeking motives can operate intelligently. By extension, an implicit system of prices and rental rates arises covering the entire social system — implicit because there are no formal markets. (Indeed most of these transactions are not formally exchanges at all but things like grants, promises and threats and implicit understandings — cf. M Mauss *The Gift*, Boulding *The Economy of Love and Fear*).

47. This implicit price system satisfies two broad conditions. First, it satisfies — at least approximately — the same discounted stream and marginal product conditions as above. Second, people who act as if they were reading these prices accurately and responding to them will gain in wealth relative to those who act as if they were misreading these prices or ignoring them.

48. ~~Many~~ The human capital literature estimates some of these prices — e.g. the value of a college education to a person with certain characteristics. The point is, however, that these prices exist and exert an influence whether they are known or not. If the price is x, and one believes it is $y \neq x$ one is apt to make decisions that get punished. The groups, ideologies, etc. that prosper are the ones whose values are roughly in accord with the true, implicit relative prices. The ones that fail are those whose values are distorted relative to the implicit system whom the gods would destroy they first make mad.

49. Consider again the discussion of the propensity to save. This, we argued, was correlated with wealth redistribution. But which saving concept is proper? Consider S as conventionally measured in the national accounts with a ~~more inclusive~~ broader concept that includes changes in human capital, ~~etc~~. The latter is superior in the sense that it will be more closely correlated with changes of fortune than the former. This suggests a program in the revisions of national accounts. (See Nordhaus and Tobin Is Growth Obsolete? for some steps in this direction)

~~I~~ Fox, Social System Accounts, gives a framework for extending prices ~~beyond~~ in the non-market system

50. At this point it's worth to take stock. We have argued that the principle of natural selection applies to the entire social world, working by redistributing wealth among people with a range of characteristics. But doesn't this prejudge the question of the appropriate scale of measurement? As discussed above, in addition to wealth, other possible scales are: number of people, military power, religious influence, etc., some scales making more sense in some spheres, some in others (e.g. number of votes in the political sphere, number of dollars in the economic). Having given a taste of the natural selection argument in practice, we are now ready to tackle this question head-on.

51. ~~Different appropriate scales in different domains~~ A scale is appropriate to a given domain if equal weights are given to things of equal influence. Different scales then imply different relative influences in different domains, or different exchange rates. This suggests an analogy to international trade. Price ratios of two commodities differ from one region to another. Only in the case of no frictions to movement between regions would we expect them to be equal. (Frictions include both natural barriers — transportation costs — and contrived barriers — tariffs, customs, diverse standards, etc.). Now think of different domains — business, political, military, religious, familial — as being different regions, and we have an overall picture of the plurality of measurement scales. There are forces pulling them into conformity: all trade tends to equalize price ratios; but they ~~maintain~~ have a certain autonomy maintained by frictional barriers.

52 "Trade" among these domains takes the form of converting influence in one sphere into influence in another. Military power converts to political power:

campaign contributions, etc. Conversely, the political sphere regulates these, and all others. People of eminence in religion or intellectual life can influence other spheres, but within limits. (How many divisions does the Pope have? asked Napoleon). Friction includes the natural difficulties of portraying advantages from one field to another, together with a vast overlay of custom and legislation. (See Sorokin Social Mobility, _____, for the interconnections & influences.)

53. The important point is that the various scales of measurement fit together to form one overall coherent system, just as the local price structures fit into the world price system. Most of these prices are implicit, the actual prices of the business world forming a small but important fragment. It is this overall system of prices on which natural selection operates.

54. Over the centuries it seems likely that friction between these spheres have been diminishing, a trend fueled by the reduction of transportation, and especially of communication, costs. And of the various spheres it seems likely that the business world has grown in relative importance, taking over many functions of the others (with the deepening division of labor, the extension of the market, the influence of innovations). Not that alternative ways of organizing society, with alternative scales of measurement ("values"), are themselves subject to natural selection.

55. Incidentally, the use of an interactive trade picture does not ~~~~ entail the judgment that the business world is "top dog". The logic of the argument applies as well to the time of the Pharoahs or to the Mongol Empire. In the present world, however, it is not unreasonable to refer to "wealth" as the appropriate measurement scale, even though in actuality

56. After this extended digression, we return to the main line of inquiry. What about ~~newborn children~~ about intergenerational transmission? The individual persists but the living organism characteristics may be transmitted by heredity, but also by precept, imitation and tradition. The extent to which a characteristic ~~which can be~~ is so transmitted is the measure of its selective fortune, by the principle of the closed cycle. Families are not the only vehicles of transmission. Potentially immortal organizations, such as corporations, armies, states and churches, also serve. In these the "generation" cycle is marked by recruitment, promotion and retirement rather than by births and deaths.

