Economic Thought 3.1: 21-41, 2014

From Rational Choice to Reflexivity: Learning from Sen, Keynes, Hayek, Soros, and most of all, from Darwin

Alex Rosenberg, Department of Philosophy, Duke University, USA
alexrose@duke.edu

Abstract

This paper identifies the major failings of mainstream economics and the rational choice theory it relies upon. These failures were identified by the four figures mentioned in the title: economics treats agents as rational fools; by the time the long run equilibrium arrives, we are all dead; the social, political and economic institutions that meet most urgent human needs most effectively could not have been the result of rational choice, but their 'spontaneous order' needs to be explained; human uncertainty and reflexivity prohibit a predictively useful rational choice approach to human affairs, and even limit its role in institution design. What unifies the perspectives of all four of these critics of neoclassical economics, however, is their implicit reliance or on need for a Darwinian perspective on human affairs.

Keywords: uncertainty, reflexivity, function, strategies, Darwin, Soros, frequency-dependent selection

1. Introduction

Rational Choice models (hereafter RCT or, for fun, Rat Choice), and the microeconomist's approach to employing them are in the ascendancy among social scientists. Political scientists have been expounding it for 25 years. In the last decade or so its application has extended to experimental social psychology and even neuroscience. Among economists rational choice models have been the only game in town for at least a century.

As is so often the case in the social sciences, this influence – hegemony might be a better word – is more a matter of fashion than achievement. It is mostly the result of theoretical tractability, mathematical elegance, and ideologically convenient rationalisation. It certainly is not owing to the predictive success of theories and models inspired by rational choice theory and the way in which economists employ it. Why is it that Rat Choice is so appealing despite the absence of much of a pay-off to using it?

The strength of the temptation to adopt the RCT approaches to explain human affairs is overwhelming. Introspection tells each of us, you and me, that we are rational creatures, who choose among alternatives on the basis of our beliefs and desires – in Rat Choice speak – our expectations and preferences. Similarly, we explain other people's behaviour by *interpreting* it, that is, making guesses about what desires and beliefs they must have had that worked together to bring about their behaviour. RCT is just folk psychology formalised. Since we can't shake folk psychology, we are suckers for Rat Choice. It has all the allure of our most psychologically satisfying stories. The stern admonition of science – that the mere reduction of feelings of curiosity is no mark of explanatory power – falls on our deaf ears. But we had better be able to give it up, if we want a useful social science.

The problems of RCT are four fold: Sen's problem of rational fools, Keynes' problem about the long run, Hayek's problem of spontaneous order, and Soros' problem of reflexivity/uncertainty.[1]

2. Sen's Problem of 'Rational Fools'

This is Amartya Sen's (1977) label for the charge that RCT is not only incapable of explaining a great deal of the most characteristic of human behaviours. What is worse, Sen argues, it would be foolish to substitute the choices RCT dictates for the ones we actually make.

Rat Choice is surprised by the degree to which people cooperate, mutually provide public goods – ones that are nonexcludable and non-rivalrous. The source of this cooperation Sen identified as their 'commitment' (1977). RCT almost always recommends free riding, and other strategies that unravel cooperative institutions. But these institutions persist. In order to reconcile itself with reality RCT must make unreasonable *ad hoc* assumptions about the shape of preference curves, and equally *ad hoc* ones about probabilistic expectations. RCT continues to struggle in the quest to explain away three facts: the frequency with which we succeed in providing ourselves with public goods; the frequency with which we honour norms that inhibit self-interest; and the net-costs we willingly impose on ourselves to police their violation.

Economists didn't start to take Sen's critique seriously until it began to emerge from computer simulations in game theory and human experiments in cognitive social psychology. Even now, these two sources of evidence are met with much resistance by mainstream economists. Sen's insight was the need for a richer psychology than Rat Choice allows, one that has room for commitment, among other features. A decade after Sen's original paper, Robert Frank (1988) advanced this insight in detail. In *Passion within Reason* he argued that there are a variety of crucial social-interaction problems people regularly solve in ways RCT cannot accommodate. Rational choice theory makes honesty the best policy, except where you can get away with dishonesty. In a straight contest between unconditional honesty and RCT's qualified honesty, the latter wins and unravels most of our social institutions. If we really were Rat choosers, we'd still be in Hobbes' state of nature. We aren't, so RCT must be wrong about the most fundamental facts of human psychology and social life. Sen, Frank, and a generation of cognitive social psychologists following them, have shown that human affairs are driven not by Rat Choice but by emotions harnessed to norms of fairness, equality, and real non-opportunistic altruism.

This work has combined with another line of research to undermine, if not unravel, RCT. Start with the most profound regress problem Rat Choice faces, one first identified by Sidney Winter (1975) and Jon Elster (1978): to make a rational choice you need to have correct expectations – accurate information about alternatives. Acquiring knowledge about alternatives costs resources and presents an optimisation problem. How much should you spend to acquire the information you need? This is a problem for RCT. How to solve the problem of figuring out how much to spend in a particular case? Use RTC? How to solve the problem of figuring out how much to spend to figure out the problem of how much to spend to acquire the information in the first place, and so on…

How, in fact, does this regress get cut short? Herbert Simon answered the question in the general case even before Winter and Elster articulated the problem. Humans don't

[1] No one should suppose that the argument to follow constitutes a full or even balanced account of the contributions to economic theory of the four figures. This paper is not a contribution to the scholarship of their work. I shamelessly pluck from their manifold contributions themes that work together in proving a telling account of the limits of received mainstream neoclassical economics.

maximise, as RCT requires: they satisfice. This is an insight Tversky and Kahnemann (2011) developed into a Nobel Prize winning insight about the role of heuristics in decision making, what Gigerenzer calls fast and frugal cognitive strategies for making choices. Humans are as fully committed to these cognitive norms as they are to the emotionally driven moral norms that prevent us from making rational fools of ourselves. (The relevance of this work to Sen's earlier independent insights was not lost on Kahneman, 2003, p. 152.)

RCT started out life as a psychological theory, an account of how people make choices. Just ask Jevons (1877) or Edgeworth (1895, 2003) or Wicksteed (1910). By the time Milton Friedman wrote 'The methodology of positive economics' he, and his more farsighted colleagues (for example Gary Becker), had recognised they needed quite a different rationale for their attachment to it than its adequacy as a theory of individual human behaviour. So they surrendered any interest in the project of explaining individual behaviour. They insisted rather that RCT was a powerful tool for explaining and predicting the behaviour of markets, industries, economies. This brings us to John Maynard Keynes.

3. Keynes' Problem of Long Run Equilibrium

Lord Keynes famously said, 'In the long run we are all dead.' The full quote is worth reproducing: 'The long run is a misleading guide to current affairs. In the long run we are all dead. Economists set themselves too easy, too useless a task if in tempestuous seasons they can only tell us that when the storm is past the ocean is flat again.' [1923, Ch. 3.] The line is probably his most famous in a life of *bon mots*. Keynes' point was that a theory that could enable us to predict only long-run outcomes was of little use, even if it was correct. One could go further and argue, with Popper and Soros, that there is no way to tell if such a theory is correct and so it is not really a scientifically respectable theory at all.

Adam Smith first hypothesised that Rat Choice would, via the invisible hand, exploit self seeking to produce an outcome that would make every one better off. Economists sought to convert Smith's hypothesis into a mathematical theorem for the next 150 years. They succeeded, but at great cost.

Several factors conspired with Smith's insight to drive RCT to an unshakeable commitment to the existence, uniqueness and stability of a market clearing general equilibrium. To begin with, the tools of differential calculus that Walras and other 19th century economists shared with physics (and later also evolutionary biology) made it natural to search for equilibrium solutions to sets of simultaneous equations. More important was the evident fact of price stability – change one price and the result is not a market spiralling out of control, but what looks like a smooth readjustment. Over-generalise and the result is that one sees equilibrium everywhere

Equilibrium outcomes, if we can get them, provide several things economists want: first, like all social scientists, economists seek explanatory regularities in the chaotic swirl of human affairs. If underneath the booming, buzzing confusion, there are equilibria among significant social forces, then there are generalisations about them that we can discover: to start with, the laws of supply and demand. If there are no equilibria, the prospect of uncovering laws governing human affairs is much reduced. Second, the equilibrium outcomes of the interaction of rational choosers are probably allocatively efficient – they direct inputs to their optimal use in meeting the real, attainable wants of economic agents. As such, equilibrium analysis provides guidance to policy – public and private – about how to arrange matters to attain the beneficent outcome Smith's invisible hand hypothesis envisions. Of

course, the fact that the equilibrium is unique and stable means that usually nothing need be done by government to attain it. *Laissez-faire*: left to itself the economy will get there.

