What Economists can learn from physics and finance

Joseph L. McCauley

University of Houston

October 2004

Online at http://mpra.ub.uni-muenchen.de/2240/
1. Modeling in Finance and Economics

Some economists (Mirowski, 2002) have asserted that the neo-classical economic model was motivated by Newtonian mechanics. This viewpoint encourages confusion. Theoretical mechanics is firmly grounded in reproducible empirical observations and experiments, and provides a very accurate description of macroscopic motions to within high decimal precision. In stark contrast, neo-classical economics, or ‘rational expectations’ (ratex), is a merely postulated model that cannot be used to describe any real market or economy, even to zeroth order in perturbation theory. In mechanics we study both chaotic and complex dynamics whereas ratex restricts itself to equilibrium. Wigner (1967) has isolated the reasons for what he called ‘the unreasonable effectiveness of mathematics in physics’. In this article we isolate the reason for what Velupillai (2005), who was motivated by Wigner (1960), has called the ineffectiveness of mathematics in economics. I propose a
remedy, namely, that economic theory should strive for the same degree of empirical success in modeling markets and economies as is exhibited by finance theory.

2. Existence Proofs without Dynamics are Dangerously Misleading

I begin with a topic of much interest to an economist: existence proofs of equilibrium in the absence of dynamics may be completely misleading. I provide an example to back up my claim.

Consider Osborne’s model of lognormal market prices, used by Black and Scholes (1973) to price options based on the assumption of Gaussian returns. The stochastic differential equation generating the model is

$$dp = rpdt + p\sigma_p dB(t)$$

(1)

where $r$ and $\sigma_p$ are constants, and $dB(t)$ is the Wiener process. If we would take $r<0$, negative expected gain rate, then the drift term would provide us with an example of a restoring force, an example of the Invisible Hand (McCauley, 2004). Does the Invisible Hand pull the market toward equilibrium? From (1), the corresponding Fokker-Planck equation describing the price density $g(p,t)$ is
and indeed has a very simple equilibrium solution $g(p)$ with fat tails in price $p$. However, the time dependent solution of (2), the lognormal density $g(p,t)$, spreads without limit as $t$ increases and does not approach statistical equilibrium at all! In particular, the second moment $<p^2>$ increases without limit. The reason that equilibrium is not approached is that the spectrum of the Fokker-Planck operator defined by (2) is continuous, not discrete. Imposing finite limits on $p$, price controls, would yield a discrete spectrum so that statistical equilibrium would follow asymptotically. We therefore expect that market equilibrium and stability are inconsistent with deregulation.

Having whetted the reader’s appetite, let me now get down to business.

3. Invariance Principles and Mathematical Laws of Motion

Data collection and analysis are central to physics. Data collection in the attempt to describe the motion of bodies began with the ancient astronomers, who used epicycles to describe planetary orbits. The mathematical description of empirically discovered laws of nature began with Archimedes’ discovery of the conditions for static equilibrium. Galileo and Kepler revived
the Archimedian tradition in the seventeenth century and provided the empirical discoveries from which Newton was able to formulate nature’s dynamics mathematically in a very precise and general way (Barbour, 1989). Why have we been able to discover strict mathematical laws of inanimate nature, but haven’t discovered corresponding mathematical laws of socio-economic behavior? Wigner (1967) discussed these questions in his beautiful essays on symmetry and invariance, where he identifies the basis of the seemingly unreasonable effectiveness of mathematics in physics.

Following Galileo and Kepler, scientists have discovered mathematical laws obeyed by nature via repeatable, identical experiments (physics, chemistry, genetics) and repeatable observations (astronomy). The foundation for the invariance of experimental results performed at different places and times and in different states of motion lies in the local symmetry principles that form the basis of Newtonian and quantum mechanics, and general relativity: the simplest predictions of mathematical laws of nature are invariant under translations, rotations, time translations, and transformations among Galilean/inertial frames. These symmetries produce periodic orbits in integrable systems like the Newtonian two-body problem. Laws of nature were first discovered by Galileo and Kepler from careful observations of very simple orbits of period zero and period one.
Given enough symmetry principles obeyed by prices, we should in principle be able to discover mathematical laws obeyed by markets. We know only one invariance principle for markets, and will discuss it below. There’s a fundamental difference between economic motions, like price changes (or GNP growth), and motions of inanimate bodies described by inviolable mathematical laws of nature.

