The End of Certainty in Economics

W. Brian Arthur

Talk delivered at the Conference *Einstein Meets Magritte*, Free University of Brussels, 1994.
Appeared in *Einstein Meets Magritte*, D. Aerts, J. Broekaert, E. Mathijs, eds. 1999, Kluwer Academic Publishers, Holland.
Reprinted in *The Biology of Business*, J.H. Clippinger, ed., 1999, Jossey-Bass Publishers.

The story of the sciences in the twentieth Century is one of a steady loss of certainty. Much of what was real and machine-like and objective and determinate at the start of the century, by mid-century was a phantom, unpredictable, subjective and indeterminate. What had *defined* science at the start of the century—its power to predict, its clear subject/object distinction—no longer defined it at the end. In the century just past, science after science lost its innocence. Science after science grew up.

What then of economics? Is economics a science? I believe it is. It is a body of well-reasoned knowledge. Yet until the last few years it has maintained its certainty, it has escaped any loss of innocence. And so we must ask: Is its object of study, the economy, inherently free of uncertainties and indeterminacies? Or is economics in the process of losing its innocence and thereby joining the other sciences of this century?

I believe the latter. In fact, there are indications everywhere these days in economics that the discipline is losing its rigid sense of determinism, that the long dominance of positivist thinking is weakening, and that economics is opening itself to a less mechanistic, more organic approach. In this chapter I will show my own version of this loss of certainty. I will argue that there are major pockets of uncertainty in the economy. I will show that the clear subject/object distinction in the economics often blurs. I will show that the economy is not a gigantic machine, but a construct of its agents. These are not "anomalies" to be feared, they are natural properties of the economy, and if we accept them, we will have a stronger, not a weaker science.

High Modern Economics

The fundamental ideas in economics stem from the thinking of the eighteenth century, in particular from the thinking of the English and Scottish Enlightenment. In 1733, at the height of the intoxication of enlightenment thinking, Alexander Pope condensed its essence in one stanza of his poem, *An Essay on Man*:

All Nature is but Art unknown to Thee
All Chance, Direction, which thou canst not see
All Discord, Harmony, not understood
All partial Evil, universal Good:
And, spite of Pride, in erring Reason's spite
One truth is clear, "Whatever IS, is RIGHT."

In this context "Art" means artifice. It means technique or mechanism. And so, all the intricate wonders we see in nature, says Pope, are in fact a gigantic machine, an artifice like the mechanical automata figures of his time. All that looks unkiltered really has direction behind it. All that looks complex and discordant, like the movements of planets before Kepler's and Newton's times, has a hidden simplicity. All that affects each of God's creations adversely, in some unspoken way works to the good of the whole. Quoting Socrates, "Whatever is, is right."

These were not merely the ideas of Pope. They were the ideas that filled the intellectual air when Adam Smith was growing up. And Smith went on to enshrine them in *The Wealth of Nations*, that magnificent work that uncovered the hidden simplicity behind the traffickings of traders and manufactories and butchers and bakers. The economy was indeed Art, and its principles were now unhidden. The selfish interests of the individual were guided as by an invisible hand to the common interest of all. Whatever was, was right. Two centuries later, the philosopher of science, Jacob Bronowski, was to comment glumly that economics never recovered from the fatally rational structure imposed on it in the eighteenth century. But we inherited more than Smith's rational structure. Deep in some recess of our minds, we inherited the thinking that the economy is but Art, a gigantic machine, that if we merely understood its parts, we could predict the whole. Certainly when I was studying economics in Berkeley 25 years ago, many economists hoped (as I did) that a Grand Unified Theory of economics was possible. From the axioms of rational human behavior, a theory of the consumer could be constructed. From this and a corresponding theory of the firm we could construct a consistent microeconomics. From this, somehow, we could construct an aggregate theory of the economy: macroeconomics. All this would constitute a Grand Unified Theory of the economy.

There have always been two embarrassments to this hope of constructing a theory of the economy from its reductionist parts. One was that the economy relies on human beings, not on orderly machine components. Human beings with all their caprices and emotions and foibles. The second embarrassment was technology. Technology destroys the neatness because it keeps the economy changing. Human behavior was finessed in economics by the device of Economic Man, that perfectly rational being who reasons perfectly deductively on well-defined problems. And technology change was not so much finessed as ignored, or treated as exogenous. And so to make an orderly, predictive theory possible, Economic Man (the subject) needs to operate on well-defined Problems (the object). There should be no blurring of agent and problem. And the well-defined problems should have well-defined Solutions. And the solutions would comprise the building blocks for the next aggregated level of the theory.