As noted, because of their attraction to equilibrium analysis, economists spent about 150 years trying to make a mathematical theorem out of Smith's metaphor of the invisible hand making us all better off through the self seeking of each of us. They succeeded and congratulated themselves by awarding one of their earliest Nobel Prize to the economists who did it (Arrow and Debreu, 1954). What they proved was a weaker result than they wanted, but it was the best Rat Choice could do: In a perfectly competitive market of rational agents, the prices of all goods and services will arrive at a unique stable 'general' equilibrium that is allocatively (Pareto-) efficient. But the cost of providing the proof was draconian: the existence of the unique stable allocatively efficient equilibrium required so many false and impossible assumptions (five of them) besides those of RCT that the proof mainly explains why *actual* economic outcomes are neither welfare maximising nor allocatively efficient.

Observation suggests strongly that human affairs are rarely in equilibrium. At most patterns in human affairs are very local, temporary equilibria – broken up annually, monthly, daily, indeed sometimes hourly. Rat Choice theorists bid us to disregard or deprecate this inconvenient fact. Most changes, they argue, are temporary stochastic departures from the unique stable long-term equilibrium they have proved to exist. These temporary departures are fated to be cancelled out by equally random movements back towards the unique stable equilibrium.

It's rationality of the sort RCT asserts to be widespread that assures us of the existence of this happy outcome. Rational agents choose under conditions not of certainty, but of risk. Risk is the condition under which agents can assign probabilities to all alternative outcomes in accordance with three relatively weak axioms of probability theory. It's a direct implication of RCT that agents optimally allocate resources to the search for information they need. Recall the regress problem of Winter and Elster. An equally obvious implication of RCT is that they employ this information optimally to maximise their expected utility. If agents obey the laws of probability, new information is always incorporated into their expectations in accordance with Bayes' theorem. Two startling implications follow from these assumptions: first, under these circumstances, in the long run, everyone's probability assignments will converge on the same subjective probability assignments to all alternatives no matter where they start out. Second, because the errors people make are randomly distributed on a bell-shaped curve, the average value of their probability judgments will always be close to the objective (i.e. correct, actual) probabilities. Every individual's expectations may be wrong all the time, but the aggregate average of these wrong expectations will be the right expectation. Thus, in the long run, the market's expectations about outcomes are always correct. This is the basis of the continued confidence that markets really are allocatively efficient because they are informationally efficient – the efficient markets hypothesis.

Keynes' famous epithet reflects several criticisms of economics' attachment to equilibrium. Most have taken 'In the long run we are all dead' to mean that *laissez-faire* solutions to economic problems arrive too late to help the people who need it, if they arrive at all. But Keynes famously argued that instead of one unique, stable market clearing, allocatively efficient, equilibrium, there are many local equilibria which are far from allocatively efficient, and that governments can actually move economies out of these local equilibria in the short and medium run. Finally, and most subversively, he gave a reason to think that economies are never really on the move towards the long-run equilibrium of which the rational choice theorist dreams. The reason he gave was that humans often cannot act in accordance with the requirements of RCT. This fact deprives economics even of a theoretical assurance of the existence of an equilibrium that observation never detects.

The heart of Keynes' critique of mainstream equilibrium thought was his diagnosis of what RCT gets wrong and why. The diagnosis was perhaps not completely original with Keynes (Frank Knight (1921) prefigured Keynes and George Soros (2003) came at the same point, perhaps independently). It begins with a distinction between risk and uncertainty and explains the crucial role of money in the economy. Agents face conditions of risk if the alternatives facing them can be assigned probabilities that behave in accordance with the three axioms of probability theory, and which they can update in accordance with Bayes' theorem. Agents face conditions of uncertainty when it is impossible to assign probabilities to alternatives in this way.

Equilibrium economics is predicated on two assumptions: that risk is the rule and uncertainty the exception, and that probabilistic expectations of agents are distributed normally around the objective probabilities of events, cancelling out individual errors and making markets allocatively efficient. For this reason, *there is no room in mainstream economic theory for the existence of money*, a remarkable fact on which most microeconomists are silent.

In fact, humans generally face uncertainty, not risk. The difference between risk and uncertainty is the difference between the casino – in which all probabilities can be calculated, and living on an earthquake fault-line where no one has the slightest idea when the big one will hit. Exogenous – outside – events, big and small, intervene in almost all social processes almost all of the time. Agents don't, can't probabilify these events. Even if there are equilibria around which outcomes are moving, these exogenous events destroy them, substitute others, and destroy them too, in a continual process. It is a process that another great opponent of equilibrium thinking, Schumpeter (1942), called 'creative destruction,' though he should also have recognised the process of 'destructive destruction.' In the end, it's this continual destruction of general equilibrium trajectories before they reach their end points, that Keynes's pithy observation draws our attention to.

Uncertainty was the key to Keynes explanation of why money exists and what its real role in an economy is. The prevalence of uncertainty is one reason humans employ cognitive heuristics in decision making, instead of the RCT tools suited only to quantifiable risk. Uncertainty, and the way humans deal with it, produces multiple stable and unstable local equilibria, none of which are allocatively efficient, and all of which obstruct the economy's march to the mainstream economists' nirvana: general equilibrium.

So, social scientists', especially economists', commitment to equilibria is equal parts: wishful thinking about the invisible hand, attraction to mathematical elegance and tractability, and overconfidence in the rationality of human beings. It's a recipe for retrospective rationalisation and prospective impotence. There are, however, many *local* equilibria, some relatively long lasting. The existence of money is one such, and it raises another fundamental problem for RCT.

4. Hayek's Problem of 'Spontaneous Order'

It's not just that Rat Choice does not explain several of the important features of human life. *It cannot do so*. This Nobel Prize winning insight is due to an economist revered by many mainstream (i.e. Chicago-school) economists, Friedrich Hayek.[2] Here are three examples, all

[2] Hayek's insight about the problem of spontaneous order and its explanatory solution did not of course prevent him from embracing RCT as an account of individual economic choice. His commitment to RCT of course endeared him to his colleagues at the University of Chicago, the epicentre of such commitments. Following Hayek some Chicago economists early invoked Darwinian processes to explain how rational choice is imposed on individual choice and eventuates by aggregation into unintended spontaneous but efficient outcomes such as the market. More lately

from economics, where you would assume rational choice has an explanatory role: the firm, money, and the price-system. Each of these institutions fulfils an important *need* individuals have. None emerged from a rational choice process. Hayek's problem was to figure out how they could have emerged and why they persist. He called them cases of 'spontaneous' emergence, persistence or order. But that is just to label the problem, as we'll see.

In the case of the firm, the human need is to solve the transaction-cost problem, as Ronald Coase (1937) first noticed. Without a solution to this problem, the division of labour must come to a standstill and with it almost all the productivity increases humans have contrived since the Middle Ages. No rational agent recognised what the problem everyone faced was. No one decided to invent the firm in order to solve this problem. It emerged 'spontaneously' to 'order' exchanges between individuals that faced a transaction-cost problem. The firm is an example of 'spontaneous order.'