57. Groups and collectivities, then, must be brought in to complete the discussion. As a first step consider groups as statistical aggregates, classified in a hierarchical pattern - say individuals grouped into families, families into communities, and communities into nations. The wealth - or other appropriate measure - of a group is then the sum of the wealth of its components ~~and the~~ The relative rate of growth of the wealth of a group is then a weighted average of the growth rates of its components, each of these a weighted average of its subcomponents, etc. (Furthermore, if differential growth persists, the weights themselves must shift in favor of the faster-growing components, simply because these come to be more and more in relative terms.) Natural selection operates at all hierarchical levels simultaneously, but in a bottom-up fashion, so to speak. Redistribution of wealth among aggregates is the resultant of differential growth among their components.

58. But this is not the whole story. ~~Their existence~~ Organizations are entities in their own right, having their own characteristics and their own size is appropriate

But why ~~should~~ do organizations come into being in the first place? How does the invisible hand of natural selection operate here? We have been arguing up to now as if the relative growth of wealth of entities depended solely on their own attributes. But it also depends of course on the attributes of the environment in which the entity finds itself, including the presence of neighboring entities ~~with~~ with their attributes and activities. Entities can enhance each other's success; for example, by entering into joint productive activities, or simply as a by-product of each others activities. Or they can harm each other, as there can be harm going in one direction or another in the other. (In the biological world these are the categories of symbiosis, mutualism, parasitism, etc.) In economics these are externalities, using the term in a broad sense.

59. ~~The presence of externalities implies that~~ But even so, there is a mutually-beneficial contract possible, each party modifying its conduct to take account of the effects on the other parties. But getting to such a contract may be infeasible, as a group ~~in which case there is 'market failure'.~~ (e.g., the parties agree to coordinate their actions to realize a common plan). This contract "internalizes" the externalities

60. Recall, however, that, even without conscious planning in that direction, natural selection shapes ~~actions~~ things so that people act — at least approximately — as if they were wealth enhancers. Similar forces arise when externalities are present. In the first place, entities that are

59. If the price system already reflects these externalities, no problem arises. If not (perhaps because there are missing prices), then ~~there is~~ 'market failure'.

mutually enhancing will do better if (by chance) they find themselves in each other's presence - hence they will tend to become associated. Entities which harm each other will do better apart, hence become "negatively associated". But more than this (Patterns of migration that attain these results will also be reinforced).

61. But more than this. Suppose an entity behaves in the way that "enhances its enhancers". Then it will be rewarded indirectly, and such behavior will tend to be stamped in. This process is itself self-reinforcing and tends to lead to a community of mutually-beneficial entities. Similarly, harming or destroying that which harms you gets reinforced. Revenge! "Vengeance" is a kind of wild justice" - (F. Bacon). A modified golden rule is supported by natural selection: do unto others as they do unto you.

62. What about the third case, where A harms B but B benefits A? This situation is unstable. Selection favors those who can moderate or reverse these effects. It is common in the natural world for parasitism to turn into mutualism. The domestication of plants and animals is an example - even the predator turns into even the protector).

63. The upshot of this argument is that there is a kind of implicit contracting that arises by natural selection, that supplements the explicit contracts arrived at by conscious effort. The process is, however, slow and important. Any institutions that arises to aid this process will itself be favored. Resolving disputes, keeping the peace, and enforcing

64. At a minimum, an organization involves some coordinated pattern of activities by its participants that persists over time — a vortex in the flux of events. Selection favors the development of stabilizing mechanisms that return the organization to the normal mode of functioning after disturbance; that is, it favors organizations that preserve themselves. One can think of organizations themselves as arising from the implicit contracting process discussed above. Once in existence they acquire a life of their own. It is less costly in general for existing organizations to acquire new functions than to build up new organizations from scratch.

65. We will not pursue this line of inquiry, but instead return to the foundations of economic theory. The basic postulate is that people are rational, in the sense that they optimize some coherent preference order among the opportunities available. Can something be said about the content of such preferences — e.g. attitudes toward the future, or toward risk? But first consider the basic postulate itself. Can this be derived from something even more basic — namely, that we live in a world in which natural selection operates?

Becker, Samuelson, Append. to [illegible], [illegible], the most critically important element

66. What are the alternatives to the rationality postulate? Consider the behavior of an animal, or of any physical object for that matter. It behaves in certain characteristic ways in certain situations — it has habits, in a broad sense of the term (CS Peirce). In game theory the concept of a strategy fits: It need not be derived from any preference order, and indeed when "nature" is a player, her strategy is not so derived.

67. Suppose then ~~in a world~~ that entities have a range of habits or strategies, some deriving from a coherent preference ordering, some not. How will the former reveal themselves in practice? By consistently driving in one direction, while the others bounce around at random, so to speak. They are the ions in an electric field amidst a crowd of neutral particles.