Unlike natural law, we act on human wishes and expectations to create all of economic behavior. Without actions determined by our brains, wishes, and actions, markets and prices would not exist. Nature, e.g., stars, planets, atoms, and DNA are not invented and manipulated in that way. *Mathematical laws of nature are beyond human invention, intervention, and convention.* Without human agreement and/or regulation, in contrast, markets and prices do not exist. Given that human decisions and actions create markets and money, and even that self-fulfilling prophecy is possible, to what extent can we hope to discover an even approximately correct dynamics of markets? And bear in mind that nonuniqueness due to limited precision in data analysis can lead us not to a single model, but at best to some nonuniversality class of models. That is still better than nonempirical postulation, which exhibits far worse nonuniqueness problems.

4. Invariance principles in markets
Let’s start with the dynamics of price \( p \) and quantity \( x \) of assets in some real market or in a hypothetical market model,

\[
\frac{dp}{dt} = \varepsilon(p,t)
\]

(3)

where \( \varepsilon(p,t) = D(p,t) - S(p,t) \) is the excess demand, and \( x_D = D(p,t) \) and \( x_S = S(p,t) \) are the demand and supply at price \( p \) respectively. **There is only one dynamically correct definition of equilibrium:** nothing changes with time. In deterministic dynamics, \( dp/dt = 0 \), or excess demand vanishes.

For a stochastic description of markets, as in parts 2, 5, and 6, the condition \( d\langle p \rangle/dt = 0 \) is necessary but not sufficient for equilibrium, where \( \langle ... \rangle \) denotes the average. Also necessary for equilibrium is that all moments of the price distribution are time independent, which means that the price distribution \( g(p,t) \) is time invariant. **No other definition of equilibrium is consistent with dynamics.** Contrary to confusion rampant in the economics and finance literature (see, e.g., McAdam and Hallett, 2000), a limit cycle is not an equilibrium, nor is a strange attractor. Neither a Wiener, lognormal, nor Levy stochastic process defines an equilibrium. More than seven different misuses of the word “equilibrium” in the economics and finance literature are exposed in McCauley (2004).
In order to arrive at a completely different invariance principle, consider next a distribution of markets for a single asset, like gold or globalized autos (Ford, Toyota, GM, VW, or BMW, e.g.) on the face of the earth. The price density $g(p,X,t)$ depends not just on price $p$ and time $t$, but on location $X$ as well, and $g(p,X,t)$ is a conditional probability density for prices, a ‘Green function’ in the language of physics. The ‘no-arbitrage’ principle is equivalent to the assumption of translational and rotational invariance (McCauley, 2004) of the price density on the earth. The absence of arbitrage is a purely geometric principle that guarantees nothing other than that the probability distribution $g(p,t)$ for the price of the object traded is independent of position $X$. In particular, ‘no-arbitrage’ has nothing to do whatsoever with ‘market equilibrium’. Market equilibrium would be equivalent to time translational invariance of the price distribution: in equilibrium or in a driven steady state, $g(p,t)$ would also be independent of $t$, would define a statistical equilibrium with price density $g(p)$.

Falsifiability of a model via empirical data is a scientific necessity. The idea of falsifiability is not a new idea. Karl Popper only put into words what ‘hard science’ since Galileo has practiced. In physics, a new model will not be accepted unless it makes falsifiable new predictions. As an example of its predictive power, Newtonian mechanics was used to predict the existence of an ‘extra’ unobserved planet before Neptune was discovered. The SU(3) model in field theory was used to predict the $\Omega$ particle before it was observed. The neo-classical model
was perhaps once an example of science: it made definite predictions that have been falsified (Osborne, 1973; McCauley, 2004). So why is it still taught, since it cannot be used to predict or even explain any observable phenomenon correctly?

