Is Economics the Next Physical Science?

An emerging body of work by physicists addressing questions of economic organization and function suggests new approaches to economics and a broadening of the scope of physics.

J. Doyne Farmer, Martin Shubik, and Eric Smith

In the past decade or so, physicists have begun to do academic research in economics. Perhaps a hundred people are now actively involved in an emerging field often called econophysics, and two new journals and frequent conferences are devoted to the field. At least ten books have been written recently on econophysics in general or on specific subtopics. The University of Fribourg in Switzerland maintains an extensive bibliography of books and archived articles at its econophysics website, http://www.unifr.ch/econophysics. Physics departments worldwide are granting PhD theses for research in economics, and in Europe several professors in physics departments specialize in econophysics. The international consulting firm McKinsey and Co sponsors a new annual research prize, the Young Scientist Award for Socio- and Econophysics.

Is all this activity just a fad, or is something more substantial happening?

If physicists want to do research in economics, why don’t they just get degrees in economics in the first place? Why don’t the econophysicists retool, find jobs in economics departments, and publish in traditional economics journals? Perhaps the growth of econophysics is just a temporary phenomenon, driven by a generation of physicists who made bad career choices. Is there any reason why members of economics departments should pay heed to the methods of physics? What advantage, if any, is conferred by a background in physics? And most important, how does econophysics differ from economics, and what unique contribution can it make?

Social physics

The involvement of physicists in social science has a long history, going back at least to Daniel Bernoulli, who introduced in 1738 the idea of utility to describe people’s preferences. Pierre-Simon Laplace, in his Essai philosophique sur les probabilités (1812), pointed out that events that might seem random and unpredictable, such as the number of letters in the Paris dead-letter office, can be quite predictable and can be shown to obey simple laws. Laplace’s ideas were further amplified by Adolphe Quetelet, who was a student of Joseph Fourier and who studied the existence of patterns in data sets ranging from the frequency of different methods for committing murder to the chest size of Scottish men. It was Quetelet who, in 1835, coined the term “social physics.” Analogies to physics played an important role in the development of economic theory through the 19th century, and some of the founders of neoclassical economic theory, including Irving Fisher, were originally trained as physicists; Fisher was a student of Willard Gibbs. In 1938, Ettore Majorana presciently outlined both the opportunities and pitfalls in applying statistical physics methods to the social sciences.

The range of topics that have been addressed by physicists spans many different areas of economics. Finance is particularly well represented (see the article by Joseph Pimbley, PHYSICS TODAY, January 1997, page 42); sample topics include the empirical observation of regularities in market data, the dynamics of price formation, the understanding of bubbles and panics, methods for pricing options and other derivatives, and the construction of optimal portfolios. Broader topics in economics include the distribution of income, the emergence of money, and implications of symmetry and scaling for market functioning. We believe that a union of the methods of physics and economics, and collaborations between physicists and economists, can add value to the science of economics. However, overselling that view has its dangers; econophysics is far from well established.

Despite the fields’ long history of association, the substantial contribution of physics to economics is still in an early stage, and we think it fanciful to predict what will ultimately be accomplished. Almost certainly, “physical” aspects of theories of social order will not simply recapitulate existing theories in physics, though already there appear to be overlaps. The development of societies and economies can be contingent on accidents of history and at every turn hinges on complex aspects of human behavior.

Nonetheless, striking empirical regularities suggest that at least some social order is not historically contingent, and is perhaps predictable from first principles. The role of markets as mediators of communication and distributed computation, which underlie the collective processes of price formation and allocation of resources, and the emergence of the social institutions that support those functions, are quintessentially economic phenomena. Yet the notions of markets’ communication or computational capacities, and the way differences in those capacities account for the stability and historical succession of markets, may naturally be part of the physical world with its human social dynamics.

Markets and other economic institutions bring with them concepts of efficiency or optimality in satisfying human desires. While intuitively appealing, such ideas have proven hard to formalize even if some progress has been made. As with most new areas of physical inquiry, we expect that the ultimate goals of a physical economics will be declared with hindsight, from successes in identifying, measuring, modeling, and in some cases predicting empirical regularities.