68. Points to be noted. First, coherence is a matter of degree: one's behavior can approximate to that deriving from a preference ordering. Second, ~~the~~ preference may be implicit - one acts as if one were consciously striving for something. The selective process will still operate in these cases.

This echo previous comments. The next point is new, and important. It is not coherence per se that is selected for, but coherence together with the positive content of the preference order. Captains Ahab had very coherent preferences; so did Macbeth and Richard III. Self-destructive preferences are selected against precisely because of their coherence. Indeed, more randomized, relaxed or less driven behavior would be an improvement in such cases. "A foolish consistency is the hobgoblin of little minds" (Emerson).

69. There is, then, a selection in favor of coherent preferences of the right sort - roughly, wealth-enhancing preferences taking account of the price system, explicit and implicit. This should not necessarily be interpreted as meaning that "economic man" must emerge. On the contrary, if organizations effectively neutralize externalities, then egoistic behavior merges into altruistic. This is the content of Becker's "rotten-kid" theorem in the family context

Becker A Treatise on the Family

And in these conditions

selection may well favor consciously altruistic preferences.
(Buchler on love) & Spinoza Social Stoics

Note the phenomenon of conscience (Freud's superego), an apparently inborn mechanism inhibiting the doing of harm to one's fellows. (It can be overcome with training).

70. How far will this tendency toward internal coherence be carried? To the point where the (marginal) gains from further coherence balance the (marginal) costs of internal coordination. The latter rise prohibitively past some point, so we may get only a general similarity of rationality in human behavior. The same "strain toward consistency" (Sumner, Folkways) also occurs at the supraorganic level, among the various competing institutions of society.

71. We now take a closer look at 'coherence'. ~~Broadly~~, this is persistence of preferences over time. Indeed, preferences that have large momentary fluctuations would yield behavior that is in effect random, and would be selected against relative to coherent "good" preferences.

Over a fixed period of time, however, would it not be more advantageous to change preferences to adapt to changing circumstances? Or, should distinguish between action ends and means here. A single fixed goal will involve twists and turns in following the route that attains it. What appears to be a change of preferences may actually be a change in tactics to match circumstances with a single underlying invariant preference order. Stigler & Becker, "De Gustibus non est disputandum" ___ carry this argument to its limit, postulating a unique universal invariant ordering underlying all human actions.

Does natural selection favor such a structure? Recall that

coherence per se is not selected for, but only coherence with respect to a "good" ordering. So any such universal ordering must have the property of being "good" under all circumstances — one must never be in a situation where one says (in effect) "wait, let's modify our goals here".

We believe, in fact, that there is such a universal ordering, but it would require much additional machining even to make intelligible the meaning of such a concept. ~~the sense in which it exists. There is a similar parallel here to the problem of the "universal prior" in the foundations of probability, to be discussed later.~~

77. We now turn to the problem of time-preference. We need not specify the exact form of the utility function. People with high time-preference favor high living over capital accumulation. (The artist Modigliani put this ideal thus: une vie brève mais intense.) Low time-preferrers will tilt in the opposite direction. It is the grasshopper versus the ant.

Natural selection favors low over high time preference, the lower the better.

77. This calls for several comments. First, one might think that too low a time-preference would lead to such a high rate of saving that the agent would perish of starvation. This is false, and results from thinking of time preference in terms of a utility function like $\int_0^\infty v(c)\, e^{-bt}\, dt$, c being consumption at time t, b the rate of time preference. Such a utility function is not in the spirit of human capital

would then be interaction of c(t) at different t, and in particular adequate food would be a sound investment.

74. A deeper issue is involved here – the relation of pleasure (and pain) to natural selection. The standard tradition in economic theory is to postulate an opposition between pleasure and success: the ~~pain of abstinence from consumption~~ the balance between consumption and saving is struck by the trade-off between the pain of abstinence and the reward of accumulation. (W. Senior; the tradition carries on through all the major theories of the interest rate – A Marshall, I Fisher, etc, – through the literature on economic development, as well as in the novels of Dickens and the films of Frank Capra)

From a broader perspective, however, pleasure and success must be positively correlated, not opposed. All organisms pursue what is pleasant to them and shun what is painful. ~~The resulting behavior~~ Those are selected for which the resulting behavior happens (by chance) to promote survival and progression. The outcome is that, on the whole, what feels good to the organism is good for it in terms of success: those organisms in which this correlation is reversed will have died out.

This argument derives from the logic of the natural selection concept, hence applies to its operation in human society as well. Our attention is drawn to those relatively rare situations of opposition between instinct and success (overeating, debauchery, drug-abuse), while the more common harmony between the two is overlooked.