Money solves the biggest problem of barter: what the economists call 'double coincidence of wants'. Without money, if I want oranges and have only banana, I need to find someone who wants bananas and has oranges. What's more, if we can't divide and store bananas and oranges, I'll need to find someone who wants to trade in exact whole numbers of bananas and oranges that match up with the amounts I am prepared to trade. This is a problem that becomes intractable very early in human exchange. How does it get solved? Several times in distant cultures the same solution was hit upon: the emergence of a commodity with common features: portability, divisibility, durability, utility or widespread desirability, and short-term limits on its quantity. When money emerged no one around consciously recognised that money would have to have these features. (Something Hayek (1978) also noted.) No one rationally adopted some commodity in order to solve the problem of the double coincidence of wants. Rat Choice can't explain how it happened

The emergence of money requires that agents solve another problem: one of coordination. Sooner or later they must all converge on the same commodity. People must solve a 'common knowledge' problem. Somehow each agent must be willing to adopt a certain commodity as money, and must come to believe that everyone else will adopt the same commodity, and must believe that everyone else will be confident that every other agent has adopted the same commodity. You can see that this is a set of problems that can't be solved by individual rational choice, that were not solved by some explicit social contract. The institution of money is another example of order emerging without anyone intending it, or taking steps to bring it about. Of course to say money emerged spontaneously is simply to label the problem and exclude an obvious Rat Choice explanation of how it emerged.

The third example, Hayek's (1945) example, the system of market prices, is the most important – but the most difficult to understand – of these problems of spontaneous order. It was this realisation that earned Hayek his renown among economists.

The unsolvable problem of socialist central planning is informational. Central planning faces the mathematical problem of converting a list of available inputs and a list of desired outputs into a list of production orders, and then continually updating this list as input availability changes and desired outputs change. Central planning faces the further problem of sending information about each of the changes in inputs and outputs, only to those who need to have this information, in order to change their production plans. The central planner can't send the changes to everyone: we'd have to spend the better part of every day just trying to find the information we would need from a daily massive data dump. But the central planner can no more figure out to whom exactly to send the updated information than it can figure out the initial production order. These are all what mathematicians call NP-hard

---

Chicago economists have substituted rational expectations and representative rational agents operating in the market to effect efficient market outcomes, while foregoing any account of how markets could have arisen to begin with.

problems ('nondeterministic polynomial-time hard problems'). There is no known algorithmic, computerisable solution to such problems, and a good chance than none exists. Yet the problems are all solved all day and every day, instantaneously, by the system of market prices. The market price system is an information storage, retrieval and calculation system – a vast virtual computer – that provides the closest approximation to mathematically correct solutions to the central planners' calculation problems and at no cost whatever.

The market price system performs a function indispensable – not just to modern life – but to all human life beyond the Pleistocene. It is a function meeting a need that cannot have been foreseen by humans, no matter how rational; it is a solution to that need that no human or coalition of humans could have intentionally contrived. Indeed, it is a solution that rational choice would have led individuals to try to undermine or subvert in their own interests. It is a solution to the problem people face that is so ingenious it automatically and successfully responds to such subversion attempts.

The market price system operates continually to meet a need that no human or set of humans could, by intentional and deliberate action, fulfil, no matter how rational they are, and no matter how powerful and expensive their information storage, retrieval and computational resources are. And the market price system emerged, like money, spontaneously, independently, repeatedly and without malice of human forethought, throughout human affairs, across the globe.[3]

These three examples of spontaneous order highlight the economist's version of a problem facing all social sciences, a problem Rat Choice is incapable of dealing with. The problem is deep, and pervasive.

First, why pervasive? Because the three cases identified here are just the tip of an iceberg. Almost every phenomenon of interest to the social scientist manifests the problem of spontaneous order. Almost every human institution, almost every long-standing social practice, almost every organisation of individuals, and of their coalitions, fulfils a function, solves a problem, confers a benefit or advantage on something or other. Think of any of the variables of macroeconomics: the interest rate, the rate of inflation, the money supply, the fiscal deficit. These are institutions, or the properties of institutions with functions.

Unbeknownst to the agents who participate in them, the macroeconomic institutions fulfil important functions for the economy, for industries, for markets, and for their individual participants. Most of these functions are unrecognised, unintended and unforeseen most of the time, by most of their participants. But the functions fulfilled by these institutions are crucial to their emergence, persistence, change over time, and to their eventual disappearance. In this respect economic institutions are no different from almost all the political, social, cultural institutions, organisations, and practices that order the behaviour of individuals and groups. That means all social sciences face the problem of spontaneous order, not just economics. No social institution, organisation or practice could exist long enough even to be noticed by social scientists unless it had a function. Since most of the functions of most of the institutions that make human affairs possible go unnoticed, as well as unintended and undesigned by their participants, they all raise the problem of spontaneous order that Hayek noticed and that confronts the economist.

Almost everything of interest to the social scientist has a function, usually unintended and unforeseen and continually unrecognised. This observation was recognised, dimly and imperfectly, by functionalist social scientists, like Durkheim (1895) and Parsons (1951), in the first half of the 20th century. They recognised that most functions of most institutions escape

[3] There are of course domains, in which rational calculation is required to design 'incentive compatible' institutions, for example electronic bandwidth auctions in which there are a small number of bidders and strong pay-offs to collusion. Designing such institutions requires designers assume individuals are rational egoists and organise the institutions to defend themselves against undermining by such egoists.

the notice of their participants. These they called 'latent' functions, by contrast with the 'manifest' functions recognised and often designed, intended, and sustained by conscious deliberation and perhaps even by something approaching rational choice. The written US Constitution has manifest functions, some of them quite different from those of the unwritten British constitution. The former fulfils important latent functions not intended and not widely recognised. One reason British people are in certain respects – e.g. health care – far better off than American people is because the British constitution fills latent functions the US constitutions does not.

The 20th century functionalists were right about the functional character of almost all social institutions. But a serious oversight in their analysis condemned it to implausibility, and it went into eclipse long ago. The simple error functionalists made, which made their view sound so implausible, was to mis-identify the *beneficiaries* of the functions that institutions, practices, and organisations fulfilled. They assumed, quite myopically and wrongly, that the function of institutions, practices, organisations, was to fulfil the needs of people, of human beings. But it was obvious that many institutions, practices, organisations are in fact are harmful to people, confer no net advantage on them, for instance most religions, or Chinese foot-binding, or tobacco smoking. This Panglossianism about all social institutions made functionalism a laughing stock when it was not pilloried as an invitation to complacence and conservatism: if almost all human institutions fulfilled functions for us, then it is tempting to reason that we should not change them lest we deprive ourselves of the benefits they confer on us. Whence the charge of complaisance.

Only in the late 20th century did it become apparent that in these and other cases, a change in perspective – a *Gestalt switch* – would enable us to see what was not previously apparent: the relevant beneficiary of those features of institutions, practices, organisations harmful to people were the institutions, practices, organisational structures themselves, that parasitise people, that treat people as niches, environments to be exploited. Think of people as the environment and think of types of institutions, practices and organisations as the things that survive, replicate, and spread or recede and become extinct owing to the degree their features *exploit* human characteristics. Then the functionalist perspective becomes irresistible: many socially significant institutions, practices, organisations, confer huge net benefits on people – money, the firm, the market price system. Many others confer huge net harms on people, but in so doing ensure their own persistence – think again of foot binding or tobacco smoking or heroin addiction. Other institutions confer benefits on some people, and harms on others – slavery for example. Most institutions – religions, for example – confer a mixture of harms and benefits on different mixtures of persons over time.

One way to effect the gestalt switch necessary to accept thorough-going functionalism about human affairs is to employ the game theorist's notion of a 'strategy'. A 'strategy' is simply a rule, norm, procedure, of the form 'Under condition X, do Y'. Strategies may be reflexive or voluntary, moral, or ritual, matters of fashion or style, short-lived or not, obligatory or optional, complex or simple, consciously followed or not, beneficial to the agent employing them or harmful to him or her. People's behaviours are determined by strategies they internalise. These strategies are traits, like left-handedness, or speaking French, or wearing miniskirts, that can come and go. They are acquired, by social learning, by imitation, by unconscious classical and operant conditioning, and transmitted from person to person, and they interact with other strategies, cooperating with them, competing with them, subordinated to them, or subordinating them. Human social institutions, from a book club to Feudalism, are nested sets of coordinated strategies. Think of practices like patrilateral cross cousin marriage or purdah or the incest taboo. Think of organisations like the free masons or the parish council. 'Human affairs' is a matter of nested institutions, organisations,

practices, all composed of the strategies individuals employ. Then there are the strategies each individual employs to navigate through these institutions, organisations, practices. The institutions, organisations, practices have functions. They thrive or perish depending on how well the strategies they impose on people enable the institutions, organisations and practices to fulfil these functions for their beneficiaries – often themselves.

