COMPLEXITY, THE SANTA FE APPROACH, AND NON-EQUILIBRIUM ECONOMICS

Author(s): W. Brian Arthur

Source: *History of Economic Ideas*, 2010, Vol. 18, No. 2, COMPLEXITY AND THE ORGANIZATION OF ECONOMIC LIFE (2010), pp. 149-166

Published by: Accademia Editoriale

Stable URL: https://www.jstor.org/stable/23723515

COMPLEXITY, THE SANTA FE APPROACH, AND NON-EQUILIBRIUM ECONOMICS

W. Brian Arthur*

Santa Fe Institute for the Study of Complex Systems and PARC

Over the last twenty or more years a different way of doing economics has been slowly emerging. It goes by several names: complexity economics, agent-based computational modeling, generative economics, Santa Fe economics, and each of these has its own style, its own followers, and its own nuances. What I want to do in this paper is to recount my own involvement with what has happened, and the part that the Santa Fe Institute has played. And I also want to ask what this new movement in economics is: what complexity economics really is.

After two centuries of studying equilibria – patterns of consistency that call for no further behavioral adjustments – economists are beginning to study the unfolding of patterns in the economy. That is, we are starting to study the economy out of equilibrium. This way of doing economics calls for an algorithmic approach. When viewed out of equilibrium, the economy reveals itself not as deterministic, predictable and mechanistic; but as process-dependent, organic and evolving.

O VER the last twenty or more years a different way of doing economics has been slowly emerging. It goes by several names: complexity economics, agent-based computational modeling, generative economics, Santa Fe economics, and each of these has its own style, its own followers, and its own nuances. What I want to do in this paper is to recount my own involvement with what has happened, and the part that the Santa Fe Institute has played. And I also want to ask what this new movement in economics is: what complexity economics really is.

Let me start with some personal history. Throughout the 1980s, I had been involved with the problem of increasing returns in economics. My work had come to Kenneth Arrow's attention, and in April 1987 he stopped me in the quad at Stanford and invited me to a meeting at a small institute in the Rockies that September. I arrived in August, along with John Holland. We were the nascent Santa Fe Institute's first visiting fellows. The institute was still in its early days, and was operating in an old convent.

The meeting had been convened at the behest of John Reed of Citibank, to discuss the idea of the economy as a evolving, complex system. Kenneth Arrow, one of the organizers had invited ten theoretical economists, among whom were Tom Sargent, Larry Summers, William (Buz) Brock, and Jose Sheinkman. Philip W. Anderson, Arrow's counterpart in physics, had invited ten people from physics, biology and computer science. Among these were John Holland, Doyne Farmer, David Pines, Richard Palmer and Stuart Kauffman. New ideas had been bubbling in the natural sciences, loosely tied under the rubric of 'the sciences of complexity', and the hope was that these might stimulate new ways of thinking about economic problems.

For ten days, the economists and natural scientists took turns talking about their respective worlds and methodologies. While physicists grappled with general equilibrium analysis and non-cooperative game theory, economists tried to make sense of spin-glass models, Boolean networks, and genetic algorithms.

The meeting generated enough excitement on both sides that the Santa Institute decided to found a research program – its first research program – and this would be on the Economy as an Evolving, Complex System. It would start a year later, in the summer of 1988, and I was asked to head it up. I spent the intervening academic year recruiting, with a great deal of help from both Arrow and Anderson. Among the first people I managed to lure to the Program were the statistician David Lane, physicist Richard Palmer, computer scientist John Holland, and theoretical biologist Stuart Kauffman (who was already resident at the Institute).

When we convened in the fall of 1988, we had no clear plans. At best all we had were instincts about what we wanted to see happen in economics. The mission was to encourage the understanding of economic phenomena from a complexity perspective, but it was far from clear what this meant. I asked Arrow and Anderson what they wanted, and was told that we should look at the foundations of economics and see what needed to be done. I asked John Reed what he wanted – Citibank was funding the Program – and was told to «do anything, as long as it is not conventional».

Over several weeks, and several months, some desiderata began to emerge. Many assumptions in conventional economics had been chosen over time not for their reality but to ensure an equilibrium, and an analytical solution. We would feel free to replace these standard assumptions in economics – perfect rationality, identical representative agents, convexity, equilibrium – with more realistic ones where required. We would pick standard problems in economics (the asset pricing problem say, the double-auction problem) and see how more realistic assumptions would change the conventional solutions. In

doing so we would go for analytical solutions, but would allow computational ones if analytical ones could not be worked out. And we would avoid working with chaos theory, then fashionable; it was being amply researched elsewhere.

Our mission, as we conceived it, was not directly to change economics – we weren't central enough for that. Rather it was to catalyze a series of changes that seemed to us inevitable in the field. In the 1980s, the neoclassical program was at a high water mark, and its reign, we believed, would not persist forever. We were aware that other groups in other institutions would be thinking along lines similar to ours too. But we had several luxuries. The Santa Fe Institute had no students, and therefore no teaching. So we had the luxury of time to think. It had no departments, and no set of colleagues with locked-in ways of thinking. Hence we had no colleagues objecting to our lack of convention, or censoring ideas. At one stage, the theoretical biologist Stuart Kauffman asked, «Why do you guys do everything at equilibrium?». We came up with a ready enough answer, but Kauffman's question nevertheless lingered.