The one thing that should be added is that the great human capacity ~~5. There is an exploration to the opposition of values.~~ for learning and adaptation means, in effect, a certain capacity to find pleasure in what is good for you, and to feel bad in doing what is bad, partially

35. There is an explanation for the opposition between the views of economists and the conclusions of the natural selection approach. In a period of major change, old successful habits may no longer work. There then arises the literally painful choice of trade-off between loss of wealth or imbalance (at least in relative terms) and the learning of new modes of living.

To be specific, the rise of the modern economy opened new investment opportunities in physical capital, and presumably tilted the optimal balance between consumption and saving in favor of the latter. A simple, lifestyle under old conditions now looks like "high living", overindulgence, and undersaving. This is a long-run disequilibrium condition that takes decades, perhaps centuries, to work itself out. One piece of evidence for this is the close positive correlation across countries between saving rates and growth rates. (Khoury --- Kuznets--

36. Another piece of evidence lies in interest rates. This requires some discussion. The natural selection approach suggests a virtual long-term tendency for the interest rate to approximate the growth rate of the economy. The reason is that if, for example, the interest rate exceeds the growth rate, an agent could enlarge his share of the total pie indefinitely simply by taking and holding a net creditor position in the market, while debtors would - ceteris paribus - find their share shrinking.

Now for qualifications. First, the capital-money market is necessarily beset with frictions and very imperfect. One would also have to cut through the relatively complications of risk, taxes and liquidity effects to isolate "the" pure interest rate. Second, to be consistent with the way natural selection works, the size of the economy must be measured in terms of wealth, not income, and must include wealth

embodied in human savings (by far the biggest component).

Notwithstanding these problems, one might hazard the guess that "the" interest rate has indeed stood above the growth rate fairly consistently over the last few centuries. If so, it would be consonant with the view of long-term disequilibrium suggested above. What is happening is that ~~business~~ firms have an incentive to borrow at these elevated rates because rates of return on business investment opportunities are higher still. These firms gain, and lenders also gain, a growing share of the pie, at the (relative) expense of those who save little, and even more so those who take out consumption loans.

(This argument supports Schumpeter's contention that the interest rate in a stationary state would be zero. Theory of Economic Development), ——
On the other hand, the idea of a stationary state is itself not congenial to the natural selection approach)

The consequent long-term transfer of wealth among these groups should eventually move the interest rate toward equality with the growth rate.

11. Next we come to attitudes toward risk. Will natural selection operate to transfer wealth systematically among those who are risk-averse, risk-neutral, and risk-loving? Yes — on the average and with the usual assumption of other things being equal — wealth moves toward those who are approximately risk neutral, and away from those who are strongly risk averse or risk loving. The key word here is "average": Expected wealth is maximized by the risk-neutral. The risk averse pass up favorable risky investments, the risk-loving go after unfavorable ones.

78. Some comments. First, this assumes no systematic misjudgment of probabilities associated with ~~attitude~~ attitudes toward risk - part of the ceteris paribus clause above. Second, people may be inconsistent in their risk attitudes. This raises no problems — the tendencies above apply to each particular act of risky investment, the overall result being an average over these. Third, things look different if you concentrate on dispersion of wealth rather than averages: the risk averse shelter the least dispersion, neutral next, and risk lovers the most. Indeed an irrational gambler will occasionally make a fortune, but will also more often go bankrupt.

79. One more big comment is called for. This conclusion appears to conflict directly with a large part of economic theory as enshrined in finance books — the whole theory of portfolio selection collapses, for instance. The usual ~~assumption~~ assumption is that people are normally risk averse.

There is an illusion here, and it arises from the importation of the capital market. The latter must be important even in principle, else one could walk into a bank and borrow unlimited funds at a fixed interest rate. (Instituting collateral requirements already destroys perfection) But an important market requires that we distinguish carefully between short-term and long-term wealth distribution. The latter is the one that is relevant for natural selection. Suppose you wanted to maximize expected wealth as of, say, ten years in the future, and suppose that bankruptcy would put you in a hole you couldn't easily climb out of (e.g. would destroy favorable opportunities that would otherwise exist.) Then you might not be wise to accept a 51-49 % chance of doubling your fortune tomorrow or losing it all — or even a 90-10 % chance. In short, long-term risk neutrality is compatible with the appearance of short-term

economies, given an imperfect capital market.

Ev. A theory must go beyond the purely verbal level we have been operating at so far. We have no room here to develop the formalities very far, but it is important to get a taste of what these models might look like.

We are dealing with physical measures and their redistribution over time. Formally, physical measures are indeed measures $(X, \underline{S}, m)$, $\underline{S}$ being a σ-field of subsets of a space X, and m a countably additive nonnegative set function with domain $\underline{S}$ (see Fadin, Elements of Space and Time, Chp 2 for technical background and interpretation of physical measures.) The interpretation is that X is a range of attributes, and $m(E)$ is the size — in an appropriate scale of measurement — of the entities embodying the attributes in subrange $E \subseteq \underline{S}$.