To repeat: almost everything of interest to social scientists has a function, fulfils a need, confers a benefit, or is the direct consequence of something with a function. The pervasiveness of this feature of human affairs makes Hayek's problem of spontaneous order much more serious than even he supposed. Recall that problem: how do cases of spontaneous order emerge and persist? Rat Choice is not an option here. The institutions, practices, organisations that mainly interest us in social science have functions, attain ends, goals, confer advantages. Yet almost none of them were designed by men or gods. Hayek's achievement was to show that the economically most important of them *could not* have been the intended, designed or foreseen result of intentional action by rational agents. The problem of spontaneous order is: where did all these apparently well designed (but not actually designed at all) institutions, practices and organisations come from and why do they persist?

Hayek had the answer.

5. From Hayek to Soros, via Darwin

Wherever the appearance of design is to be met with, in nature or nurture, in the biological realm or the social realm, on the watchmaker's work bench or nature's laboratory, the source is never real foresight, but always tinkering – blind variation and environmental filtration. This is a lesson already well established in biology. But it is equally in force for the social and behavioural sciences. The lesson is resisted only because of the same mistake that obstructed the functionalist social scientists, combined with an equally egregious error of supposing that Darwinian processes are restricted to the domain of genetically hard-wired functions.

Every significant (unintended) feature of social life that has a function (and they almost all do) has been built by a Darwinian process. Why? Because there is no other alternative. Long ago, science – especially physical science – excluded the possibility of real goals, ends, purposes in nature. It revealed that future states couldn't reach back into the past and pull events in its direction. The Aristotelian conception that purposes explain anything at all has been progressively read out of every scientific domain until it is left only in folk psychology and its Rat Choice formalisation. (The idea is that correct expectations about the future, together with attainable desires about the future produce achieved futures. These futures are therefore part of the explanation of the processes that bring them about.) Purpose is hard for *hoi polloi* to shake. Part of the grip of RCT trades on its formalisation of common sense purposive explanation.

It was another 200 years after physics expunged purpose from its domain, before it was banished from the biological realm. Until 1859 the hand of the benevolent, omnipotent designer, God, was the favoured – indeed the only – explanation of the appearance of design in the domain of living things. This, in effect, made biology incompatible with physics and chemistry, sciences that had no need for the deity. The solution to the inconsistency was Darwin's discovery of the purely natural causal process that produces the appearance of design while showing that the appearance is not a reality. He showed biology had no more need of a deity than physics. There are no purposes at work in the biological realm. All

biological functions are just adaptations produced by blind variation and passive environmental filtration.

There is no underestimating how powerful this result was – and remains – for reordering all the nonphysical sciences. It reconciles them with the most fundamental facts about nature physics discovered – that there are no purposes, goals, ends, designs out there waiting to be realised and playing a role in bringing about their realisers. Once Darwin showed how purely causal processes *could* bring about the appearance of design, biologists set about showing exactly how causal processes *did* bring them about: a 150 years of this work produced genetics and the molecular biology of the gene, protein, enzyme, neuron. It made thoroughly mechanical reproduction, respiration, development, and cognition.

If social processes, and all the interesting aspects of human affairs fulfil functions – for us, for themselves, for something else – if they show the appearance of having been designed to deliver some benefit to something or other, then they have to be the product of a Darwinian process of blind variation and passive environmental filtration. Why? Because that is the only way things with functions, adapted traits, can come about.[4]

Recall the suggestion above that we need to treat human institutions, groups, practices as packages of strategies employed by people. The features, characteristics, traits of institutions, organisations, practices, are composed of these packages of strategies. At the basement level of individual agents, the strategies they employ are their own individual adaptations – traits that have pay-offs for them or for someone else that result in these strategies persisting – being used over and over, and spreading – by imitation or instruction, reinforcement or coercion, or receding by operant punishment, or legal sanction, etc. Individual strategies are traits of individual people. Their cognitive equipment is what passes them on, modifies them. The human environment, including all the nested packages of strategies that constitute institutions, organisations and practices, select among these strategies in ways that result in the emergence, persistence, and change – rapid and slow – of individual strategies, and nested groups of them.

Here game theory (the scientific study of strategic interactions) is a pedagogic help. Types of games are characterised by pay-offs and strategies available to be played. In the prisoner's dilemma, one can cooperate or defect. The rational strategy is to defect. Bigger institutions, practices, organisations are composed of strategies played by their smaller component institutions, practices and organisations, and in the end, by the individual participants, people, whose interaction produces these larger social units and their features, as the unintended, unforeseen result of their individual strategies. (Students of the philosophy of social science will recognise this claim as the thesis of methodological individualism, a thesis familiar to economists and Popperians.)

Given the pay-offs – costs and benefits – that institutions, practices, organisations impose on the use of various strategies, there is selection for those that do better, regardless of whether the people who play them recognise the pay-offs or are motived by them. Who decides on the pay-offs to various strategies? Almost always no individual does. It's nature that decides in the earliest, simplest institutions. For example, the strategies of males and females in the hunter-gatherer domestic division of labour, were selected for by their impact on off-spring survival. As institutions, practices, and organisations emerge, they increasingly set the pay-offs to participating strategies, to other strategies that may undermine or unravel

[4] At this late date it may not bear mentioning that there is no commitment to secular 'progress,' continual improvement, or some biological, social or moral betterment of later evolutionary outcomes over earlier ones. Darwin's insight was that all adaptation is local, that today's adaptation in the current environment will be tomorrow's maladaptation in a different environment, and that there is absolutely nothing morally better or improving about increased fitness. Sen (2002) effectively criticizes mistaken conceptions about Darwinian progress. But neither Darwin nor latter-day exponents of Darwinian cultural evolution make any such commitments.

them, and to strategies that compete with them. Of course institutions, practices, roles, spawn new institutions, groups, and practices, often by combining into larger units with new functions, and also by selecting for smaller component units, cooperative ones, and exploitative ones.

In biological adaptation by natural selection there is a well-understood process that operates through adaptation to produce both increasing complexity and diversity. The same forces operate in human affairs. The persistent processes of random variation and passive filtration produce complex and diverse social institutions, practices and organisations – ones with new adaptations, new functions.

Think about strategies that people employ as the traits out of which the traits of all the rest of social processes are composed. Think about these strategies and packages of them that constitute institutions, as having traits that enable them to colonise humans, to spread, to compete, to cooperate, to synergise and support, or subvert and exploit one another.

Hayek realised that this Darwinian approach to the domain of the social sciences is not merely a useful metaphor, a suggestive trope, a way of looking at things we had not noticed before. There are two reasons a Darwinian cultural mechanism is an unavoidable reality in human affairs, and therefore an indispensable tool for understanding them.

First, as noted above, a Darwinian approach to human affairs is 'the only game in town'. We know, with great confidence, that there is only one way that functions, adaptations – the appearance of purpose or design, can emerge in a world like ours: the way Darwin discovered. Unless we are prepared to deny that social institutions, practices, organisations, have functions, we are stuck having to apply Darwin's discovery to human affairs. Applying the Darwinian approach requires a lot of work. We cannot expect to simply apply the details of how Darwinian selection works in biology to human affairs. Darwinian cultural processes will not employ the mechanisms (especially the genetic ones) Darwinian biological processes do. The Darwinian approach to human affairs does not require genetic determinism of human differences, the innateness of important human traits, the hard-wiring of human culture, or an evolutionary psychology about human cognitive and emotional traits. Blind variation and passive environmental filtering are as much matters of culture as matters of nature.

Second, and for present purposes more important, recognising the Darwinian character of all processes in the domain of the social sciences helps solve the three problems that daunt RCT: the problem of rational fools; the problem that by the time equilibrium arrives we are all dead; and the mystery of spontaneous order. And it does all these three things in a way that vindicates an important insight George Soros has been articulating in lectures, papers and books for about 30 years.