In the first year or two of the Economics Program we were not really aware of themes. Rather we thought of ourselves as exploring problems. Still, looking back, I can see three themes that emerged throughout our work. The first of these was that of patterns slowly forming – market solutions slowly forming, in our case. Pattern formation is very much at the heart of complexity studies, so this was not surprising. But seeing solutions as patterns in formation did force an immediate break with standard equilibrium thinking. The patterns we were looking at did not need to reach any stasis, nor did they need to be permanent. Pattern formation, as John Holland instructed us, could be a matter of ongoing, perpetual novelty.

A second theme that emerged was that of making models based on more realistic cognitive behavior. Neoclassical economic theory treats economic agents as perfectly rational optimizers. This means among other things that agents perfectly understand the choices they have, and perfectly assess the benefits they will receive from these. If there is uncertainty, they evaluate it probabilistically, revise their evaluations in the light of new information, and choose the course of action that maximizes their expected utility. Where there are multiple parties involved, each agent is usually assumed to have common knowledge about the others' possible choices and assessments of these. Our approach, by contrast, saw agents not as having perfect information about the problems they faced, or as generally knowing enough about other agents' options and payoffs to form probability distributions over these. This meant that agents need to cognitively structure their problems – as having to 'make sense' of their problems, as much as solve them. And they

have to do this with cognitive resources that are limited. It follows that agents generally do not optimize in the standard sense, not because they are constrained by finite memory or processing capability, but because the very concept of an optimal course of action often cannot be defined. We were greatly guided here by the thinking of John Holland and his colleagues on induction. We could use Holland's ideas to model 'making sense' in economics, to model inductive behavior.

Usually, we found, such formation of cognition – such 'making sense' of an economic problem or situation – was too complicated to be handled analytically. And this brought in the third theme. There are many ways to make sense of a situation and agents might naturally differ in these. So we found we often modeled agents and their sense-making as individually differing, and resorted to the computer to allow these differing 'cognitions' to interact, update themselves, and produce ongoing aggregate outcomes. At the time such agent-based modeling did not yet have a label, and versions of element-based simulation were being developed in physics, biology, and other fields. So the idea was in the air. But by allowing the elements of simulation to be human economic agents, for the first time we could not just study equilibria but ask how they formed.

I will not talk further about this early history. In 1991 Mitchell Waldrop conducted a series of interviews on these early days, and his book, *Complexity*, details our efforts. And the paper in this volume by Magda Fontana provides a further account. The Program we started in 1988 lasted to the mid-1990s, and was headed by a succession of economists. One of its most effective actions, I believe, was to bring in pre-doctoral students, both as researchers and in an annual summer school organized since 1994 by John Miller and Scott Page. The students could absorb the approach, and infect others at their home institutions.

One natural question in retrospect is how effective was the Santa Fe Institute's Program? I will leave it to others to assess this. And certainly Santa Fe was not the sole pioneer of this perspective. Parts of it were popping up in other places. What I will say is that research related to the complexity perspective is under active development now in a number of different institutes and university departments. Today many of approaches developed at SFI – non-convex economics, agent-based modeling, computational economics – have either entered the mainstream or at least are no longer controversial.

But there remains a larger question, and I want to devote the bulk of this essay to it. Is complexity economics not just a fashion, a minor adjunct to neoclassical theory? Is it one of these temporary fads that fields go through – catastrophe theory, chaos theory, complexity theory? Is this way of looking at the economy, with its heavy reliance on different

assumptions and heavy use of computation undisciplined – *ad hoc* – a retreat from theory? What does this way of doing economics really provide?

In what follows I want to argue that this movement is not a minor adjunct to neoclassical economics; it is something more than this. It is a shift from looking at economic problems at equilibrium to looking at such problems out of equilibrium, a shift to a more general economics – to a non-equilibrium economics.

Let me step back from our Santa Fe experiences and at the risk of some repetition argue this from first principles.

1. Beyond Equilibrium

Economic agents – banks, consumers, firms, investors – continually adjust their market moves, buying decisions, prices, and forecasts to the situation these moves or decisions or prices or forecasts together create. To put this another way, individual behaviors collectively create an aggregate outcome; and they react to this outcome. There is nothing new in saying this. Economists have seen the economy this way at least since Adam Smith. Behavior creates pattern; and pattern in turn influences behavior.

It might be natural in such a setting for economic theorists to study the unfolding of patterns that economic agents create. But this obviously is complicated. And therefore to seek analytical solutions, historically economics chose to simplify its questions. It asked instead what behavior caused an outcome or pattern that leads to no incentive to change that behavior. In other words, it asked what patterns in the economy would look like if they were at equilibrium – were consistent with the micro-behavior (actions, strategies, expectations) that creates them. Thus, for example, general equilibrium theory asks: What prices and quantities of goods produced and consumed are *consistent with* – would pose no incentives for change to – the overall pattern of prices and quantities in the economy's markets? Game theory asks: What strategies, moves, or allocations are consistent with – would be the best course of action for an agent (under some criterion) – given the strategies, moves, allocations his rivals might choose? Rational expectations economics asks: What forecasts (or expectations) are consistent with – are on average validated by – the outcomes these forecasts and expectations together create? Partial-equilibrium economics – say in international trade theory – asks: what local behaviors would produce larger patterns that would support (be consistent with) those local behaviors.