This approach works. But it runs into difficulties when problems start to involve more than one decision maker and any degree of complication. Then heroic assumptions must be made. Otherwise well-definedness unravels, agent and problem become blurred, and pockets of uncertainty start to bulge.

Let me illustrate what I mean in the context of a typical microeconomic situation in modern economics. (I have chosen it from the mid-1980s literature on industrial organization.) Consider this problem: We have a circle that we might think of as a 24-hour clock. A number of firms, say twenty airline companies, have to decide in which time slot of this clock their planes will take off, say from La Guardia Airport to go to Washington. Of course the different airlines have different preferences about when to take off. They know their preferences and are going to book such take-off slots. The choices will be made once and for all. But there is a trade off (in every decent economic problem there is always a trade-off) between where they really want to take off versus not being too close to other airlines' choices of their time slots. So, given the airlines preferences, which time slots will they choose? This is the problem.

We might feel uneasy about saying much with certainty here. But I want to show the modern version of the Enlightenment approach, where we find the Harmony of a solution within the Discord of the situation. This High Modern approach is called rational expectations. I will first spell it out, then shine a bright light of realism on it, so that it starts to unravel and pockets of uncertainty appear. Let's go ahead. In the modern approach, we begin by supposing we know the order in which the airlines will submit their choices. Now imagine airline number 20, the last to choose, reasons like this: knowing where the first nineteen airlines are, I will know where I will want to be. So regardless of any arbitrary choice of the first nineteen airlines, I will know which time-slot to choose. This is an easy problem for me as the twentieth. What about airline number 19? Well, airline number nineteen, when choosing, will know the chosen positions of the previous eighteen airlines and can figure what it should do, given that the twentieth will choose an optimal position given the positions of the eighteen other airlines and number 19's choice. What about the number 18? Well, the eighteenth, knowing what the previous seventeen have chosen, arbitrarily can solve the problem of selecting an optimal placement knowing what the nineteenth will do, given that the nineteenth makes his optimal choice, given what the twentieth will do as a result of number 19's choice. Getting complicated? Yes. But you can work the whole logic in reverse order by backward deduction, or more properly by dynamic programming, and deduce how all twenty airlines will place themselves.

Notice the properties of this procedure: The problem is well defined by making it sequential and assuming the airlines use logical backward deduction. The solution is precise and clean in a mathematical sense. The problem becomes a mathematical one. (Indeed all such problems become mathematical. And economics in turn becomes mathematics.) Another property that we normally have in this kind of problem is that the individual act comes to good of the whole, that is, partial evil is universal good. It is not quite true in this case, but nevertheless this is a generic property that often holds in economics. However, the Solution comes with a lot of fine print. Airlines must know their preferences exactly. Not only that, they must know the preferences of all other airlines. Further they must know that every other airline accurately know the preferences of every other airline. They also must know that every airline knows that every airline knows the preferences of every other airline, and so on in an infinite regress. Also, each airline must be rational enough to work out the solution. Further, each airline must believe that every other airline is rational and will use perfect rationality to work out the solution. Further each airline must know in an infinite regress that every other airline is using this rational way to work out the problem, because if one of these airlines fails to do so, it messes the solution up for every other airline. Further the optimal placement of each airline using this backward deduction must be unique. If any link of this network of requirements breaks, the solution ceases to exist.

This type of multi-agent choice problem is pervasive in economics. So let us take this solution approach seriously. What if we are airline number 3 and we feel uncertain as to what airline number 17 is going to do? As airline number 3, we might say: “I don't think the people of airline number 17 are that bright, and I'm not sure whether they are going to solve this problem by this rational method. And if they don't work it out in this way then I am not sure what my optimal choice would be as the third bidder in the process.” This is sufficient to upset the situation. But worse, airline number 3 may communicate its uncertainty to other airlines and they may no longer rely on number 3 or number 17. The entire solution is starting to unravel. In fact the Solution as defined by rational expectations theory is a function of airlines’ expectations or predictions of what other airlines are going to do. The problem is that if I am a representative airline I am trying to figure out what my expectations ought to be—I am trying to predict a world that is created by the expectations of myself and everybody else. There is a self-referential loop here. The outcome each airline is trying to predict depends on the predictions it and others might form. In other words, predictions are forming a world those predictions are trying to forecast. Barring some coordinating device, by which an airline can logically determine the predictions of others (such as the tortured solution-reasoning above), there is no logical way it can determine its prediction. There is a logical indeterminacy.