Probability measures, are, formally, the special case where $m(X) = 1$. A finite measure m (we consider only finite measures here) is completely describable by its total size or scale, $m(X)$ and its relative distribution p, defined by $p(E) = m(E)/m(X)$, $E \subseteq \underline{S}$. (Thus, p is a probability on the same domain as m.)

This decomposition is very important for two reasons. On the formal side, it allows us to make contact with probability theory. And in interpretation this whole paper has revolved around its conclusions in terms of changes in relative distributions. (A return to Ricardo, by the way, who thought he could say more about relative shares than about total size.)

81. Suppose then that at some time there is a cross-sectional physical measure representing a size distribution over a range of attributes.

The entities having attribute y will be multiplying or growing at an average rate $x = g(y,t)$ at this time t. (g is in relative terms - say percent per year; it can be negative). For example, consider an array of families or countries growing at different rates, size measured by total wealth.

§2. We now make the simplest possible assumption: that $g(y,t)$ does not depend on t. That is, the growth rate $g(y)$ persists over time, possibly different for different y's.

The space of attributes Y can be quite complex, and without further assumptions it is much easier and more elegant to deal, not with the measure m' on Y, but with the induced measure m on the space of growth rates itself, X. Now X itself is simply the real line (we take Σ to be the real Borel field), and, for any $E \subset \Sigma$, $m(E) = m'\{y \mid g(y) \in E\}$.

Take the interval (a,b) for example. $m(a,b)$ is precisely the total size (wealth, say) of those entities which are growing at a rate that is larger than a and smaller than b. This should really be written m_t since it varies with time. (The only case in which m_t does not change in time is when it puts all its mass on $x = 0$, the no-growth, no-decline case.)

§3. Call the initial time 0. By time t, whatever mass at x has expanded by a factor e^{xt} (Remember that x is a growth rate). Hence the total mass at time t is given by

$$M(t) = m_t(X) = \int_{-\infty}^{\infty} e^{xt}\, m_0(dx) \qquad (1)$$

Let p_t be the relative distribution at time t, so that p_t is the probability $m_t / M(t)$. Now for this whole set-up to make sense, $M(t)$ must be

finite, at least in some time-interval containing 0. It is clear from (1) that this is equivalent to P_0 having a moment generating function (mgf):

$$mgf(P_0, u) = \int_{-\infty}^{\infty} e^{xu} P_0(dx) \tag{2}$$

is to be finite for u in some interval containing 0. (It follows that all moments of P_0 exist). All further results refer to the class of measures m_0 whose relative distributions have moment generating functions; this is no restriction in practice.

84. The cumulant generating function (cgf) is the key to the following results. This is given by

$$cgf(t, u) = \log mgf(t, u) \tag{3}$$

where we have written $mgf(t, u)$ in place of $mgf(P_t, u)$. The cgf expands in an infinite series written as follows

$$cgf(t, u) = c_1(t)\, u + \frac{c_2(t)\, u^2}{2!} + \frac{c_3(t)\, u^3}{3!} + \cdots \tag{4}$$

valid for t, u in some interval containing 0. The $c_n(t)$'s are cumulants; in particular $c_1(t)$ is the expectation of P_t, $c_2(t)$ its variance, $c_3(t)$ its third central moment or skewness, and the remaining cumulants are more complicated functions of the moments of P_t. (H Cramér, Mathematical Methods of Statistics, Princeton UP, 1946, p 185-87)

85. Here is the key result

Theorem 1: $\dfrac{d\, c_n(t)}{dt} = c_{n+1}(t)$, $n = 1, 2, \cdots$, Also $c_1(t) = \dfrac{dM}{dt} \Big/ M$.

Proof: Expressed as an indefinite integral, (1) gives a representation of m_t, hence of P_t. Taking the mgf of this yields

$$mgf(t,u) = \int_{-\infty}^{\infty} e^{x(t+u)}\, m_0(dx) \Big/ M(t) \qquad (5)$$

Take the logarithmic derivative of (5) with respect to u, and separately with respect to t, and compare. This yields

$$\frac{\partial\, cgf(t,u)}{\partial u} = \frac{\partial\, cgf(t,u)}{\partial t} + \frac{dM}{dt}\Big/ M. \qquad (6)$$

Plug (6) into (4) and compare coefficients of u on both sides. The result follows. QED

86. This says in particular, that the relative growth rate of total mass equals the expected growth rate, the rate of change of the expectation equals the variance, the rate of change of the variance equals the skewness, etc. (The middle result is essentially R A Fisher's: the rate of evolution equals the variance. The Genetical Theory of Natural Selection, .)