This equilibrium approach lends itself to expression in equation form. And because an equilibrium by definition is a pattern that does-

n't change, in equation form it can studied for its structure, its implications, and the conditions under which it obtains. Of course the simplicity that makes such analytical examination possible has a price. To ensure tractability we usually have to assume homogeneous (or identical) agents, or at most two or three classes of agents. We have to assume that human behavior – a notoriously complicated affair – can be captured by simple mathematical functions. We have to assume agent behavior that is intelligent but has no incentive to change; hence we must assume that agents and their peers deduce their way into exhausting all information they might find useful, so they have no incentive to change. Still, as a strategy of advancement of analysis, this equilibrium approach has been enormously successful. As it evolved into the neoclassical structure we know today, it has built a degree of understanding that is the envy of other social sciences.

It is natural to go beyond this equilibrium approach and ask how agents' behavior might not just be *consistent with* the aggregate pattern it creates, but how actions, strategies, or expectations might in general react to – might *endogenously change with* – the patterns they create. In other words, it is natural to ask how the economy behaves when it is not at a steady state – when it is *out of equilibrium*. At this more general level, we can surmise that economic patterns might settle down over sufficient time to a simple, homogeneous equilibrium. Or, that they might not: they might show ever-changing, perpetually novel behavior. We might also surmise they might show new phenomena that do not appear in steady state.

By its very nature this new approach calls for detailed instructions on how individual behavior adjusts as the situation unfolds; in this sense it is algorithmic. And since there is considerable scope for learning or reacting in different ways, this approach sees no reason to treat adjustments in behavior as identical. Agents must therefore be separately considered; hence the approach is based on individual agents. Consideration of economic patterns out of equilibrium therefore naturally introduces algorithmic updating and heterogeneity of agents. On both these counts it is best handled by computation.

One possible objection to doing economics this way is that because the approach is computational, it does not constitute theory. But this statement is too facile. If working out the implications of a set of assumptions is theory, then whether this is done by hand or by computer does not matter. Both methods yield theory. But certainly there is a difference in style. Equation-based methods call for equation-based dissection of the results – and equation-based discovery of telling implications – and this dissection and analysis can be accomplished rigorously. Of course often the rigor is specious. Implications match reality only as

well as the chosen assumptions and chosen functional forms do; and functional forms are always abstractions of reality – often gross ones when closely examined – so there is plenty of scope for rigorous deduction based upon faulty assumptions. Computer-based modeling is different but parallel in these regards. It calls for statistical dissection of the phenomena discovered, and in many computer-based models it may be difficult to discern phenomena through the thicket of events. There is also scope for unrealistic assumptions and for needless complication. And doing computer-based economics well is not necessarily easier than doing analytical economics well. Good work here shows an eye for elegance of experiment for the telling, simple, computational model that demonstrates a phenomenon clearly; and for extracting a phenomenon from the jumble of other data that obscure it.[1]

A different objection is that because non-equilibrium studies require detailed modeling of how individual behavior adjusts (and how agents interact), they encourage behavioral assumptions that are *ad hoc*. The point has some merit: assumptions are sometimes adopted for convenience. But we need to remember that the standard assumptions of 'rational behavior' themselves are highly stylized versions of reality. If modeling agent adjustments forces us to study and think rigorously about *actual* human behavior, this is actually a strength.

Non-equilibrium studies of course do not answer all possible questions. They do not tell us usually about the formation of tastes, or of technologies, or of structure. David Lane (1993) notes that such studies «offer only very limited scope to the emergence of new structures – and, so far, none at all to the emergence of higher-level entities». What emerges is pattern, not hierarchical structure.

One thing noticeable about complexity-based studies is that they are nearly always evolutionary in approach. Why should this be? I said earlier that an assumption common to most studies is that agents differ in the way they react to aggregate patterns; they have different circumstances, different histories, different psychologies. That is, agents are adaptive and heterogeneous. On first thought, this might seem to yield at most a trivial extension to standard homogeneous theory. But consider. If heterogeneous agents (or heterogeneous strategies or expectations) adjust continually to the overall situation they together create, then they adapt within an 'ecology' they together create. And in so adapting, they change that ecology. So providing we use 'evolution' in

---

[1] The two styles can of course be mixed. If a phenomenon shows up computationally, often it can be reproduced in a simpler analytical model. If it shows up analytically, it can be probed computationally. Properly carried out, computation does not replace theory. It allows more realistic assumptions and accommodates non-equilibrium behavior. It thereby extends theory.

the broadest sense of the word, which I interpret as elements adapting
their state to the situation they together create, we see that in this sense
evolution emerges naturally from the very construction of such mod-
eling. It need not be added as an adjunct (of course in any particular case
we would need to define precisely what we mean by 'elements', 'adapt-
ing', 'states', and 'situation'). Because non-equilibrium economics is by
its nature evolutionary, it resembles modern evolutionary biology more
than it does 19th century physics.