5. The Invisible Hand is a Falsifiable Proposition

Adam Smith’s Invisible Hand is the idea that supply in a free market should tend to rise to meet demand. Neo-classical economics refined the idea of the Invisible Hand to mean that price changes occur at or near equilibrium, that prices should tend to equilibrate so that market stability is implicitly assumed. Stable markets could exhibit only small fluctuations about statistical equilibrium, or near a steady state. The neo-classical assumption of stable equilibrium is falsifiable. Price changes near equilibrium, under the influence of noise traders, could be described mathematically by a stationary process in stochastic dynamics, one where the Gibbs entropy of the market

\[
S(t) = - \int g(p,t) \ln g(p,t) dp
\]

would necessarily become asymptotically constant as \( t \) increases, achieving an entropy maximum. Both the average return and variance/volatility of a stationary process are constants. Here, \( g(p,t) \) must be understood the correct
empirically deduced price density. Financial markets are typically very liquid and in that limit can be approximately described by a stochastic differential equation

\[
\frac{dp}{dt} = \epsilon(p) = rp + p\sqrt{d(p,t)} \frac{dB}{dt}
\]

(5)

where, as we will show below, the p-dependence of the price diffusion coefficient \( p^2 d(p,t) \) must be extracted from the observed time-dependence of the empirical price density \( g(p,t) \). Here, the excess demand \( \epsilon(p,t) = D(p,t) - S(p,t) \) is described as drift plus noise, in agreement with the fact that price changes are not deterministic even on the shortest time scales. But let us ignore the empirical data for the moment and ask first what would be the practical implications of the economists’ assumption of market equilibrium.

Stationarity would demand an asymptotically time invariant price density \( g(p) \). This defines statistical equilibrium. In this case, both the mean \( <rp> \) and standard deviation \( \sigma = \Delta p^2 = <p^2> - <p>^2 = <p^2 d(p,t)> \) would be constants (\( <rp> = 0 \) is necessary if \( <e> = 0 \)). Equilibrium markets would therefore be both stationary and nonvolatile.

Why should anyone care about equilibrium? If we could locate equilibrium in a real market, then we could define ‘value’ unambiguously. ‘Value’ would simply be the equilibrium price
In statistical equilibrium we could take the equilibrium price \( p^* \) to be either the average or most probable price, with fluctuations about equilibrium described by \( g(p) \). This would permit the construction of a trading strategy: buy the stock if \( p < p^* \) and sell it if \( p > p^* \). One could refine this to argue that one should trade outside the range \( \Delta p^* = p^* \pm \sqrt{\sigma} \). Stationary stochastic dynamics is ergodic, whereas nonstationary dynamics is not. The return to equilibrium demanded by the assumption of stationarity guarantees that such buying and selling are possible, and to be more precise one could calculate the distribution of first passage times.

However, if we study the returns variable \( x = \ln p(t)/p_o \) where the returns density is given by \( f(x,t) = g(p,t)dp/dx = pg(p,t) \) and \( p_o \) is some initial or other reference price, then the observed returns variance is given by \( \sigma^2 = \langle \Delta x^2 \rangle = \langle (x - \langle x \rangle)^2 \rangle = \Delta t^{2H} \) with \( H = O(1/2) \). Thererfore, financial markets are nonstationary. Another way to say it is that financial markets are unstable, they never approach statistical equilibrium. In spite of this simple fact, some economists continue to write papers about ‘stationary financial markets’. Physically, outstanding limit orders prevent financial markets from clearing: empirically seen, there is no daily clearing price in a financial market.