Writing E, V for expectation and variance, we have $\frac{dE}{dt} = V$.

The average growth rate is thus constant if and only if the distribution is degenerate – i.e. everything grows at the same rate. In every other case $\frac{dE}{dt} > 0$: the average growth rate must itself grow.

87. Under what conditions will this redistribution process lead merely to a shift of location with no change in shape – that is, when will the P_t's be translates of each other? Another question that might be asked is:

when will the expectation change linearly with time, ($E(t) = a + bt$)? It turns out that both these questions have the same answer: precisely when the overall distribution is normal!

Theorem 2: The following are equivalent

(i) all the P_t's are translates of each other

(ii) expectation changes linearly with time

(iii) variance is constant

(iv) P_0 is normal

(v) P_t is normal for all t

Proof: (i) implies (v): Translation changes only the first cumulant c_1; hence c_n is constant, $n \geq 2$. By Theorem 1, c_n is actually 0, $n \geq 3$; this characterizes the normal distribution

(v) implies (iv): clear

(iv) implies (iii): The differential equations in Theorem 1 starting off with $c_n = 0$, $n \geq 3$ clearly remain in that state, all t. Hence also c_2 is constant.

(iii) implies (ii): from $dE/dt = V$

(ii) implies (i): $c_1(t) = a + bt$. By Theorem 1, $c_n(t)$ is constant, $n \geq 2$; this characterizes translation. QED

88. It is fascinating that a purely qualitative characterization, or a condition on one moment, determines the entire distribution. Note from theorem 1 that $M(t)$ is proportional to $\exp(at + bt^2/2)$ so the mass rises hyper-exponentially. (It will however decline at first if $a < 0$). Incidentally, degenerate distributions, as a limiting case of the normal, satisfy Theorem 2,

Equally fascinating is the case of a bimodal mixture of two normals. The redistribution process is linear, so each bump models independently, the one with larger variance moving faster and eventually swamping the other.

Exercise: Suppose the two bumps have equal variances. Then they remain equidistant. Why does this not violate Theorem 2, since the mixture is not itself normal?

Eg. Next, a harder problem of the same sort. ~~Suppose all the P_t's come from the same location-scale family. Then each P_t must be gamma, or reverse gamma, or normal. (Y has a reverse gamma distribution if $-Y$ has a gamma distribution)~~

Theorem 3: The following are equivalent

(i) All P_t's come from the same location-scale family

(ii) P_0 is gamma or reverse gamma or normal (Y has a reverse gamma distribution if $-Y$ has a gamma distribution; we also allow location-scale shifts of these).

(iii) all P_t's have this property (up to the point where the distribution blows up).

Proof: (i) implies (ii); Let Y have the distribution P_0; then $a(t) + b(t)Y$ has the distribution P_t, for some functions a, b ~~$a(s)=0$ $b(s)=1$~~. In terms of mgf's, this reads

$$mgf(t, u) = e^{ua(t)} \cdot mgf(0, ub(t)) \tag{7}$$

Taking logs, we find the following relation holds among the cumulants:

$$c_n(t) = b(t)^n c_n(0), \quad n = 2, 3, \cdots \tag{8}$$

Theorem 1 now yields

$$c_{n+1}(t) = n\, b(t)^{n-1} c_n(0) \frac{db(t)}{dt}, \quad n = 2, 3, \cdots \tag{9}$$

For $t = 0$ this reads (since $b(0) = 1$)

$$c_{n+1}(0) = n\lambda c_n(0), \quad n = 2, 3, \cdots \tag{10}$$

where we have written $\lambda = db/dt$ at $t = 0$. ~~$\lambda = 0$ brings us back to~~ the normal case. This difference equation system has the solution

$$c_n(0) = (n-1)!\, \lambda^{n-2} v, \quad n = 2, 3, \cdots \tag{11}$$

v being the variance of P_0. Substituting back into the cgf yields

$$cgf(0, u) = c_1 u + \sum_{n=2}^{\infty} \frac{\lambda^{n-2} v u^n}{n} \tag{12}$$

The normal arises from $\lambda = 0$, $\lambda > 0$ yields the cgf of a (shifted) gamma ~~with such parameter λ and degrees of freedom~~, and $\lambda < 0$ yields the reversed gamma.

(ii) implies (iii): The normal case is already disposed of. For the others, one verifies that

$$c_n(t) = c_n(0)(1 - \lambda t)^{-n} \tag{13}$$

$n = 2, 3, \cdots$ solves the differential equations of Theorem 1 with initial conditions (11). (By integration one also finds a $c_1(t)$ satisfying them.) (13) still yields a (shifted) gamma or reversed gamma, respectively.