Agent-based, non-steady-state economics is also a generalization of
equilibrium economics. Non-equilibrium systems may converge to or
display patterns that are consistent – that call for no further adjustments.
If so, standard equilibrium behavior becomes a special case. It follows
that non-equilibrium economics is not in competition with equilibrium
theory. It is merely economics done in a more general, generative way.

I have made a large claim so far, namely that a new form of econom-
ics is a-birthing – a generative or non-equilibrium economics. If the
reader accepts this, a natural question to ask is what it delivers. What
novel phenomena do we see when we do economics out of equilib-
rium? Are there questions that equilibrium economics can not answer,
but that this more general form of economics can? In Kuhnian lan-
guage, are there anomalies that this new paradigm resolves?

The answer to this last question is yes. In the remainder of this essay
I want to look at two characteristic anomalies – two indeterminacies, to
be precise – in equilibrium economics and show that these disappear
under the new approach. Along the way, I want to point to some char-
acteristic phenomena that arise in the new approach. I will base the dis-
cussion mainly on a study by Lindgren and on three topics I have been
heavily involved with, because these address directly the points I want
to make (and because I am most familiar with them). There are cer-
tainly other studies that widen the scope of agent-based economics be-
yond the discussion here.[1] These also, I believe, corroborate the argu-
ments I will make here.

2. PERPETUAL NOVELTY

Let me begin with a phenomenon, one often we see in this sort of eco-
nomics. That is the absence of any equilibrium, or more positively, the

---

[1] For some early studies see BAK *et alii* 1993, DURLAUF 1993, LINDGREN 1992, MARIMON *et alii* 1990, SARGENT 1993 and SCHELLING 1978. See also YOUNG 1998. The earliest agent-based studies I know of were by MILLER 1988 and MARKS 1989. From the most recently available col-
lection (ARTHUR *et alii* 1997), the reader might consult the papers of Blume, Durlauf, Kirman, Kollman *et alii*, Ioannides, Lane and Maxfield, and Tesfatsion. The collection of Blume and Durlauf (eds) 2006 contains more recent work. For the literature on network interactions, see WILHITE 2005.

presence of ever-changing, perpetually novel behavior. For an example, consider the classic study of Kristian Lindgren (1991). Lindgren sets up a computerized tournament where strategies compete in randomly chosen pairs to play a repeated prisoner's dilemma game. The elements in his study are therefore strategies rather than human agents. Strategies that do well replicate and mutate. Ones that lose eventually die. Strategies can 'deepen' by using deeper memory of their past moves and their opponent's. A strategy's success of course depends on the current population of strategies, and so the adaptive elements here – strategies – react to, or change with, the competitive world they together create.

In his computerized tournament Lindgren discovered that the simple strategies in use at the start went unchallenged for some time. Tit-for-tat and other simple strategies dominated at the beginning. But then other, deeper strategies emerged that were able to exploit the mixture of these simple ones. In time, yet deeper strategies emerged to take advantage of those, and so on. If strategies got 'too smart' – that is, too complicated – sometimes simple ones could exploit these. In this computer world of strategies, Lindgren found periods with very large numbers of diverse strategies in the population, and periods with few strategies. And he found periods dominated by simple strategies, and periods dominated by deep strategies. But nothing ever settled down. In Lindgren's world the set of strategies in use evolved and kept evolving in a world of perpetual novelty. This is unfamiliar to us in standard economics. Yet there is a realism about such dynamics with its unpredictable, emergent, and complicated sets of strategies. Chess play at the grand master level evolves over decades and never settles down. Lindgren's system is simple, yet it leads to a dynamic of endless unfolding and evolution.

When, in general, do we see perpetually novel behavior in the economy? There is no precise rule, but broadly speaking perpetual novelty arises in two circumstances. One is where there is frustration (to use a physics term) in the system. Roughly this means that it is not possible to satisfy the needs of all the agents (or elements) at the same time and that these jostle continually to have their needs fulfilled. The other is where exploration is allowed and learning can deepen indefinitely – can see better and better into the system it is trying to understand. In this case collective behaviors can explore into constantly new realms, sometimes mutually complicate, sometimes simplify, but not settle down.

3. EQUILIBRIUM INDETERMINACY AND THE SELECTION PROCESS

In the Lindgren case, the situation shows no equilibrium; it is always in perpetual novelty. In other cases equilibrium is possible, but there may

be more than one natural pattern of consistency: there may be multi-
ple equilibria. This situation arises naturally in the presence of positive
feedbacks or increasing returns – or more technically, under non-con-
vexity. Here multiple equilibria are the norm. At first sight this does not
seem to pose any major difficulty to equilibrium economics. Instead of
a unique equilibrium there are several. But there *is* a difficulty. Equilib-
rium economics can identify consistent patterns, but can not tell us how
one comes to be chosen. Standard economics therefore runs up against
an indeterminacy.