So it is in the economy. People are creating a world that forms from their predictions, but if they try to form these expectations in a perfectly logical deductive way, they get into a self-referential loop. There is a logical hole in standard economic thinking. Our forecasts co-create the world our forecasts are attempting to predict. And if I do not know how others might determine their forecasts, mine are indeterminate. There are some cases in economics where it is pretty obvious that everyone can figure out what to do, where something like the above given scheme does work. But otherwise the problem is fundamental. When our ideas and preferences co-create the world they are trying to forecast, self-reference renders the problem indeterminate. The idea that we can separate the subjects of the economy—the agents who form it—from the object, the economy itself, is in trouble. Pockets of indeterminism are present everywhere in the economy. And the high modern form of economic determinism fails.

Economics under Indeterminacy

There are two questions we want to ask now. One question is: Does it matter ? Maybe all of this happens on a set of measure zero, maybe this difficulty is confined to some trivial examples in economics. The second question is: If there are pockets of indeterminism how should we proceed? To answer these I want to turn to the field of capital markets, to asset pricing theory—an area of economics that does matter. There is a well worked out efficient-market economic theory for financial markets and there is a very different set of ideas that financial traders use. Let me first outline the standard theory. The standard efficient markets theory says that all and any information hinting about the future changes of the price will be used by investors. By an argument very much like the airline argument, each stock’s price is bid to a unique level that depends on the information currently available. Using past patterns of prices to forecast future prices (technical trading), in this view, cannot lead to further profits. Otherwise the information inherent in past prices could be used to make further profits, and by assumption investors have already discounted all useful information into current prices. So the standard theory says investors use all information available to form expectations. These will determine stocks’ prices which on average will uphold these same expectations. Rational expectations again. Thus there is no way to make any money, and the market is efficient. Traders, on the other hand, believe that the market is forecastible. They believe they can spot patterns in past prices helpful in prediction—they believe in technical trading. They believe the market is anthropomorphic, that it has a psychology, that it has motives. “The market was nervous yesterday. But it shrugged off the bad news and went on to quiet down.” Economists are skeptical of this, and so the two viewpoints sit badly with each other.

The standard theory is wonderfully successful. It has its own logic. And this logic is complete and has desirable properties such as uniqueness of solution. But the standard theory must face some unexplained phenomena—or so-called empirical anomalies. Crashes and bubbles seemingly with no cause. The fact that the volume of market trades is an order of magnitude higher than theory predicts. The fact that econometric tests show that that technical trading is indeed profitable statistically (Brock, Lakonishok, and Le Baron). The phenomenon of GARCH behavior, (GARCH means Generalized Auto Regressive Conditional Heteroscedasticity), which means there are periods of high volatility in stock prices interspersed randomly with periods of quiescence. In sum, the standard theory does not explain at least half a dozen major statistical "anomalies" in real markets. This has recently led to a great deal of modern thinking, some using ad-hoc behavioral observation, some more sophisticated theorizing.

Let me now show, as in the airline problem, how the standard theory breaks down and leads to pockets of indeterminacy. Suppose investors can put some portion of their money in a single stock that pays a dividend every time period (a day, a year, say), and they cannot perfectly predict this dividend. The investors are buying the stock for the dividend plus any capital appreciation (tomorrow's price), and they face the problem of forecasting these. To make the standard solution work, we assume homogeneous, identical investors who have identical forecasts of the dividend at the end of the period and identical forecasts about the stock's price in the future—forecasts that are on average unbiased and are therefore rational expectations. A little economic reasoning then shows today's price is equal to the common expectation of tomorrow's price plus dividend (suitably discounted and weighted). This yields a sequence of equations at each time, and with a pinch or two of conditional-expectation algebra, we can solve these for the expectations of future prices conditioned on current information, and wind up with today's price expressed as a function of expected future dividends. The problem is solved. But it is only solved, providing we assume "identical investors who have identical forecasts of the dividend at the end of the period and identical forecasts about the stock's price in the future." But what if we don't? What if we assume investors differ?