Sen's problem of rational fools is that rational agents won't provide themselves with the fruits of cooperation since cooperation is just a set of strategies that the rational agent can free ride upon. Even when cooperation is a Nash equilibrium strategy, the rational agent will continually seek opportunities to change the pay-offs, take advantages, free-ride, secure rents. Enduring cooperation robust enough to withstand threats requires commitment, usually driven by emotions, that override rational choice. This is the lesson of much of the research on the evolution of cooperation: research that explicitly employs Darwinian dynamics – mechanisms of random variation and natural selection to identify strategies that maximise fitness. These strategies are rarely ones RCT recommends.

Besides Sen's rational fools problem there is the calculation/implementation problem that faces RCT. Besides the regress in applying it to decide how much to invest in acquiring information, the absence of risk/presence of uncertainty in choice, make Rat Choice often maladaptive as a real-world decision strategy. What we employ instead are heuristics – rules of thumb and rules of thought that provide quick and dirty solutions to real-time problems. Our

hardwired cognitive strategies are the ones natural selection found through geological eons of tinkering. Our learned cognitive strategies are ones Darwinian cultural selection has produced by trial and error and transmitted by teaching. In Herbert Simon's (1955) terms, humans don't optimise, they satisfice, just as Darwinian natural selection – the satisficing process *par excellence* – would have them do.

Darwinian cultural selection enables us to fully understand both spontaneous order and the real role of equilibria in human affairs. Once we see how it does these two things we will be in a position to appreciate Soros' conception of reflexivity and how pervasively it influences all aspects of human affairs.

Turn for a moment to Darwinian processes in the biological realm. Here there is a considerable role for equilibrium analysis and it is an important tool in both the mathematical modelling of biological processes and in the explanation of biological regularities or laws.

An illustration will help greatly. It is a regularity that in almost all vertebrate species, indeed in almost all sexually reproducing species, the sex ratio is 1:1 – 50 % males, 50 % females. That there is almost always, almost exactly the same number of men as women, was long treated as strong evidence of the benevolence of God. The 20th century British geneticist, R.A. Fisher, showed that the 1:1 sex ratio generalisation is a stable equilibrium which results from a Darwinian process of blind variation and passive environmental filtration. Women have varying hereditary predispositions to give birth to males or to females. Whenever the sex ratio departs from 1:1 in favour of more females, those mothers who disproportionately bare male children will have more and fitter grandchildren, since their sons are scarcer relative to females and can be choosier. More grandchildren carrying genes that favour having boys results in more boys and so moves the sex ratio back to 1:1. When ratio begins to favour males over females the same process in reverse shifts it back to 50% of each. Whence the stable equilibrium and the biological law that sex ratios remain the same and in balance.

Actually it's not a law, because it is false for a small number of species. In humans the long run equilibrium sex ratio at birth is 1.05 to 1, slightly favouring male births. Why? Because boys' mortality rates are higher than girls', or at least were higher in the environment that selected for homo sapiens. Darwinian natural selection had to fine tune the sex ratio to make it 1:1 at sexual maturity. Doing that required more boys at birth than girls. Additionally, there are several species of insects in which the sex ratio is heavily biased towards females.

How does the fine-tuning on the one hand, and the cases of complete abrogation of the apparent biological law on the other, happen? There are many examples of fine-tuned equilibria in the biological realm. The Fisher sex ratio is but one very easily understood example. Any two traits of organisms that work together well, such as flying and good eye-sight in birds, or symbiotic traits – cleaner fish and cleaned shark, remain in equilibrium for a long time once established. Competing traits do so as well: think of predator and prey species that maintain a long term relationship, neither becoming too rare or too numerous to make the other extinct. And of course the traits of parasites and hosts show the same process of fitness maximising equilibriation over time: a parasite so virulent that it kills a host before it can jump to the next host becomes extinct, leaving the less virulent form to spread, and so it goes until parasite and host can just live with each other, as in the Simian version of the AIDS HIV.

These are all cases of 'local equilibrium'. But underneath the appearance of changelessness, balance, calm stasis in the fixed relationship between traits, there is an underground guerilla war taking place. Each population of traits among cooperative ones or competing ones, in the same species or in different interacting ones, is in constant variation – random mutation. Almost all of these mutations are, of course, unfavourable in their local environments – the combination of the competing or cooperating trait and the rest of the niche

in which the mutation finds itself. Very rarely, one of these random variants confers an advantage to a trait and the organism that bears it. The new trait enables the organism to exploit its hitherto cooperating partner, or to suddenly achieve an advantage over its hitherto equally fit competitor. At this point, the local equilibrium begins to break up. The only thing that can prevent it from eventually unravelling completely is a counter-variation in the traits that hitherto cooperated or competed effectively enough: a counter-variation fit enough to preserve the equilibrium. For obvious reasons, the process that ensues when local equilibria break up, are called *arms races*.

The natural history of the planet is a history of local equilibria broken up and followed by arms races in the biological domain. The local equilibria are imposed by natural selection operating through very small variations over enormous time scales in very slowly changing environments. They produce regularities that short-lived creatures might mistake for fixed laws of nature: giraffes have long necks, polar bears are white, Australian mammals are marsupial. And when environments are constant for long enough some of the regularities reflect equilibria that approach fitness optima very closely, for example the 1.05 to 1 sex ratio in humans.

Sometimes these local equilibria last for only a brief time and are broken up quickly. The best examples of such rapid evolutionary change, where arms races are the rule and equilibria are the exceptions, is the evolution of bacterial drug resistance. The AIDS virus varies so rapidly, owing to instability of its RNA genome, that it can quickly defeat any single retroviral agent. For this reason the only effective treatment of AIDS requires the use of three different drugs in combination which reduces the probability of a variant arising to a low enough level to prevent resistance building up.

So, natural selection produces traits that are locally adapted, i.e. that perform functions, confer net benefits to whatever bears these traits. Natural selection packages these traits together into local equilibria that endure for varying time periods – hundreds of millions of years in some cases, depending on the constancy of the environment. But these equilibria are always liable to be ended when environments change or when persistent though blind variations that add or change traits in ways better able to exploit or even destroy the local equilibrium. Since environments change slowly, many local equilibria emerge and increasingly approach local optima. Therefore they resist undermining by random variations, and the resulting biological regularities usually endure for eons. But when environments begin to change rapidly, the geographic range and the life-times of local regularities begins to shorten. By the time you get to disease-causing bacteria and host, the local equilibria are very short lived.

The application of all this to human affairs is obvious, direct and highly significant. To begin with, it solves Hayek's problem. Spontaneous order is relatively long lasting and widespread local equilibrium. It is the result of Darwinian cultural selection operating on strategies, packaging them together in ways that pay off for participants – often cooperating or competing individuals, sometimes coalitions of them. Others of these strategies get packaged together into institutions that parasitise all their participants or only some, symbiotically benefiting others, or more likely doing both at the same time, to differing degrees, to all participants, practitioners, group members.

Consider the three examples from economics developed above: money, the firm, and the price system. We could have picked others, examples of political institutions – parliamentary democracy, social practices like the Indian caste system, complex cross-cousin marriage rules that anthropologists have uncovered, historically long-lived practices such as primogenitor, or the fashionability of an innovation, like the iPhone. Each of these reflects a local equilibrium, some very long lived such as the caste system, some not quite as old, such

as parliamentary democracy, others fleeting reflections of Schumpeter's creative destruction. Our three very long lived examples – the firm, money, and the price system have been around, and will continue to be, owing to the importance of the functions they fulfil, and the low probability of environmental change or variants emerging that could unravel them. Over the eons, people have consciously and unconsciously adopted strategies that attempt to take advantage of each of them – rent seeking – by counterfeiting, or currency debasement in the case of money, by a range of business frauds in the case of the firm, by market cornering, or price controls, or insider trading in the case of the price system.

In each case the institution has responded, through new variations in the strategies that compose it, in ways that successfully resisted subversion. The local equilibrium each of them constitutes has persisted, as the institutions have found ways to adapt to changes in their environments. Institutions like money, the firm, the price system last long enough to provide an environment, a framework within which many more local, more short-lived equilibria come and go. These more local equilibria emerge as individual environmental adaptations, and co-adaptations, temporarily well matched competitors, or combinations of them. The environments within which these packages of strategies are co-adapted, are ones created by institutions and practices such as money, the firm, and the market price system. Keynes's long run equilibrium has arrived for a few long-lived fundamental human institutions. Local equilibria are nested within them, and the more local they are, the easier to break up, till at some level of strategic interaction, there are no local equilibria or none lasting long enough to exploit.