    This indeterminacy has been an embarrassment to economics over
the years. «Multiple equilibria», wrote Schumpeter in his 1954 book, «are
not necessarily useless, but from the standpoint of any exact science the
existence of a uniquely determined equilibrium is, of course, of the ut-
most importance, even if proof has to be purchased at the price of very
restrictive assumptions; without any possibility of proving the existence
of uniquely determined equilibria – or at all events, of a small number
of possible equilibria – at however high a level of abstraction, a field of
phenomena is really a chaos that is not under analytical control». Faced
with this potential 'chaos', different subfields of economics took differ-
ent approaches. Some – especially within game theory in the 1960s and
'70s – added restrictive (and somewhat artificial) assumptions until on-
ly a single solution remained. Others, contrary to Schumpeter, accept-
ed the chaos. They statically determined the possible equilibria in a
problem and left the choice of equilibrium open and therefore indeter-
minate. An example is the international trade theory of Helpman and
Krugman (1985) which allowed increasing returns and settled for multi-
ple static, but indeterminate, equilibria.

    A more natural approach, I believe, is to tackle the issue generatively
(Arthur 1989, 1994b): to see the problem not as one of equilibrium *se-
lection* but as one of equilibrium *formation*. Economic activity is quan-
tized by events that are too small to foresee, and these small 'random'
events – who sits next to whom on an airplane, who tenders an offer
when, who adopts what product when – can over time cumulate and
become magnified by positive feedbacks to determine which solution
was reached. This suggests that situations with multiple equilibria can
best be modeled by looking at what happens over time – what happens
in formation. That is, they are best modeled not as static deterministic
problems, but as dynamic processes with random events, with natural
positive feedbacks or nonlinearities. With this strategy the situation can
then be 'observed' theoretically as its corresponding process unfolds
again and again to 'select' or determine an outcome. Sometimes one
equilibrium will emerge, sometimes (under identical conditions) an-
other. It is impossible to know in advance which of the candidate out-

comes will emerge in any given unfolding of the process, but it is possible to study the probability that a particular solution emerges under a certain set of initial conditions. In this way the selection problem can be handled by modeling the situation in formation, by translating it into a dynamic process with random events. With an non-equilibrium approach, the anomaly disappears.

In this sense a whole realm of economics – increasing returns problems – requires a non-equilibrium approach. This realm, by the way, is not small. Increasing returns arise in economic geography, finance, economics of markets, economic development, economics of technology, and economics of poverty; and the literature in these areas is becoming large. Interestingly, in most of the important cases the work has been analytical, not computational. The reason is that most increasing returns problems lend themselves to sufficient homogeneity of agents to be handled by analysis.

Whatever their topic of focus, increasing returns studies tend to show common properties: a multiplicity of potential 'solutions'; the outcome actually reached is not predictable in advance; it is 'selected' by small events; it tends to be locked in; it is not necessarily the most efficient; it is subject to the historical path taken; and while the problem may be symmetrical, the outcome is usually asymmetrical. These properties have counterparts in a different science that emphasizes the formation of pattern: solid-state physics. What economists call multiple equilibria, non-predictability, lock-in, inefficiency, historical path dependence, and asymmetry, physicists call multiple meta-stable states, unpredictability, phase- or mode-locking, high-energy ground states, non-ergodicity, and symmetry breaking. Some of these properties can be identified by static analysis (multiplicity, possible non-efficiency, non-predictability, and lock-in). But to see how they come about, and to see symmetry breaking, selection, and path-dependence in action, requires looking at the situation as the solution forms – out of equilibrium.

4. EXPECTATIONAL INDETERMINACY AND INDUCTIVE BEHAVIOR

Multiple equilibria cause one type of indeterminacy in static economics. Expectations can cause another, and this also requires nonequilibrium resolution. Let me explain.

All economic actions are taken on the expectation of some outcome. And in many situations this outcome is determined collectively – it depends upon the results of other people's actions. Thus an entrepreneur may have to decide on whether to invest in a new semiconductor fabrication plant today, based upon what he forecasts supply in the market to be like in two years' time. And his competitors may have to make

similar decisions. But the collective result of their choices today will determine the aggregate supply (and hence prices and profits) in two years' time.

In cases like this, agents attempt to forecast what the outcome will be; but their actions based on their forecasts determine this outcome. So the situation is self-referential: agents are trying to form expectations about an outcome that is a function of their expectations. Or, to collapse this further, their choices of expectation depend on their choices of expectation. Without some additional conditions imposed, there is no logical or deductive way to settle this self-referential choice. This is a fundamental indeterminacy in static economics.

It is tempting to dismiss this as a minor anomaly, but the situation that causes it pervades economics: it occurs anywhere agents' decisions affect other agents.[1] It confronts economics with a lacuna – how expectations might logically be formed in multi-agent situations. And it is the main reason economists feel uneasy about problems with expectations.

Static economic theory, of course, *does* deal with problems where multi-agent expectations must be considered; it has evolved a theoretical method – a sort of analytical workaround – to do this: the rational expectations approach. Rational expectations asks, within a given economic problem, what expectational model (if everyone adopted it) would lead to actions that would on average validate that expectational model. If such a model existed, agents' expectations would be on average upheld, and this would solve the problem of selecting suitable expectations.