(iii) implies (i): The gammas or reversed gammas must satisfy Theorem 1, since (13) [and (11)] for some λ, v. But (13) also satisfies (8) with $b(t) = (1 - \lambda t)^{-1}$, so these distributions are all in the same location-

90. Again we see that a qualitative characterization pins down the entire distribution family.

The gamma has the unfortunate property of blowing up at $t = 1/\lambda$, when the mean and all moments simultaneously become infinite (see (13)). Similarly, the reverse gamma descends from infinity at this time. Thus if the P_t's form a location-scale family but are also to be well-defined for all time, then they must be normal.

91. We finish with a few further results whose verifications are left as exercises.

Moment dynamics. Let $a_k(t)$, $b_k(t)$ be the k-th central moment and raw moment, respectively, of P_t. Then

$$\frac{da_k}{dt} = a_{k+1} - k \cdot v \cdot a_{k-1}, \quad k = 1, 2, \cdots, \tag{14}$$

where v is the variance a_2,

$$\frac{db_k}{dt} = b_{k+1} - E \cdot b_k, \quad k = 0, 1, 2, \cdots, \tag{15}$$

where E is the expectation b_1.

Let $E(t)$ be exponential: $E(t) = be^{ct}$ for some constants b, c. Then the P_t's are all Poisson-distributed, or reverse Poisson, with a change of scale. (The converse is also true)

92. The foregoing gives, as we stated, a taste of the kind of formal structure arising from natural selection. We will finish by giving a quick sketch of further developments.

The key assumption made above was that growth rates persist over

time. This is not literally correct and becomes a ~~poor~~ poor assumption for long term prediction. Thus the results above cannot be applied directly. (They are also sensitive to the exact shape of distribution upper Tails). But they are a good first step and suggest what has to be done next.

93. For example, if growth rates $g(y,t)$ drift over time, say according to an Ornstein-Uhlenbeck process, one already gets a fairly plausible model. (One implication is that the distribution of growth rates approaches normality no matter what it starts out as).

94. For deeper results, however, one must go back to the space of attributes y itself, and break down the growth process into its components - births and deaths, transformations from one attribute to another by migration, learning, imitation and contagion, etc., production transformations, etc. The very structure of attribute space arises from natural selection, a fact that gives a guiding thread on how to proceed.

Uncertainty | Probability

95. We cannot go into any further detail here. One aspect is so important, however, that it must be touched on briefly: uncertainty. The cognitive style of conscious beings, the manner in which they handle information and experience it itself an attribute, and one powerfully subject to natural selection. Cognitive rationality gets shaped just as behavioral rationality does.

96. People differ in the number of concepts they can distinguish, in the complexity of sentences they can understand, in their degree of perceptual discrimination. In general, the richer these categories, the greater is the variety of strategies one can follow. On the other hand, complexity has its costs, and requires that many

in §102

to §102 resources is devoted to information processing. Selection will favour the level of complexity that optimizes net benefits. (See Fodor the boundaries of psychology)

97. The modern apparatus of ("subjective" or "Bayesian") decision theory consists of three interlocking principles: (i) degrees of belief are represented as probabilities (ii) these are updated by conditioning on observations (iii) actions are taken to maximize expected utility (expectation with respect to current probability) (de Finetti, Ramsey, Savage, Lindley, Good, Raiffa, Schlaifer, Mersekak et al or e.g. Kyburg & Smokler) de Groot Optimal Statistical Decisions (~~Raiffa Decision Analysis~~ J O Berger Statistical Decision Theory)

These principles have been justified by arguments that plausibly must be obeyed by anyone acting in a coherent rational manner. For the natural selection approach this is not good enough. It must be shown that people who follow these principles will be favoured on the average — ceteris paribus — over those who violate them.

~~people following these principles be favoured?~~ Yes, but with some major reservations. First, the

98. ~~Will they be favoured?~~ ~~The answer appears to be yes, but with some~~ positive content of these principles counts as much as the form. With a disfavoured utility function — say a self-destructive one — the extra coherence imparted by these principles might make things even worse. (See above § ___). And the same applies to probabilities — any prior distribution is not as good as any other.

99. Those having distributions that are "close" to how the world really is will do better than those with less realistic distributions. In the short-run this may be a matter of luck, but in the long-run those having structurally

add to §99. Actually, two processes are at work simultaneously here. Individuals are undergoing changes in taste, fortune and numbers in response to the adequacy of their belief systems. At the same time, within individuals beliefs are being modified by experience, so that they "migrate" to more realistic distributions. Which of these processes is more important varies with circumstances. A famous remark by Planck is that the outlook of science changes by proponents of the older systems dying off. (e.g. Priestley and the phlogiston theory). If so, the external forces are dominant over the internal.

sound priors will tend to prevail.