6. Sorosian Uncertainty and Reflexivity

The impermanence, instability, multiplicity and indeed the absence of equilibria in day-to-day, or even month-to-month human affairs brings us to Soros, and his insights. These insights do two things: most important, they give us the mechanism through which the Darwinian processes operate to make and break spontaneous orders. Less important, they vindicate Popper's thesis of the unity of science, the one doctrine of Sir Karl that Soros rejects.

Soros makes two two claims:

> *The human uncertainty principle*: humans are fallible, in fact usually mistaken in their expectations, including their probabilistic ones. They predict inaccurately and these predictions cannot be improved, for example by honouring the principles of probability theory more fully.
>
> *Reflexivity*: Agents' earlier expectations about future outcomes combine with their preferences in ways that change the future outcomes, often so greatly as to bear no resemblance to their earlier expectations and to fail to satisfy their preferences.

In *The Alchemy of Finance* Soros uses the combined, iterated, and cyclical operation of these two regularities to undermine the confidence economists have that we are ever near the 'long-run-all-dead' general equilibrium. He uses them to explain how several obvious facts absurdly denied by RCT-dominated economic theory in fact obtain – e.g. bubbles and busts. And Soros employs them to show that prediction is impossible – in financial markets in particular.

An example illustrates the uncertainty/reflexiveness process at its starkest, where it produces bubble and busts in the stock market. In thinking about the extreme cases it is

important to keep in mind that the process operates everywhere, and produces *routine* changes as well as *non-routine* ones.

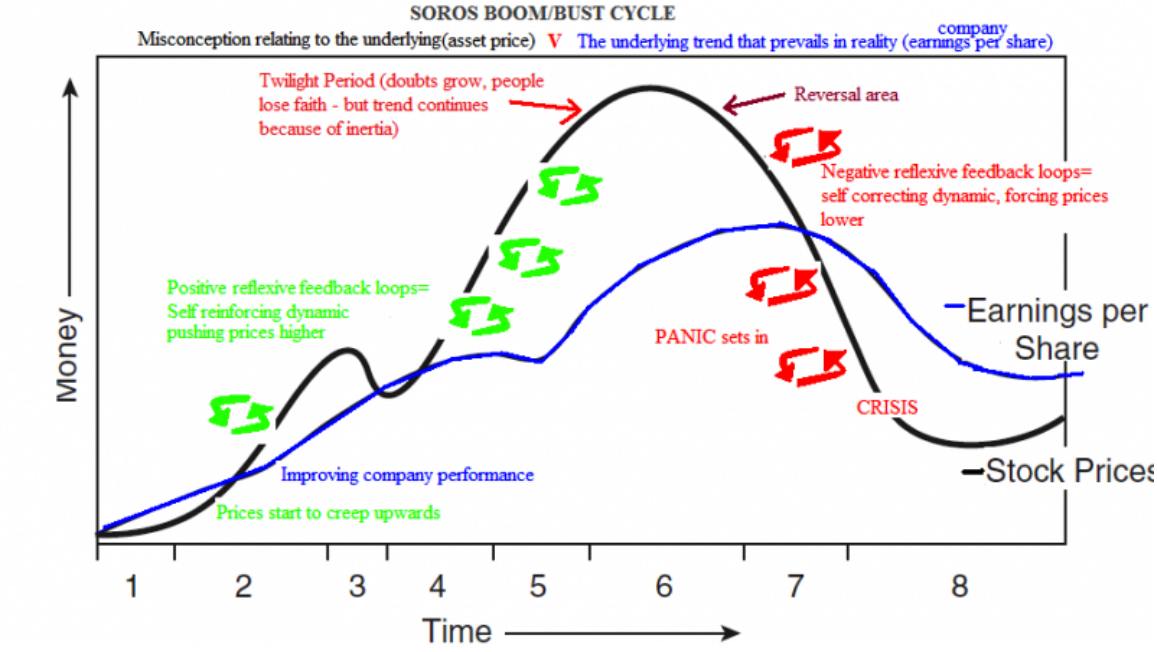


Figure 1. Stock prices track earnings and vice versa – i.e. reflexively.

The graph is based on one in *The Alchemy of Finance*, p. 56. In brief, the red curve of stock prices reflects the strategies of agents – mutual funds, hedge funds, individuals, etc. This curve reflects the aggregate of agents' uncertain, unforesighted expectations about companies' futures. The blue line imperfectly reflects the business success of companies, i.e. the equally uncertain strategies and packages of strategies of CEOs, managers, sales reps, and the shop-floor workers, and consumers, who effect the companies actual earnings per share.

Following Keynes and Knight, Soros insists that the scope for probability is extremely limited: errors do not fall on a bell-shaped curve around the truth, and new evidence does not drive it in that direction either. For that reason RCT's substitution of risk for certainty is not a significant improvement on the standard assumption of complete information. Equally important, in figure 1, the shapes of the two curves reflect the fact that earlier stock prices influence later earns per share, and vice versa. Peoples' (fallible) expectations about future states of affairs have effects on how those future states turn out, and these future states effect people's later (and always fallible) expectations. The combination of reflexivity and uncertainty, when not held in check, produce swings in two (or more) factors locked in a reflexive relation, and usually much wider swings on the expectations side of the relationship.

Expectations by themselves won't effect anything. They need to be acted upon. (Sometimes Soros calls this the 'manipulative' function of thinking, sometimes he calls it the 'participating' function.) So, reflexiveness is a relationship between strategies. But strategies are driven by expectations that lack foresight. They are almost always individually wrong, very often also wrong on average, and when the expectations are right and drive successful strategies, they are right by accident!

Now, the combination of strategy-uncertainty and strategy-reflexiveness does not simply produce wild swings in financial markets – bubbles and bubble-bursts. It operates everywhere in human affairs, because reflexivity is the rule and not the exception in these affairs. Strategies that one set of agents and organisations employ to exploit other peoples' and other organisations sets of strategies effect the second set of strategies and these in turn effect the success and thus the spread and persistence of the first set of strategies. This makes human affairs unpredictable to participants owing to the ineliminable combination of uncertainty and reflexivity that drives the choice of strategies reflexively linked.

Contrast the picture of Sorosian reflexivity/uncertainty with the mainstream economists' picture: Rat Choice economic theory would have these curves move very closely together, since all parties are hypothesised to employ probability theory and the average of their expectations cancel out to the actual objective probabilities. The two curves should move in lock-step. This is the efficient markets hypothesis, reflected in figure 2.