The lack of equilibrium in market data means that ‘value’ does not exist as an unambiguous idea, only price exists uniquely (to within arbitrage). Because neither dynamical nor statistical
equilibrium can be found in real market data, assertions that an asset is either undervalued or overvalued are subjective. But wishful thinking acted on collectively (self-fulfilling expectations widely-held) can lead to big price swings, as in the phenomenon of ‘momentum investing’ and the corresponding U.S. stock market bubble of from 1994-2001. This psychological condition, the inability to know ‘value’, combined with the easy availability of money as credit (and especially via leveraging) surely contribute to both nonstationarity and volatility. One can imagine noise traders changing their minds frequently, and so trading frequently, because they’re very uncertain of the ‘value’ of a financial holding like a stock, currency, or bond. This proposition could be simulated via an agent based trading model. An interesting exercise would be to introduce a trading model where equilibrium ‘exists’ mathematically in the model but is in some sense noncomputable (it could simply be NP-complete, not necessarily Turing (1936) noncomputable), and see what would be the effect on the market. The liquidity bath term $\sqrt{d(p,t)}dB(t)/dt$ in (5), which does not generate a lognormal process in $p$ when $d(p,t)$ depends on $p$, approximates the effect of the ‘noise traders’. Successful traders like Warren Buffet have zero weight in (5), they do not provide the daily liquidity that allows us to trade frequently, even on a time scale of a second, with small bid/ask spreads.
In the language of statistical physics, equation (5) with \( d(p,t) \) chosen correctly to reflect the market data provides us with something that may be roughly analogous to a mean field approximation to a complex system of interacting agents. Real agents have PC’s or Macs, high computational capability, but generally can’t do any worthwhile calculations when trading because they can’t distinguish knowledge from noise, and can only make guesses about future prices in the absence of ‘value’. Long Term Capital Management (LTCM) nearly brought down the world financial system (Dunbar, 2000) by assuming that 1. they could determine value, 2. by taking seriously the Modigliani-Miller “theorem” that the debt/equity ratio doesn’t matter, and 3. by combining these two assumptions with Black’s assumption that there is an equilibrium in the market, that ‘price always tends to return to value’. But what is ‘volatility’?

The first approximately quantitatively correct description of stock market returns was proposed in 1958 by the physicist turned finance theorist M.F.M. Osborne (Cootner, 1964), who plotted rough price histograms based on Wall St. Journal data in order to try to deduce the empirical distribution of stock prices. He inferred that stock returns seem to do a random walk, so that prices are distributed lognormally. The lognormal price distribution is generated by the stochastic differential equation (5) with variable local price volatility \( p^2 d(p) = (\sigma_p p)^2 \), where \( \sigma_p \) is constant. The corresponding returns distribution is
Gaussian because the stochastic differential equations for returns $x$ is given (via Ito calculus) by

$$dx = (r - \sigma_p^2 / 2) + \sigma_p dB(t)$$

(6)

Because Osborne's stochastic model is Markovian, the Hurst exponent $H$ in the variance or 'average volatility' $\sigma^2 = \langle(x - <x>)^2\rangle = \Delta t^H$ is $H=1/2$. We know from empirical data analysis that $H=O(1/2)$ (Mantegna and Stanley, 2000), but whether $H=0.4$, $0.5$, or $0.6$ is hard to decide empirically. The choice $H=1/2$ yields models obeying the 'efficient market hypothesis', which means simply that the market is very hard to beat: for $H=1/2$ there are no long time correlations in the market. There is also evidence from stock indices for $H \neq 1/2$ (Skjeltorp, 1996). A Hurst exponent $H \neq 1/2$ implies fractional Brownian motion and yields long-time correlations that could, in principle, be exploited for profit.

The Black-Scholes (1973) model of option pricing assumes Osborne’s Gaussian returns model. The Black-Scholes model is based on only two empirically measurable parameters, $\sigma_p$ and $r$, and is falsifiable. In fact, the model has been falsified on several grounds. One of them is that when the model is force fitted to option prices, the constant $\sigma_p$ must be varied as if it would depend on the strike price $K$. This is so-called ‘implied volatility’, and indicates that, in order to understand what the
market is telling us, we should start with the more general stochastic differential equation
\[ dx = (r - D(x,t)/2)dt + \sqrt{D(x,t)}dB \]

(7)
corresponding (via Ito calculus) to (5), where the diffusion coefficient, or ‘local volatility’ D depends on x, and the returns diffusion coefficient is D(x,t)=d(p,t). We’ll show in part 6 below how an (x,t) dependent diffusion coefficient can be deduced from the empirical density of returns f(x,t). We know three important empirical facts about financial markets: they’re nonstationary/unstable, and they’re volatile. Also, f(x,t) exhibits ‘fat tails’, but we’ll discuss the asymptotic behavior of f later.