Actually, this last assertion came too fast. To be rigorously exact, if such a model existed it would demonstrate at least one set of expectations consistent with the outcome. Whether this translates into a *theory* of expectations formation matched by reality is another question, one that leaves even supporters of this approach uncomfortable. To suppose that this solution to a given problem would be reached in a one-off non-repeated problem, we would need to assume that agents can somehow deduce in advance what model will work, that everyone 'knows' this model will be used, and everyone knows that everyone knows this model will be used, *ad infinitum* (this is the common knowledge assumption). And we would further require a unique solution; otherwise agents might coordinate on different expectations.

The net effect is that unless there is good reason for agents to coordinate somehow on a single set of expectations, rational expectations become theoretically singular: they resemble a pencil balanced on its point

---

[1] For some history and commentary on this indeterminacy see KOPPL and ROSSER 2002.

– logically possible but in reality unlikely. The situation worsens when agents differ. They must now form expectations of an outcome that is a function of expectations they are not privy to. Whether behaviorally or theoretically, barring some obvious coordinating set of expectations, the indeterminacy can not be avoided. Deductive equilibrium economics therefore faces an anomaly.

As a theory of expectations formation, rational expectations begin to look better if the situation is repeated over time, because we might suppose that agents 'learn' their way over time into on-average correct expectations. In this case rational expectations would at least form a solution to which expectations converge. But it is possible to construct repeated situations in economics where rational expectations are not a guide – where in fact they *must* fail. Consider the El Farol bar problem (Arthur 1994b). One hundred people must decide independently each week whether to show up at their favorite bar (El Farol in Santa Fe, say). The rule is that if a person predicts that more that 60 (say) will attend, she will avoid the crowds and stay home; if she predicts fewer than 60 she will go. We see at once the self reference I mentioned above: agents attend based on their predictions of how many agents will attend.

Will rational expectations work here? Suppose for a moment they do. Suppose that a rational expectations prediction machine exists and all agents possess a copy of it. Such a machine would take a given history of attendance (say, ten weeks back) and map it into a forecast of the coming week's attendance, and by definition it would on average predict correctly. Suppose now this machine predicts one week that 74 will attend. But, knowing this nobody shows up, negating that forecast. Suppose the next week it predicts 44. Then 100 people go, negating that forecast as well. In El Farol, expectations that are shared in common negate themselves. Therefore forecasts that are on average consistent with the outcome they predict do not exist and can not be statically deduced. As a theory of expectations formation, rational expectations fails here. The indeterminacy is also manifest in this case. Any attempt to deduce a reasonable theory of expectations that applies to all is quickly confounded.[1]

The anomaly resolves itself in this case (and in general) if we take a generative approach and observe expectations in formation. To do this we can assume agents start each with a variety of expectational models

---

[1] This El Farol situation of preferring to be in the minority occurs in the economy anywhere pre-committed decisions have to be made under diminishing returns (to the numbers committing). In its minority game formulation, the problem is much studied among physicists (see CHALLET, MARSILI and ZHANG 2005, and COOLEN 2005).

or forecasting hypotheses, none of them necessarily 'correct'. We can assume these expectations are subjectively arrived at and therefore differ. We can also assume agents act as statisticians: they test their forecasting models, retain the ones that work, and discard the others. This is inductive behavior. It assumes no *a-priori* 'solution' but sets out merely to learn what works. Such an approach applies out of equilibrium (expectations need not be consistent with their outcome) as well as in equilibrium; and it applies generally to multi-agent problems where expectations are involved.[1] Putting this into practice in the case of El Farol means assuming that agents individually form a number of predictive hypotheses or models, and each week act on their currently most accurate one (call this their active predictor). In this way beliefs or hypotheses compete for use in an ecology these beliefs create. Computer simulation then shows that the mean attendance quickly converges to 60. In fact, the predictors self-organize into an equilibrium pattern or 'ecology' in which, on average, 40% of the active predictors are forecasting above 60 and 60% below 60. And while the population of active predictors splits into this 60/40 average ratio, it keeps changing in membership forever. There is a strong equilibrium here, but it emerges ecologically and is not the outcome of deductive reasoning.

My point in this discussion is not just that it is possible to construct problems that confound rational expectations. It is this: In multi-agent situations the formation of expectations introduces a fundamental indeterminacy into equilibrium economics; but if we allow expectations to form out of equilibrium in an inductive, agent-based way, the indeterminacy disappears. Expectation formation then becomes a natural process.

If we apply this generative approach to standard problems, do expectations indeed usually converge to the rational expectations norm? The answer is mixed: sometimes they do and sometimes they don't, depending on whether there is a strong attractor to the rational expectations norm or not. Interestingly both answers can obtain in the same problem. Different parameter sets can show different behaviors. In one set (or phase or regime) simple equilibrium behavior might reign; in another complex, non-converging pattern-forming behavior might obtain. My guess is that such phases will turn out to be common in agent-based models.