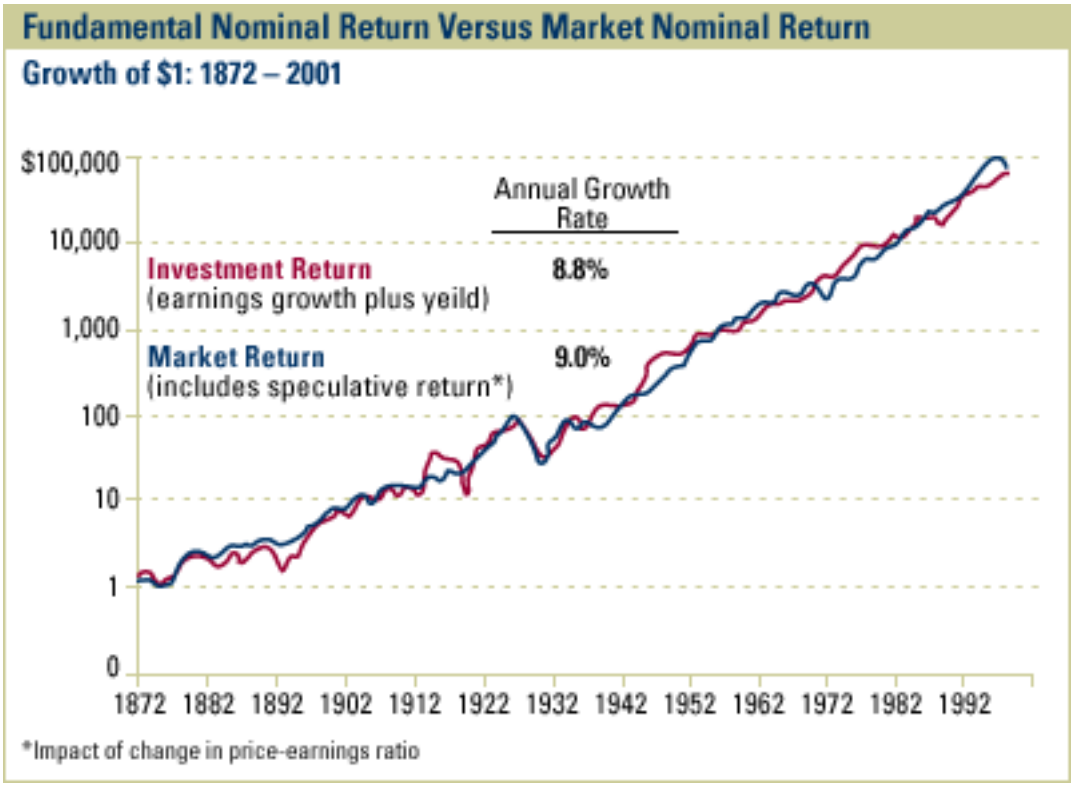


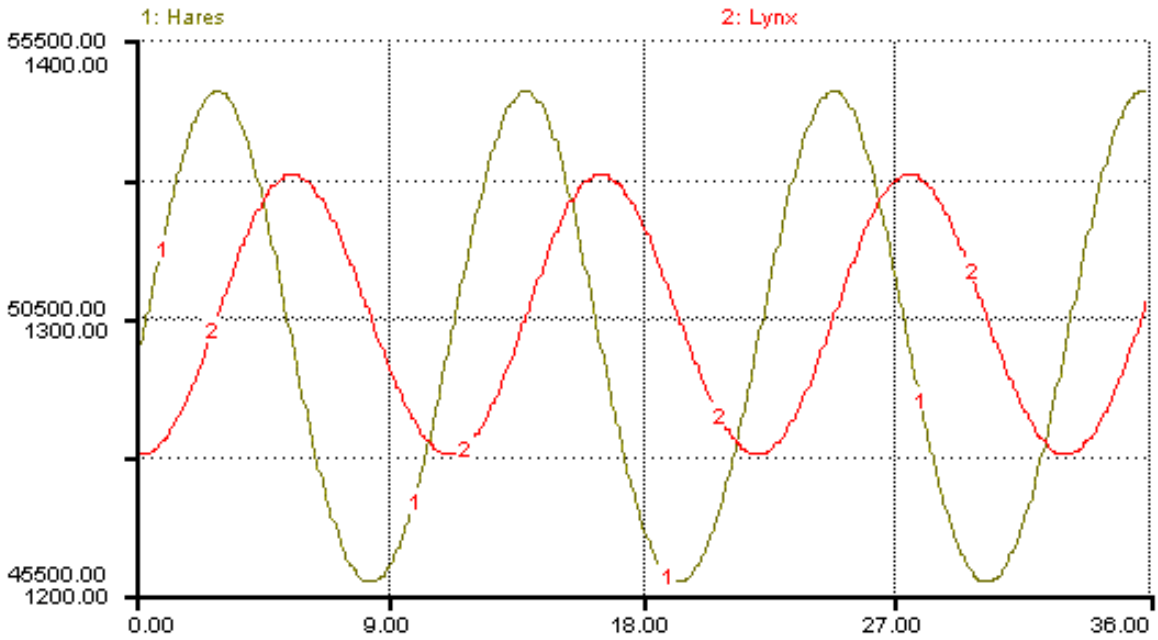
Figure 2. Stock prices track earnings over a 120 year period.

If Soros is correct the combination of reflexivity and uncertainty makes mainstream, general equilibrium-oriented, economics impossible.

How do we know Soros is correct and mainstream economics, Rat Choice and the microeconomic 'paradigm' they drive is wrong? Which graph above is correct, figure 1, the Soros boom/bust curve, or figure 2, the 120 years of efficient markets? The empirical data by themselves won't decide. There are many reasons they can't decide. But the main reason they can't is that a) it is equivocal, b) data collection is theoretically driven and doesn't point in the direction of any theory without a great deal of theoretical adjustment and interpretation. We know that Soros' insights about uncertainty in expectations and reflexivity in their effects

are right because these two processes are driven by the Darwinian processes. The same Darwinian processes prevent us from being rational fools, make us users of fast and frugal heuristics, produce spontaneous order, and shape the institutions, groups and practices through which humans navigate. Sorosian reflexivity and uncertainty is a matter of Darwinian forces acting in human culture. Let's see exactly why.

The graph below represents a typical predator-prey population cycle over time. In this case, lynx and hare populations cycle between limits with a constant six-month lag between population maxima and population minima.



The cause of this pattern is a combination of reflexivity and uncertainty. Lynx survive by employing the strategy of preying on hare. Lynx-predatory strategies select for hare strategies that are good at avoiding lynx-predation. Successful hare hunting strategies increases lynx populations, but this reduces later hare populations and so even later reflexively reduces lynx populations by reducing the pay-offs to their predation strategies.

The reflexiveness of the relationship between lynx and hare strategies is well understood in evolutionary biology: each set of strategies is subject to linked frequency dependent selection with a lag. Each is maintained within a certain minimal and maximal range by stabilising selection. Underneath this stable cycle, both lynx and hares are varying their behaviours randomly, without foresight. Mother nature faces the same kind of uncertainty we humans face.

Now, compare this curve to the Soros boom/bust cycle curve. It's the same curve of lagged reflexiveness of stock prices to earnings per share, the same relationship between strategies of stock-pickers and strategies at work in firms reflected in their earnings per share. The difference is that the predator-prey graph covers four cycles, and Soros's covers just one.

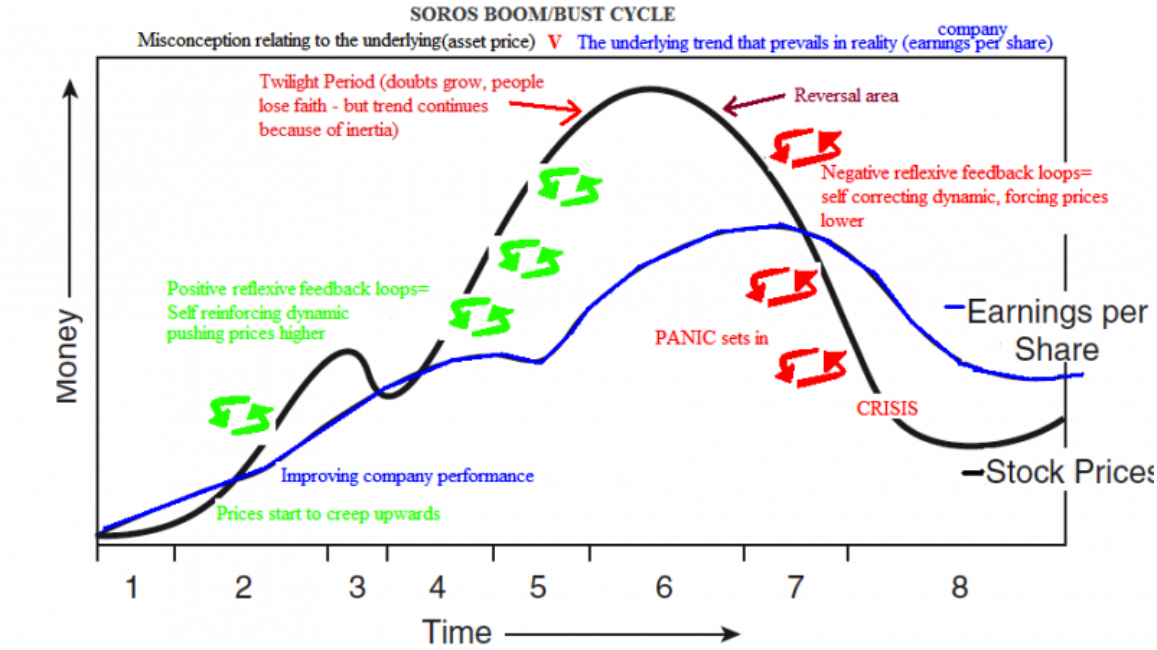


Figure 1. Again.