For any market or economy, the notion of the Invisible Hand is a falsifiable proposition: one need only test a set of price or returns data or other time series for a given market or economy for asymptotic (strong) stationarity, or at least for weak stationarity in the form of lack of growth and lack of volatility (McCauley, 2004). We’ll explain why the worst problem that one faces is that typical nonfinancial markets provide us with such sparse data that reliable testing may be difficult or even impossible (the data are too easy to fit by many different dynamics models). Because of nonuniqueness in extracting models from data, e.g., we expect that GNP and business cycle data should be relatively easy to fit by using nonstationary, volatile models. To date, there is no convincing evidence from empirical data that any known market is
asymptotically stationary, and market volatility is rather common. Instead of approaching equilibrium, we expect that empirical returns distributions for nonfinancial markets will broaden without limit as time increases.

But how can the empirically observed time series $x(t)$ of a particular market be used to infer the underlying dynamics? This question is of central importance for economics, because economic dynamics have not yet been deduced empirically except for financial markets. Instead of trying to argue that the falsified ratex model is ideal, but the data are ‘hard to describe’ (no physicist will give any weight to such an argument), we must ask what the unmassaged market data can teach us. I emphasize in advance that our approach to data analysis is not at all the method of the econometrician: instead of having limited, preconceived models in mind (Granger, 1999), we deduce a stochastic model from the data (see McAdam and Hallet (2000) for an example of an attempt to force preconceived notions on the data). I will outline our program next, where I will argue that real market data are not at all hard to fit accurately by using dynamical models. To the contrary, market data are too easy to fit: lack of uniqueness in empirically based modeling is the real problem that we face.

Returning to Wigner’s theme, given the absence of enough symmetry principles to pin down inviolate dynamical law in finance and economics, what can we do? As Osborne has shown us, the answer is the same as if there would be enough invariance
principles to pin down real mathematical laws: we can study the available data for a specific market and try to extract a dynamical model that reproduces that data. In this case, we know in advance that we’re modeling data for a particular market in a particular era, and that any model is expected to fail at some unknown time in the future. Therefore, it’s essential that the model has few enough empirically known parameters to be falsifiable, otherwise one cannot know when the market has shifted in a complex/fundamental way.

Many economists are averse to studying finance, but financial markets differ from other markets mainly in that many trades are made very frequently, even on a time scale of a second, so that very good data are available for the falsifiability of few-parameter models. For houses or cars, the time scale for a large number of trades is much greater so that the data are much sparser. Such markets are far less liquid and may vary much more from one locale to another. Because of the abundance of adequate and reliable data, financial markets provide the best testing ground for both new and old ideas. Financial markets exhibit the interesting characteristics of economic systems in general: growth and, the business cycle (see Goodwin (1993) for a discussion of these phenomena). When we speak of the business cycle, a topic where both stochastic (Cootner, 1964) and nonequilibrium nonlinear deterministic models were considered rather early (Velupillai, 1998), we no longer expect to discover any either stable or unstable periodicity. We rather expect to understand ‘the business cycle’ as volatility combined
with nonstationarity of the market distribution, where the market distribution is simply the collection of histograms obtained from real market data. Stationarity is another name for time invariance. Nonstationarity means that market entropy (4) increases without limit, that the market is far from any equilibrium. Equilibrium and stability do not exist as possibilities for financial markets: evidence for vanishing excess demand cannot be found in the empirical data.