The search for universality

Economists are typically better trained than physicists in statistical analysis, so physicists might seem to have little
to contribute in that area. However, differences in goals and philosophy are important. Physics is driven by the quest for universal laws. Partly because of the extreme complexity of social phenomena, that quest has been largely abandoned in our postmodern world, where relativist philosophies of science enjoy disturbingly widespread acceptance. Nowadays, work in social science is largely focused on documenting differences. Although the trend is much less obvious in economics, a typical paper in financial economics, for example, might study the difference between the New York and NASDAQ stock exchanges, or the effect of changing the tick size, or smallest change, of prices. Physicists have entered the field, perhaps naively, with fresh eyes and new hypotheses, and have looked at economic data with the goal of finding pervasive regularities; they emphasize what might be common to all markets rather than what makes them different. Such work has been opportunistically motivated by the existence of large data sets like complete major-exchange transaction histories that span years and that sometimes contain hundreds of millions of events.

Much of the work by physicists in economics concerns power laws. A power law is an asymptotic relation of the form \( f(x) \rightarrow x^{-\alpha} \), where \( x \) is a variable and \( \alpha > 0 \) is a constant. In many important cases, \( f \) is a probability distribution. The first power-law distribution in any field was observed in economics (see the box on page 39), and the existence and significance of economic power laws has been a matter of debate ever since. In 1963, Benoit Mandelbrot observed that the distribution of cotton price fluctuations follows a power law. Later, more detailed observations of power laws were made by Rosario Mantegna and Eugene Stanley, who coined the term econophysics. Figure 1 shows the striking fidelity often found in economic power laws. The existence of power laws in price changes is interesting from a practical point of view because it has implications for the risk of financial investments, and from a theoretical point of view because it suggests scale independence and possible analogies to critical phenomena and nonequilibrium behavior in the processes that generate financial returns.

Many other economic power laws have been discovered by physicists. They describe the variance in the growth rates of a company as a function of the company’s size, the distribution function for the number of shares in a transaction, the distribution function for the number of trading orders submitted at a specified price from the best price offered, and the size of the price response to a trade as a function of the size of the company being traded. The question of why power laws are ubiquitous in financial markets has stimulated a great deal of theoretical work.

### Long memories

One of the most famous and most used price models is the random walk, introduced for prices by Louis Bachelier, a student of Henri Poincaré, in 1900; that was five years before Albert Einstein used it to describe Brownian motion. The random walk is the basis for the Black–Scholes theory of option pricing, which won Robert Merton and Myron Scholes the 1997 Nobel memorial prize in economics. Since then, physicists have incorporated into extended analytic methods for pricing options more realistic assumptions about how prices behave.

An interesting and surprising property of the random walk for real prices is that the diffusion rate is not constant. The size of the price change at a time \( t \) is correlated to that at \( t + \tau \) even though the directions of the price changes are uncorrelated: This phenomenon is called clustered volatility. The correlation decays as a power law of the form \( \tau^{-\gamma} \). Since \( 0 < \gamma < 1 \), the size of price change is a long-memory process and thus displays anomalous diffusion and a very slow convergence of statistical averages.

Physicists have discovered that volatility is just one of several long-memory processes in markets. One of the most surprising of those concerns fluctuations in supply and demand. If one assigns +1 to an order to buy and −1 to an order to sell, the resulting series of numbers has a positive correlation that decays as a power law with an exponent of about 0.6. The positive correlation persists at statistically significant levels over tens of thousands of orders, for weeks. Thus change in supply and demand is a long-memory process, which has interesting implications for a fundamental concept in financial economics called market efficiency.

The principle of market efficiency takes many forms. A market is informationally efficient if prices reflect all available information; it is arbitrage efficient if it is impossible for investors to make excess profits; and it is allocationally efficient if prices are set to maximize everyone’s welfare. One of the consequences of informational efficiency is that prices should not be predictable. In reality, the assumption of price unpredictability is not bad—even the best trading strategies exploit only very weak levels of predictability.

The coexistence of long-memory supply and demand with market efficiency creates an as yet unresolved puzzle. Long-memory processes are highly predictable. Since the entrance of buyers tends to drive prices up, and the entrance of sellers tends to drive them down, the long memory of supply and demand suggests that price change should also be long memory. But that would be incompatible with informational efficiency. To prevent such incompatibilities, the agents in the market must somehow collectively adjust their behavior: They might, for example,
Income Distributions

In 1897 Vilfredo Pareto identified the first power-law distribution in any field. He was studying the distribution of income among the highest-earning inhabitants of the UK; now all income distributions of the form he observed are known in economics as Pareto distributions. Pareto’s subsequent studies of income distribution in Prussia, Saxony, Paris, and a few Italian cities confirmed his initial observations. The figure, courtesy of Makoto Nirei and Wataru Souma, shows modern Pareto distribution examples derived from federal income tax reporting sources for the US and Japan. (The data for the US are truncated at $1 million.) Studies conducted over the past few years have shown that not only does the income of the wealthy conform to a pattern, but so does the income of the majority of wage earners, and the two groups follow different distributions. The low- and medium-income body of the distribution is either exponential, as is the case for the US, or log-normal, as is the case for Japan. Just where the transition to the Pareto law for large incomes takes place depends on time, tax laws, and other factors as yet unknown.