The point is not simply that the curves in each graph share the same relationship. The processes they describe are the same. And the Darwinian selective process that gives rise to them is the same too: reflexiveness and uncertainty. In biological contexts *reflexiveness* is the linked frequency-dependent selection of strategies that compete, cooperate, are parasitical or symbiotic. In the biological domain the *uncertainty*, the lack of foresight in strategy-choice, is reflected in their persistence coupled with the blindness of variations among them. Both operate in Darwinian cultural processes to produce the same phenomena.

The difference between the biological and the cultural is a difference of degree – the tempo and mode of evolution.

In the biological domain reflexivity and uncertainty are usually kept in bounds that produce stable cycles, ones that even vindicate some limited confidence in predictions among ecologists, agricultural scientists, even some epidemiologists. The two reasons are obvious. Reflexivity changes the environment. But it rarely changes the environment within which a strategy is played so much that it drives itself to extinction, or to complete dominance – fixation in the biologist's terminology. Extinctions are well understood. Fixations less so. About the only example of a strategy that in nature has achieved complete domination is the strategy of coding genetic information in nucleic acids. Somewhat less dominant, but almost universal, is the 1:1 ratio between strategies of bearing males vs. females within the 'environment' created by sexual reproduction (though we don't yet know what environmental 'design problem' selected for the strategy of sexual reproduction itself). The second constraint that keeps reflexive strategies cycling in balance is that the range of random variation is limited, the differences in pay-offs to the variants is small, and their rates of replication are slow, requiring a generation or so. All this means that some local equilibria in biology last long enough that it is worthwhile locating them and trying to exploit them in applied science.

In Darwinian cultural processes, none of these constrains operate, and they produce Sorosian phenomena everywhere. The rate at which strategies replicate (through imitation as well as instruction and enforcement) has been accelerating over the entire 50,000 years of human cultural evolution, and the range of variations in these strategies has also proliferated, though it is crucial to bear in mind that they are as lacking in foresight as ever. What is more,

owing to the nested character of human institutions with functions, and the speed with which they change, there has been very strong selection for strategies that track changes in other strategies and influence them as well, that is, for reflexive strategies. Finally, environments change, and once human culture and its inevitable reflexivity crosses a threshold, most selective environments never remain stable long enough to allow for the *repetition* of the same boom/bust cycles that make them of any predictive use. The efficient markets curve above looks regular and reliable until you look at the time scale on the bottom of the curve. It's 130 years! Spread out the curve and it provides no guidance any one can use in day-to-day, week-to-week, month-to-month strategy selection!

In the biological domain, uncertainty and reflexivity are held in check by environments that change with geological slowness. This produces long-lived local equilibrium outcomes. In the human domain, the environment is cultural. It is composed of nested sets of strategies that are all affected by both reflexivity and uncertainty. The result of their operation is at most short-lived local equilibria, broken up by radical environmental change. The source of this radical environmental change in human cultural processes is obvious. It is the iterated, unsynchronised combinations of reflexivity and uncertainty in strategy variation. As the rate of change in the cultural environment increases, the lifetimes of local equilibria shorten, until in many cases they disappear altogether. Human affairs appear chaotic because many of them are far from equilibrium, even if there are any equilibria to which they may temporarily be heading.

The problem reflexivity makes for all human agents is that almost all local equilibria of interest to us are too short lived to be exploited. Owing to reflexivity, many equilibria last for an hour or a day or a week or a month. By the time people have noticed, and figured out how to exploit them, they have evanesced, disappeared, been broken up by a new variant that breaks them up and substitutes a new, equally short-lived equilibrium, or perhaps an arms race, rapidly searching through the 'space' of strategy-variations for a new impermanent equilibrium. When some innovation – a new fashion, gadget, political slogan – does manage to exploit a local equilibrium, we can be confident that it arose without foresight, and that it will probably disturb the local equilibrium immediately, eventually break it up completely, and then itself fall victim to some newer strategy.

There are of course many, very long-lasting, local equilibria in human affairs. Most of them persist owing to be benefits they accord people, some exist in spite of the net costs they impose on people. They can be exploited by policy design and implementation; some of them can even be unravelled by policy, for example, consider the fate of tobacco smoking in western society over the last half-century. Most long-lived, widespread local equilibria are hard for small numbers of individuals to exploit or undermine. It's the short-lived local equilibria that are the targets of 'rent-seeking' – in business especially. And they are too short lived for any simple recipe for securing such rents to succeed for long.

Acknowledgements

Thanks to John Latsis, Constantinos Repapis, Dorian Julien, audiences at the Central European University, Purdue University, the Philosophy of Social Science Roundtable at the University of California, Santa Cruz, and most of all George Soros, for comments on previous versions of this essay.

References

Arrow, K., Debreu, G. (1954) 'Existence of an Equilibrium for a Competitive Economy', *Econometrica*, 22, No. 3, pp. 265-290.

Buchanan, J. B. (1990) *The Economics and the Ethics of Constitutional Order* (Ann Arbor: University of Michigan Press).

Campbell, D. T., et. al. (1966) *Unobtrusive Measures: nonreactive research in the social sciences* (Chicago: Rand McNally).

Durkheim, E. (1995) *The Rules of the Sociological Method*, (New York: Free Press) 1982.

Edgworth, F. Y. (2003) *Mathematical psychics, and further papers on political economy* (Oxford University Press).

Elster, J. (1978) *Logic and Society* (New York: Cambridge University Press).

Frank, R. (1988) *Passion within Reason* (New York: W.W. Norton).

Gigerenzer, G. et. al., *Ecological Rationality* (New York: Oxford University Press).

Hayek, F. (1945) 'The Uses of Knowledge in Society,' *American Economic Review*. 35, No. 4. pp. 519-30.

Hayek, F. (1978) 'The Denationalization of Money', in Hayek, 1999, *Good Money* Part II (London: Routledge).

Jevons, W. S. (1877) *The Theory of Political Economy* (London: Macmillan).

Kahneman, D. (2003) 'A Psychological Perspective on Economics,' *American Economic Review,* 93, pp. 162-168.

Kahneman, D. (2011) *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux).

Keynes, J.M. (1923) *Tract on Monetary Reform* (London: MacMillan).

Keynes, J.M. (1937) *The General Theory of Employment, Interest and Money* (London: McMillan).

Knight, F. (1921) Risk, *Uncertainty and Profit* (New York: Houghton Mifflin).

Mackenzie, R. (1984) *Constitutional Economics* (Lexington: Mass).

Parsons, T. (1951) *The Social System*, London (Routledge: London).

Popper, K. (2004) *Conjectures and refutations: the growth of scientific knowledge* (London: Routledge).

Schumpeter, J. A. (1942) *Capitalism, Socialism and Democracy* (New York: Harper).

Sen, A. (1977) 'Rational Fools: A Critique of the Behavioral Foundations of Economic Theory,' *Philosophy & Public Affairs*, 6, (4) (Summer, 1977), pp. 317-344.

Sen, A. (1997) *Development as Freedom* (New York: Anchor Books).

Sen, A. (1993) 'On the Darwinian View of Progress,' *Population and Development Review*, 19, pp. 123-137.

Simon, H. (1955) 'A Behavioral Theory of rational choice,' *Quarterly Journal of Economics*, 69, pp. 99-188.

Soros, G. (2003) *Alchemy of Finance* (New York: Wiley).

Wicksteed, P.H. (1910) *The Common Sense of Political Economy* (London: Macmillan).

Winter, S. G. (1975) 'Optimization and Evolution in the theory of the firm,' *Adaptive Economic Models* (ed. By R.G. Day and T. Groves) (New York: Academic Press), pp. 73-118.

SUGGESTED CITATION:

Rosenberg, A. (2014) 'From Rational Choice to Reflexivity: Learning from Sen, Keynes, Hayek, Soros, and most of all, from Darwin'. *Economic Thought*, 3.1, pp. 21-41.
http://www.worldeconomicsassociation.org/files/journals/economicthought/WEA-ET-3-1-Rosenberg.pdf