6. An Empirically Based Model of Financial Markets

In a stochastic description of markets, the excess demand \( \varepsilon(p,t) \) is modeled as drift plus noise

\[
dp = prdt + p\sqrt{d(p,t)}dB(t)
\]

(8)

where \( dB \) is a Wiener process and \( p^2d(p,t) \) is the price diffusion coefficient. The stochastic differential equation for the returns variable \( x = \ln p(t)/p_o \) is given by

\[
dx = (r - D(x,t)/2)dt + \sqrt{D(x,t)}dB
\]

(9)

where the returns diffusion coefficient transforms like a scalar, \( D(x,t)=d(p,t) \). We can regard the returns diffusion coefficient \( D(x,t) \) as the ‘local volatility’ (McCauley, 2004). We will use as our independent variable the logarithmic return \( x \), not price \( p \),
in modeling because empirical analyses must be carried out using returns in order to avoid errors when \( x \) is large in magnitude (Osborne, 1958; Gunaratne, 1990; Dacorogna, 2001).

In contrast with the usual desire of economists to divide the economy into ‘system’ and ‘shocks’, the noise/shock is the main part of the stochastic dynamical system (8), otherwise, excess demand is neither correctly defined nor described. The noise dominates the dynamics: financial markets are mainly noise. The noise term in (9) is \( \sqrt{D(x,t)}dB(t) \), and this is where the interesting market dynamics lie. The Green function of the Fokker-Planck equation corresponding to (9) is the market Green function: it can be used to calculate all market predictions, including option pricing (McCauley, 2004).

To a first approximation, financial data for small to moderate returns \( x \) are neither approximately Gaussian nor Levy but are instead more approximately exponentially distributed (fig. 1)

\[
f(x,t) = \frac{1}{\sqrt{d+\Delta t}} \frac{\sqrt{d_+d_-}}{\sqrt{d_+} + \sqrt{d_-}} e^{-t+|\sqrt{d_-}}
\]

(10)

where the plus-minus subscripts refer to the regions to the right and left of the peak of the returns density \( f \), \( x>\delta \) and \( x<\delta \). The exponential distribution is generated by a Markovian
model with nontrivial local volatility (diffusion coefficient $D(x, t)$)

$$D(x, t) = d_{\pm} (1 + |x - \delta|/(d_{\pm} \Delta t)^{1/2})$$

(11)

where $d_{+}$ and $d_{-}$ are constants, and $\delta$ depends on $\Delta t$ and defines the peak of the returns density. When ‘Galilean invariance’ holds, then $\delta = \tau\Delta t$. This local volatility yields a Brownian–like average (or global) volatility $\sigma^2 = \Delta t$ at long times. The average volatility, or mean square fluctuation in return $x$, is given by

$$\sigma^2 = \langle \Delta x^2 \rangle = \int (D(x(s), s)) ds = \int \int D(z, s) g(z, s; x, t) dz ds$$

(12)

where $g(x, t; x', t')$ is the market Green function (eqn. (10) defines the Green function for one particular initial condition). The exponential model prices options correctly without the need for fudge-factors like ‘implied volatility’ that characterize financial engineering based on trying to force-fit a Gaussian returns model to the data (McCausley and Gunaratne, 2003). All of the constants in the model are fixed by empirical data, so the model is falsifiable.
For the benefit of readers who are economists, here’s specifically what we did (McCauley, 2004). Gemunu Gunaratne first deduced the exponential distribution from financial market histograms. I used the known average volatility, $\sigma^2=\Delta t$, for $\Delta t$ larger than about 10 min. of trading, to deduce the time dependence of the exponential distribution. From there, we asked which diffusion coefficient $D(x,t)$ in the stochastic differential equation (9) generates the exponential density with the observed time dependence. This is not the same as force-fitting a preconceived stochastic model to the data. A strong test of our dynamical model would be to measure the local volatility $D(x,t)$ directly. The main points are very simple but are easily misunderstood, because the method of deduction is not the usual method in econometrics.