As striking as the scale-free nature of income distribution is the fact that the overall distribution is so featureless. It is described by four or five parameters: the mean income, the Pareto exponent, the transition point between low- and high-income ranges, and either the exponential constant or the mean and variance of the log-normal in the low range. Pareto, log-normal, and exponential distributions are all limiting distributions of simple random processes and can be derived as maximum-entropy distributions for income, subject to appropriate boundary conditions.

Income distribution is a hot topic economically and politically, because it lies at the heart of a society’s notions of egalitarianism, opportunity, and social insurance. Not surprisingly, causes of income inequality such as distinctions between capital ownership and wage labor are asserted, and major policy implications are derived. Maximum-entropy interpretations of income distribution place conceptual and quantitative bounds on such arguments. They suggest that the many detailed societal features that could, in principle, affect incomes average out somehow so that their individual characteristic scales are not imprinted on the aggregate distribution. The ultimate constraints may be conservation laws or boundary conditions reflected in at most a few parameters. Such featureless averaging, like scaling relations, may suggest that a form of universality classification is fundamental to economics, as it is to thermodynamics and field theory.

The central approach to strategic interactions in neoclassical economics is the theory of rational choice. In the economists’ stylized version, a rational individual maximizes some measure of personal—usually material—welfare, with perfect knowledge of the world and of other agents’ goals and abilities, and with the capacity to perform computations of any complexity. When agent A considers any strategy, agent B knows that A is considering that strategy, and A knows that B knows that A knows, and so on. This infinite regress appears very complicated. However, a key simplifying result is that in any game there exists at least one Nash equilibrium, which is a strategy that is the best defense against itself. Every Nash equilibrium is a fixed point in the space of strategies, and it circumvents the infinite regress problem by imposing self-consistency as a defining criterion. Subject to several caveats, rational players who are not cooperating with each other will choose a Nash strategy—that is the operational meaning of rational choice. The assumption that decisions of real human beings can be approximated by rational choice dominated microeconomics—economic thinking about individual choices—from about 1950 until the mid-1980s, though it is clearly implausible for all but the simplest cognitive tasks. It also leaves unaddressed the problem of macroeconomics, the aggregation of individual choices and the behavior of large populations.

Imperfect rationality

In the past 20 years, economics has begun to challenge the assumptions of rational choice and perfect markets by modeling imperfections such as asymmetric information, incomplete market structure, and bounded rationality. Several new schools of thought have emerged. The behavioral economists attempt to take human psychology into account by studying people’s actual choices in idealized economic...
Another school addresses bounded rationality with idealizations of problem solving and learning that range from standard statistical methods to artificial intelligence. Agent-based modeling focuses on the complexity of economic interactions by using computer simulations based on idealizations of human behaviors. Yet another approach assumes that some agents, called noise traders, have extremely limited reasoning capabilities while others are perfectly rational. Physicists have joined with economists to seek new theories of not fully rational choice, and have brought new perspectives to bear on the problem.

An early effort using both agent-based modeling and artificial intelligence was the Santa Fe Artificial Stock Market. It grew out of a 1986 conference, organized by Kenneth Arrow, Philip Anderson, and David Pines, that brought together physicists and economists. Pre-saging the modern move toward behavioral economics, the physicists all expressed disbelief in theories of rational choice and suggested that the economists should take human psychology and learning more into account. The Santa Fe Artificial Stock Market, a collaboration among economists, physicists, and a computer scientist, replaced the rational agents in an idealized market setting with an artificial-intelligence model. That substitution led to qualitative modifications of the statistics of prices—for example, fat-tailed distributions of price change and clustered volatility—and suggested that nonrational behavior plays an important part in generating those modifications.

The problem with approaches exemplified by the Santa Fe Artificial Stock Market is that they are complicated, and although they capture some qualitative features of markets, they have not led to more quantitative theories. Agent-based models tend to require ad hoc assumptions that are difficult to validate.