100. Which brings us to the question of the correct prior. This question merits a treatise in itself. We confine ourselves here to a few gnomic comments.

First, priors should be conceived in a broad sense, not merely referring to parameters in a statistical model, but to the entire models themselves. These are just as subjective (if that is the right word) as the parameters in them (Bayes).

Second, a good prior should guard against all contingencies. It should have enough built-in flexibility ~~so that~~ to be robust against ~~to accommodate~~ any possible empirical data. (The situation is similar to that of an ~~ideal~~ ultimate lexical preference order — see § __ above). This suggests that the prior ~~is~~ is of a logical rather

~~Third~~,

than an empirical nature. But of course — since it is prior to experience.

~~But~~ We also know in a general way that the principle of natural selection applies, but not how it works in detail, prior to experience. This suggests that priors should be framed with that in mind. They do not dictate which world we live in, but they do give us probabilities that are structured to be compatible with natural selection.

~~101. The problem of appropriate scale of measurement is crucial here as well as in natural selection. The appropriate scale ~~First, it~~ provides units of equal influence, which are ~~and~~ also units of equal probability to which the classical "principle of indifference" applies. Further, probability at any time is approximately a weighted average over experience, the weights themselves involving discounts for dissimilarity and remoteness.~~

101. To identify probabilities with relative frequencies, or any physical measures for that matter, just invites confusion. Nonetheless, probabilities are intimately connected with physical measures. The latter, however, are not uniquely specified - this is the famous problem of the appropriate scale of measurement which keeps popping up - while probabilities are. Thus even if probabilities are a kind of expected relative frequency, a non-trivial transformation of the latter is required. The weights in the latter reflect both the disjunct between what is there and what we see - the problem of (note the word) selection bias - and also the distribution of caused connections over space and time; finally they provide "dissimilarity discounts".

The appropriate scale of measurement arising from natural selection enters in two ways. First it conditions the transformations just mentioned. Second it provides natural units for equal probability assignment (the classical "principle of indifference") (see Fader The Foundations of Probability _____ for further preliminary considerations).

102. These considerations provide constitute one of the major reservations to the use of subjective decision theory. (They suggest that the word "subjective" is misleading). The other arises from the cost of information processing itself. [§96 Here]

103. But complexity costs cut deeper. A probability measure is, in general, a very complicated object. Furthermore, if P is the "correct" probability, and Q is a probability "close" to P in an appropriate sense, then performance will downgrade only slightly if Q is used in place of P. If P is more complicated than Q, then switching to Q may reduce complexity cost enough to compensate for the slight loss in accuracy.

104. The foregoing paragraph contains in a nutshell nothing less than a program for the explicit revision of statistical inference, which we have called the post-Bayesian approach (Fedan & Ransom _____) The claim is that complexity is the only reason for departure from Bayesian inference (where the latter is specified to include the "objectively" correct prior). This leads to radical departures from current practice, both classical and Bayesian.

105. It should be stressed that the post Bayesian approach to statistics flows naturally out of the natural selection approach to social life. This suggests that people are in fact informal post-Bayesians in practice - in everyday life, in scientific research, etc. That is correct, and throws considerable light on such things as the role of theories in science, the use of idealized models, stereotypes, economy of thought, etc. We have no room to discuss this further.

106. To close the circle, note that the principle of natural selection itself is probabilistic. We have indicated this informally by referring to "tendencies" making "~~cet~~ ceteris paribus" assumptions, distinguishing "long-run from short-run" effects, etc. Thus there are two notions, each involving the other in an essential way. This suggests that their foundations can only be established simultaneously.

107. This concludes our exposition of the natural selection approach.

Conclusion

The natural selection approach leads to a surprisingly large and specific number of predictions, all in the form of long-run tendencies — toward low time preference, risk neutrality and general "rationality"; toward the equality of interest rates and growth rates; toward the growth of corporate organizations. And, in the cognitive sphere, toward Bayesian (and even post-Bayesian) inference.

On the other hand, it suggests that in general these tendencies do not go to completion. The world is full of novelty and uncertainty, frictions, and hidden costs, which enter into the selection process even if they are not in favor of them.

More generally, it gives a ~~certain~~ different perspective on the economic process. The distribution of income is seen as part of the allocative process, not something that can be manipulated in isolation from the rest of economic life (contrary to John Stuart Mill and ~~most~~ many of his successors). Consumption is seen as a form of investment — the human capital concept flows naturally out of natural selection. And so does capital theory in general, but partly in the form of a quasi-optimization that takes place behind the scenes with implicit prices.

The FOCUS program is to bring these ideas to a full-fledged reformulation of economic theory (organically connected to the foundations of statistics as well), and ultimately to supplant the general equilibrium approach with its profoundly misleading picture of the economic system.