Let us look at the same exercise assuming our investors agents are not homogeneous. Note that the standard theory's requirement of identical "information" means not just the same data seen by everyone, but the same interpretation of the data. But imagine yourself in a real financial market, like the New York stock market. Then this information consists of past prices and trading volumes, moves made by large mutual funds or large pension funds, rumors, CNN, network news, the market section of the Wall Street Journal, what other traders are doing, what they are telling you by telephone, what your friend's uncle thinks what is happening to the market. All of these things compromise actual information and it is reasonable to assume that, even if everybody has identical access to all this information, they would treat this information as a Rorschach inkblot and would interpret it differently. Even if we assume that the people interpreting this information are intelligent to any arbitrarily high degree and they are all perfectly trained in statistics, they will *still* interpret this data differently because there are many different ways to interpret the same data.

So there is no single expectational model. A given investor can still come up with an individual forecast of the dividend. But tomorrow's price is determined by this investor's and other investors' individual forecasts of the dividend and of next period's price. And there is no way for our reference investor to fathom the forecasts of the others—to figure "what average opinion expects the average opinion to be" (to use Keynes' words). To do so brings on a logical regress. "I think that they might think such and such, but realizing that I think that, they will think this." Unless we assume identical investors, once again our agents are trying to forecast an outcome (future price) that is a function of other agents' forecasts. As before there is no deductive closure. Expectations become indeterminate, and the theory collapses.

Worse, expectations become unstable. Imagine that a few people think that prices on the market are going to go up. If I believe this and I believe that others believe this, I will revise my expectations upward. But then I may pick up some negative rumor. I will reassess downward, but realizing that others may reassess and that they too realize that others, I may further reassess. Expectations become fugitive, rippling up or down whether trades are made or not. Predictions become unstable. This is the way price bubbles start. If somehow people expect prices to go up, they will forecast that other people will forecast that prices will go up. So they will buy in, and once the bubble thus starts off, people can see prices go up and their expectations of upward motion fulfilled. Therefore prices may continue to go up. Similar logic applies to "floors" and "ceilings." If, for example, the price is 894, many investors believe that at 900 there is some sort of membrane, a ceiling, and when the price reaches this ceiling it will bounce back down with a certain probability or it may "break through." Such ideas seem strange at first. But it is quite possible that many investors have sell orders at 900, simply because it is a round number. So expectations that the price will fall if it hits 900 are likely to be fulfilled. Ceilings and floors emerge as partially self-fulfilling prophesies, held in place by their being convenient sell and buy places. We are now a long way from homogeneous rational expectations. Under the realistic assumption that traders may interpret the same information differently, expectations become indeterminate and unstable. And they may become mutually self-fulfilling.

To summarize all this: If we look at a serious branch of economics, the theory of capital markets, we see the same indeterminacy that we saw in the airline problem. Agents need to form expectations of an outcome that is a function of these expectations. With reasonable heterogeneity of interpretation of "information," there is no deductive closure. The formation of expectations is indeterminate.

And yet in every market, in every day, people *do* form expectations. How do they do this? If they can not do this deductively, then is it possible to model their behavior in this area ?

In 1988, John Holland and I decided that we would study situations like this by forming an artificial stock market in the computer and giving the little agents—artificially intelligent computer programs—some means by which they can do the reasoning that is required. This was one of the very earliest artificial, agent-based markets. Later we brought in Richard Palmer who is a physicist, Paul Tayler who is a finance expert and Blake LeBaron who is a financial theorist in economics.

In this market there was no feed-in from the real stock market. It was an artificial world going on inside the machine. The artificial agents, the little artificial investors, are all buying and selling a "stock" from one another. The computer could display the stock's price and dividend, who is buying and selling, who is making money and who is not, who is in the market and who is out, and so on. The price is formed within the machine by bids and offers. And another little program—a specialist—sets the price to clear the market, as in actual stock markets.