The hypothesis of rational choice, in contrast, has the great virtue of parsimony. It makes strong predictions from simple hypotheses and, in that sense, is more like a physics theory. This has motivated a search for other simple parsimonious alternatives, one of which is often called zero intelligence. If rational choice enters the wilderness of bounded rationality from the top, zero intelligence enters it from the bottom. Agents with zero intelligence behave more or less randomly, subject to constraints such as their budget. Zero-intelligence models can be used to study the properties of market institutions and to determine which properties of a market depend on intention and which don’t. Thus they provide a benchmark to help avoid getting lost in the large space of realistic human behaviors. Once a zero-intelligence model has been constructed, it can be made more realistic by adding a little intelligence based on empirical observations or models of learning.

The zero-intelligence approach can be traced back to the work of Herbert Simon, a Nobel laureate in economics and a pioneer in artificial intelligence. The model’s main champions in recent years have been physicists, who have used analogies to statistical mechanics to develop new models of markets. A good example is the work of Per Bak, Maya Paczuski, and economist Martin Shubik (BPS), who studied the impact of random trading orders on prices in the context of an idealized model of price formation. They assumed that traders simply place orders to buy or sell at random, above or below the prices of the most recent transactions. Traders modify their orders from time to time, moving them toward the middle until they generate a transaction. The result is mathematically analogous to a reaction-diffusion model for a scenario in which the reactants annihilate each other. Although the BPS model is highly unrealistic, with a few modifications it produces some qualitative features—heavy-tailed price distributions, for example—that resemble their counterparts in real markets.

By now, econophysicists have explored many variations of the BPS model. In one variation, also based on random trading orders, Marcus Daniels and colleagues performed a
demand curves, as shown in figure 2. Reference 9 has also given insight into the shape of supply and technology. Agents are capable of remembering the outcomes rewarded, or awarded payoffs, in game-theoretic terms. Agents whose choice is the minority value are chosen to be selected for rounds 3–10. Table a gives the successful value three times (red) and table b shows success seven times (blue). Thus, the strategy of table b has the better cumulative performance. An agent playing the game will consult table b and choose “0” for the 11th round.

Figure 3. Lookup tables for the minority game in which agents base their strategies on the previous two rounds of play. (a) The table summarizes the strategy of “No matter what happened in the previous two rounds, choose 1.” (b) Here the strategy is, “If the previous two rounds gave the same result, choose the complement. If the previous two rounds differed, choose the most recent winning value.” Suppose the first 10 rounds of play yield the values 1100100110. The first two rounds serve as initial input and the lookup tables give values to be selected for rounds 3–10. Table a gives the successful value three times (red) and table b shows success seven times (blue). Thus, the strategy of table b has the better cumulative performance. An agent playing the game will consult table b and choose “0” for the 11th round.

Minority games
As another alternative to rational choice, econophysicists have developed highly simplified models of strategic interaction to capture the essence of the collective behavior in a financial market. Brian Arthur’s El Farol bar problem provides an alternative to conventional game theory. El Farol is a bar in Santa Fe that is often crowded. In Arthur’s game, agents decide each day whether to go hear music. If the bar has room, they are happy; if it is too crowded, they are disappointed. The game is constructed so that only a minority of the people can be happy, which leads to a phenomenon analogous to frustration—some desires are necessarily unsatisfiable. As an additional result of the game’s structure, an astronomically large number of equilibria can emerge.

The El Farol game was simplified and abstracted by Damien Challet and Yi-Cheng Zhang as the minority game,11 in which an odd number $N$ of agents repeatedly choose between two alternatives, which can be labeled 0 or 1. Their decisions are made independently and simultaneously. Agents whose choice is the minority value are rewarded, or awarded payoffs, in game-theoretic terminology. Agents are capable of remembering the outcomes of $M$ prior rounds of play, and they maintain an inventory of strategies, called lookup tables, that dictate a next move for each history. The lookup tables are generated at random, but the strategy used in any given round is the one with the best cumulative performance (see figure 3).

If the agents have at least two strategies, the minority game will exhibit phase transitions in $z = 2^N/N$, the ratio of the number of resolvable pasts to the number of agents. Figure 4 shows such a phase transition: If $z$ is less than a critical value $z_c$, the population is in a symmetric phase and the outcome of the next move is unpredictable from the history of play. In contrast, for $z > z_c$, the $N$ agents sparsely sample the space of strategies, the next outcome is statistically predictable from the history of play, and the population is in a symmetry-broken phase that can be understood analytically with so-called replica methods. The variance in the number of winners about the optimum, $(N - 1)/2$, measures the failure of allocative efficiency and is minimized at $z_c$.