Consider as an example the Santa Fe artificial stock market (Palmer *et alii* 1994, Arthur, Durlanf and Lane eds 1997). The model is essentially a heterogeneous-agent version of the classic Lucas equilibrium mod-

---

[1] See HOLLAND *et alii* 1986 and SARGENT 1994.

el (1978). In it heterogeneous agents, or artificial investors, form a market within the computer where a single stock is traded. Each monitors the stock price and submits bids and offers which jointly determine tomorrow's price. Agents form (differing) multiple hypotheses of what moves the market price, act on the most accurate, and learn by creating new hypotheses and discarding poorly performing one. We found two regimes: if agents update their hypotheses at a slow rate, the diversity of expectations collapses into a homogeneous rational expectations regime. The reason is simple: if a majority of investors believes something close to the rational expectations forecast it becomes a strong attractor; others lose by deviating from these expectations and slowly learn their way to them. But if the rate of updating of hypotheses is tuned higher, the market undergoes a phase transition into a 'complex regime'. Here it displays several properties seen in real markets. It develops a rich 'psychology' of divergent beliefs that do not converge over time. Expectational rules such as 'If the market is trending up, predict a 2% price rise' appear randomly in the population of hypotheses and become temporarily mutually reinforcing. (If enough investors act on these, the price will indeed go up). In this way sub-populations of mutually reinforcing expectations arise, and fall away again. This is not quite perpetual novelty. But it is a phenomenon common to such studies: patterns that are self-reinforcing arise, lock-in for some time (much as clouds do in meteorology), and disappear.

We also see another phenomenon, again common to non-equilibrium studies: avalanches of change of varying sizes. These arise because individual non-equilibrium behavior adjusts from time to time, which changes the aggregate, which in turn may call for further behavioral changes among agents. As a result in such systems cascades of change – some small and some large – can ripple through the system. In artificial markets this phenomenon shows up as agents changing their expectations (perhaps by exploring new ones) which changes the market slightly, and which may cause other agents to also change their expectations. Changes in beliefs then ripple through the market in avalanches of all sizes, causing random periods of high and low price volatility. This phenomenon shows up in actual financial market data but not in equilibrium models. One interesting question is whether such avalanches show properties associated with phase boundaries in physics, namely power laws where the size of the avalanche is inversely proportional to its frequency. Systems that display this behavior may be technically *critical*: they may lie precisely between ordered and chaotic behavior. We might conjecture that in certain economic situations behavior ensures that the outcome remains poised in this region – technically that self-organized criticality (Bak *et alii* 1988) arises.

W. Brian Arthur

5. Conclusion

After two centuries of studying equilibria – patterns of consistency that call for no further behavioral adjustments – economists are beginning to study the emergence of equilibria and the general unfolding of patterns in the economy. That is, we are starting to study the economy out of equilibrium. This way of doing economics calls for an algorithmic approach. And it invites a deeper approach to agents' reactions to change, and a recognition that these may differ – and therefore that agents are naturally heterogeneous. This form of economics is naturally evolutionary. It is not in competition with equilibrium theory, nor is it a minor adjunct to the standard economic theory. It is economics done in a more general, non-equilibrium way. Within this, standard equilibrium behavior becomes a special case.

When viewed out of equilibrium, the economy reveals itself not as deterministic, predictable and mechanistic; but as process-dependent, organic and evolving. Economic patterns sometimes simplify into a simple, homogeneous equilibrium of standard economics. But often they do not. Often they are ever-changing, showing perpetually novel behavior.

One test of a different fundamental approach is whether it can handle certain difficulties – anomalies – that have stymied the old one. Certainly this is the case with non-equilibrium economics. Within the static approach, both the problem of equilibrium selection and of choice of expectations are in general indeterminate. These two indeterminacies should not be surprising, because both problems are in essence ones of formation – of coming into being – that can not be resolved by static analysis. Both have been the source of considerable discomfort in economics. But when analyzed out of equilibrium they fall into their proper setting, and the difficulties they cause dissolve and disappear.

References

Anderson P., Arrow K. and Pines D. (eds) 1988, *The Economy as an Evolving Complex System*, Reading (MA), Addison-Wesley.

ARTHUR W. B. 1989, «Competing Technologies, Increasing Returns, and Lock-In by Historical Events», *Economic Journal*, 99, 116-131.

— 1994a, *Increasing Returns and Path Dependence in the Economy*, Ann Arbor (MI), University of Michigan Press.

— 1994b, «Bounded Rationality and Inductive Behavior (the El Farol problem)», *American Economic Review Papers and Proceedings*, 84, 406-411.

— 1999, «Complexity and the Economy», *Science*, 284, 107-109.

ARTHUR W. B., HOLLAND J. H., LEBARON B., PALMER R. and TAYLER P. 1997, «Asset Pricing Under Endogenous Expectations in an Artificial Stock Market», in Arthur, Durlauf and Lane (eds) 1997, 15-44.

Arthur W. B., Durlauf S. N. and Lane D. A. (eds) 1997, *The Economy as an Evolving Complex System ii*, Reading (MA), Addison-Wesley.

BAK P., TANG C. and WIESENFELD K. 1988, «Self-Organized Criticality», *Physical Review A*, 38, 364.

BAK P., CHEN K., SCHEINKMAN J. and WOODFORD M. 1993, «Aggregate Fluctuations from Independent Sectoral Shocks», *Ricerche Economiche*, 47, 1, 3-30.

BLUME L. E. 1997, «Population Games», in Arthur, Durlauf and Lane (eds) 1997, 425-460.