The modeling question was: If the agents cannot form their expectations deductively, how are they going to form them? We decided to follow modern cognitive theory about how actual human beings behave in such situations. So we allowed our artificial agents looking at to posit multiple, individual hypothetical models for forecasting, and to test these on a continual, ongoing basis. Each of these hypotheses has a prediction associated with it. At any stage each agent uses the most accurate of its hypotheses, and buys or sells accordingly. Our agents learn in two ways: they learn which of their forecasting hypotheses are more accurate, and they continually toss out ones that do not work and replace these using a genetic algorithm. So they are learning to recognize patterns they are collectively creating, and this in turn collectively creates new patterns in the stock price, which they can form fresh hypotheses about. This kind of behavior—bringing in hypotheses, testing them, and occasionally replacing them—is called induction. Our agents use *inductive rationality*—a much more realistic form of behavior. Very well. But now the key question is: Does our market converge to the rational expectations equilibrium of the academic theory or does it show some other behavior? What we found to our surprise was that two different regimes emerged. One, which we called the *rational expectations regime*, held sway when we started our agents off with sets of predictive hypotheses close to rational expectations. We could plot the parameters of all the predictive hypotheses on a chart, and in this case, over time, we could watch them getting gravitationally pulled into the orbit of the rational expectations solution, forming a "fuzz" around this point, as they made occasional predictive forays away from rational expectations to test different ideas. It is not hard to see why rational expectations prevailed. If the overall mass of predictions is near rational expectations, the price sequence will be near rational expectations, and non-rational expectations forecasts will be negated. So the academic theory was validated.

But there was a second regime, which we called the *complex regime*, and it prevailed in a much wider set of circumstances. We found that if we started our agents with hypotheses a little removed from rational expectations, or alternatively, if we allowed them to come up with hypotheses at a slightly faster rate then before, the behavior of the market changed. Subsets of mutually reinforcing predictions emerged. Imagine, for example, we have a 100 artificial agents each using 60 different prediction formulas, so that there is a universe of some 6,000 predictors. Some of the predictors that emerge are mutually reinforcing, some are mutually negating. Suppose many predictors arise that say the stock price cycles up and down over time. Such predictors would be mutually negating because they will cause agents to buy in at the bottom of the cycle, and sell at the top of the cycle, mutually negating profits, and therefore eventually disappearing from the population of predictors. But if a subset of predictors emerged by chance that said "the price will rise next period if it has risen in the last three periods," and there were enough of these, they would cause agents to buy, which on average would cause the price to rise, reinforcing such a sub-population. Such subsets could then take off, and become embedded in the population of predictors. This was what indeed happened in the complex regime, endowing it with much richer set of behaviors. Another way to express this is that our artificial traders discovered forms of technical trading that worked. They were using, with success, predictions based upon past price patterns. And so technical trading was emergent in our artificial stock market. This emergence of subsets of mutually reinforcing elements, strangely enough, is reminiscent of the origin of life, where the emergence of subpopulations of RNA in correct combinations allows them to become mutually enforcing.

Another property that emerged in the complex regime was the so-called GARCH behavior I mentioned earlier that occurs in real markets—periods of high volatility in the stock price followed by periods of quiescence—which is unexplained in the standard model. How did GARCH become an emergent property? What happens in our artificial market is that every so often some number of investors discover a new way to do better in the market. These investors then change their buying and selling behavior. This causes the market to change, even if slightly, possibly causing other investors in turn to change. Avalanches of change sweep through the market, on all scales, large and small. Thus emerge periods of change triggering further change—periods of high volatility—followed by periods when little changes and little needs to be changed, periods of quiescence. This is GARCH behavior.

Let me now summarize. What we found in our artificial stock market is that, providing our investors start near the academic rational-expectations solution, this solution prevails. But this is a small set of parameter space. Outside this, in the complex regime, self-reinforcing beliefs and self-reinforcing avalanches of change emerge. A wider theory and a richer "solution" or set of behaviors then appears, consonant with actual market behavior. The rational-expectations theory becomes a special case.

In the standard view of the economy, which has an intellectual lineage that goes back to the enlightenment, the economy is mechanistic. It is complicated but can be viewed as a series of objects and linkages between them. Subject and object—agents and the economy they perform in—can be neatly separated. The view I am giving here is different. It says that the economy itself emerges from our subjective beliefs. These subjective beliefs, taken in aggregate, structure the micro economy. They give rise to the character of financial markets. They direct flows of capital and govern strategic behavior and negotiations. They are the DNA of the economy. These subjective beliefs are *a-priori* or deductively indeterminate in advance. They co-evolve, arise, decay, change, mutually reinforce, and mutually negate. Subject and object can not be neatly separated. And so the economy shows behavior that we can best describe as organic, rather than mechanistic. It is not a well-ordered, gigantic machine. It is organic. At all levels it contains pockets of indeterminacy. It emerges from subjectivity and falls back into subjectivity.