The minority game is readily extended to incorporate various features of real financial markets. For example, buyers or sellers of stocks can reap larger profits when they are providing the more scarce of supply or demand, and so the minority game can include payoffs that increase as the size of the winning group gets smaller. In the “grand canonical” version of the game, players may enter and leave. With such enhancements, the game self-organizes around the critical point $z_c$. Moreover, as in real markets, payoffs display clustered volatility and have a distribution with a power-law tail. The minority game provides a fascinating example of how a very simple game can display a rich set of properties as soon as one moves away from rational choice.

Entropy methods
Finance is not the only area of economics in which physicists are active. Economics, like physics, distinguishes open systems from closed ones, and that distinction gives rise to different notions of equilibrium. Markets considered merely as conduits for goods produced or consumed elsewhere are described by theories of partial equilibrium that are largely specified by open-system boundary conditions. Financial markets are open in this sense. Economists also try to determine the general equilibria of whole societies, taking into account not only trade but production, consumption, and, to some extent, government regulation.

Economics and physics have together seen a growth in understanding of relaxation to equilibrium, when equilibria are possible, and whether they are unique. In both fields, mechanical models were used first, followed by statistical explanations.12 A recent paper has shown which subset of economic decision problems has a structure identical to that of classical thermodynamics and has discussed the emergence of a phenomenological principle equivalent to entropy maximization.13 More general equilibrium problems usually considered by economists correspond to physical problems with many equilibria, such as granular, glassy, or hysteretic relaxation. The idea that equilibria correspond to statistically most-probable sets of configurations has led to attempts to define price formation in statistical terms. As discussed in the box, the related idea that income distribution seems consistent with various forms of entropy maximization recasts the problem of understanding income inequality and interpreting how much it really tells about the social forces that affect incomes.

We expect that maximum-ignorance principles will grow into a conceptual foundation in economics as they have in physics, and that the roles of symmetry, conservation laws, and scaling will become increasingly important.14 Efforts to explain which aspects of market function or regulatory structure converge on predictable forms, in a manner relatively free of historical contingency, are likely to require characterization in those basic terms.
**Future directions**

We expect that within the next few years some physics and economics departments will design a basic course teaching the essential elements of both physics and economics. Physicists will continue to contribute to economics in a variety of subjects ranging from macroeconomics to market microstructure, and their contributions will have increasing implications for economic policy making.

One area of opportunity, in which the applicability of physics might not be evident a priori, concerns the construction of economic indices, such as the Consumer Price Index or the Dow Jones Industrial Average. Those indices provide only crude scalar summaries of very complex phenomena, but they play an important role in economic decision making. For example, pension and wage payments are based on the CPI.

Economic indices are currently constructed using essentially ad hoc methods. We believe that the accuracy of such indices could be improved by careful thinking in terms of dimensional analysis, combined with better data analysis that correlates prices and other factors to appropriate qualities related to the indices. The construction of accurate economic indices is ultimately related to understanding why the economy exhibits so many scale-free behaviors, such as the distribution of wealth or the size of firms. To shed light on that question, econophysicists need to better understand the natural dimensions of economic life, and systematic dimensional analysis will be a very useful tool. Dimensional and scaling methods were a cornerstone in the elucidation of complex phenomena such as turbulence in fluids, and all the constituents that make fluid flow complex—long-time correlations, nonlinearity, and chaos—are likely to be important factors in the economy.

At the other end of the macro–micro spectrum, ideas from statistical mechanics could make practical contributions to problems in market microstructure. Physics-style models suggest, for example, that price volatility could be lowered if market rules are changed to create incentives for “patient” trading orders that are not immediately transacted. A related practical problem concerns the optimal strategy for market makers, that is, agents who simultaneously buy and sell and make a profit by taking the difference. Though markets are increasingly electronic, automated market makers are still designed in a more or less ad hoc manner. A theory for market making based on methods from statistical mechanics could result in lower transaction costs and generally more efficient markets.

Several key ideas in physics are of economic origin. A prejudice that the books should balance was likely responsible for James Joule's accounting for heat's energy content before the principle of conservation of energy was well supported by data. The concept of a currency informs scientists’ view of the role of energy in complex systems, particularly in biochemistry. Understanding the dynamics and statistical mechanics of agency promises to expand the conceptual scope of physics.

**References**