Blume L. E. and Durlauf S. N. (eds) 2006, *The Economy as an Evolving Complex System iii*, Oxford, Oxford University Press.

BROCK W. A. and HOMMES C. H. 1998, «Heterogeneous Beliefs and Routes to Chaos in a Simple Asset Pricing Model», *Journal of Economic Dynamics and Control*, 22, 1235-1274.

CHALLET D., MARSILI M. and ZHANG Y.-C. 2004, *Minority Games*, Oxford, Oxford University Press.

COOLEN A. C. C. 2005, *The Mathematical Theory of Minority Games*, Oxford, Oxford University Press.

DURLAUF S. N. 1993, «Nonergodic Economic Growth», *Review of Economic Studies*, 60, 349-366.

— 1997, «Statistical Mechanics Approaches to Socioeconomic Behavior», in Arthur, Durlauf and Lane (eds) 1997, 81-104; and in *Journal of Economic Growth*, 1, 1, 75-93.

EPSTEIN J. M. 2005, «Remarks on the Foundations of Agent-Based Generative Social Science», in Tesfatsion and Judd (eds) 2006, ch. 34.

FUDENBERG D. and LEVINE D. 1998, *The Theory of Learning in Games*, Cambridge (MA), The MIT Press.

HELPMAN E. and KRUGMAN P. R. 1985, *Market Structure and Foreign Trade*, Cambridge (MA), The MIT Press.

HOLLAND J. H., HOLYOAK K., NISBETT R. and THAGARD P. 1986, *Induction: Processes of Inference, Learning, and Discovery*, Cambridge (MA), Bradford Books-The MIT Press.

IOANNIDES Y. M. 1997, «The Evolution of Trading Structures», in Arthur, Durlauf and Lane (eds) 1997, 129-167.

KIRMAN A. 1997, «The Economy as an Interactive System», in Arthur, Durlauf and Lane (eds) 1997, 491-531.

KOLLMAN K., MILLER J. and PAGE S. 1997, «Computational Political Economy», in Arthur, Durlauf and Lane (eds) 1997, 461-490.

KOPPL R. and ROSSER J. B. 2002, «All that I have to Say has already Crossed your Mind», *Metroeconomica*, 53, 339-360.

Langton Chr. G., Taylor Ch., Farmer J. D. and Rasmussen S. (eds) 1992, *Artificial Life ii*, Reading (MA), Addison-Wesley.

LANE D. A. 1993, «Artificial Worlds and Economics, Parts i and ii», *Journal of Evolutionary Economics*, 3, 2, 89-107 and 3, 3, 177-197.

LANE D. A. and MAXFIELD R. 1997, «Foresight, Complexity and Strategy», in Arthur, Durlauf and Lane (eds) 1997, 169-198.

LINDGREN K. 1992, «Evolutionary phenomena in simple dynamics», in Langton, Taylor, Farmer and Rasmussen (eds) 1992, 295-312.

LUCAS R. E. 1978, «Asset Prices in an Exchange Economy», *Econometrica*, 46, 1429.

MARIMON R., McGRATTAN E. and SARGENT T. J. 1990, «Money As A Medium Of Exchange In An Economy With Artificially Intelligent Agents», *Journal of Economic Dynamics and Control*, 14, 329-373.

MARKS R. E. 1989, «Breeding Hybrid Strategies: Optimal behavior for Oligopolists», in Schaffer (ed.) 1989, 198-207.

MILLER J. H. 1988, «The Evolution of Automata in the Repeated Prisoner's Dilemma», Ph.D. Diss., University of Michigan, Ann Arbor (MI).

PALMER R. G., ARTHUR W. B., HOLLAND J. H., LeBARON B. and TAYLER P. 1994, «Artificial Economic Life: A Simple Model of a Stockmarket», *Physica D*, 75, 264-274.

SAMUELSON L. 1997, *Evolutionary games and equilibrium selection*, Cambridge (MA), The MIT Press.

SARGENT T. J. 1993, *Bounded Rationality in Macroeconomics*, Oxford, Clarendon Press.

Schaffer J. D. (ed.) 1989, *Proceedings of the Third International Conference on Genetic Algorithms*, Fairfax (VA), George Mason University and San Francisco, Morgan Kaufmann, 198-207.

SCHELLING T. 1978, *Micromotives and Macrobehavior*, New York, W. W. Norton.

SCHUMPETER J. 1954, *A History of Economic Analysis*, New York, Oxford University Press.

TESFATSION L. 1997, «How Economists can get ALife», in Arthur, Durlauf and Lane (eds) 1997, 533-564.

TESFATSION L. 2005, «Agent-Based Computational Economics: A Constructive Approach to Economic Theory», in Tesfatsion and Judd (eds) 2006, ch. 16.

Tesfatsion L. and Judd K. L. (eds) 2006, *Handbook of Computational Economics, Vol. 11: Agent-based Computational Economics*, Amsterdam, Elsevier-North-Holland.

WILHITE A. W. 2005, «Economic Activity on Fixed Networks», in Tesfatsion and Judd (eds) 2006, ch. 20.

YOUNG H. P. 1998, *Individual Strategy and Social Structure*, Princeton (NJ), Princeton University Press.