In particular, in a stochastic model (9), $dB(t)$ is a Wiener process but the stochastic integral of $\sqrt{D(x,t)}dB(t)$, which appears in the solution $x(t)$, is not globally a Wiener process if the diffusion coefficient $D(x,t)$ depends on $x$ (models where $D$ depends on $t$ alone, and not on $x$, are trivially equivalent to Wiener processes by a time transformation). This is the main point: the form of the diffusion coefficient $D(x,t)$ that defines the noise term $\sqrt{D(x,t)}dB(t)$ in dynamics must be deduced empirically. The usual alternative is instead to assume a stochastic model based on a postulated, preconceived form of noise, and then try to force-fit the data by a ‘best choice of parameters’. Our program is to respect the data and therefore first to discover the form of the
empirical distribution. Then, we determined the time dependence of the distribution’s parameters from the data, and used that information to deduce a dynamical model: plugging the empirical distribution into a Fokker-Planck equation (corresponding to (9)) allows one to solve the ‘inverse problem’ to find the diffusion coefficient that generates the observed distribution (McCauley and Gunaratne, 2003). Newton solved an inverse problem to deduce the inverse square law of gravity from Kepler’s elliptic orbits (McCauley, 1997).

In contrast, the usual method of the economist is to assume a stochastic model and then try to extract a best fit of parameter values for that model from the data. E.g., the Real Business Cycle (RBC) model (Chow and Kwan, 1998) assumes a particular form for the noise term. In contrast with RBC, we deduce the form of the noise term from the observed time dependence of the empirical distribution. This is physically significant: the noise term reflects what the ‘noise traders’ are doing. The noise term that would describe a stochastic model of the GNP would reflect the nature of the noise in the economy, likewise for a sector in a business cycle model.

The exponential distribution has also been discovered in empirical studies of the growth rate of firms (Stanley et al, 1996: Bottazi et al, 2001). Those papers also start with empirical histograms and then deduce a probability distribution. The exponential distribution has fat tails in price, but not in returns.
The empirical financial distributions have fat tails in returns for large returns $x$.

We’ve discussed volatility in part 4, but the most commonly heard criticism of the Gaussian returns model is that the empirical density of financial returns has fat tails $f(x,t) = x^\mu$ (fig. 2) for large returns $x$ (Dacorogna et al, 2001), where $\mu$ is a nonuniversal scaling exponent in the range from about 3.5 to 7, it may vary from market to market. Fat tails in historic cotton prices were first discovered in Osborne’s era by Mandelbrot (Cootner, 1964), following Pareto, but Mandelbrot then assumed an infinite variance, to zeroth order, in order to try to apply the Levy distributions. Levy distributions generate the smallest tail exponents, $1 < \mu < 3$, and therefore the fattest tails.

Levy distributions can be used to generate fat tails, but with entirely different underlying dynamics than in our Markovian model. In the formula

$$L_\alpha(x,\Delta t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dke^{ikx} R^{\Delta t/2}$$

(13)

for a probability density, symmetric Levy distributions are given by $\alpha < 2$ and have infinite mean square fluctuation
The exponent $\alpha$ describes fat tails for large $x$. The tail exponent is $\mu = 1 + \alpha$, but $\alpha < 2$ is too small to generate financial data. The dynamics of the dynamics of Levy distributions especially for $\alpha < 2$ is discussed in Hughes et al (1981). For $\alpha > 2$ there is recurrence in the form of long-time anticorrelations, whereas for $0 < \alpha < 2$ there is a hierarchy of clustering in the time series $x(t)$. In the physics literature, it seems not to have been understood that the case where $\alpha > 2$ cannot be described by a ‘Langevin equation’ (stochastic differential equation) in the variables $x$ or $p$. When $\alpha > 2$ then one gets fat tails with fractional Brownian motion, where the average volatility is $< \Delta x^2 > = \Delta t^{2/\alpha}$. The Hurst exponent is $H = 1 / \alpha < 1 / 2$, so there are infinitely long-time correlations, there is no diffusion coefficient $D(x,t)$, and therefore no description of the dynamics via a stochastic differential equation or Fokker-Planck equation in either $p$ or $x$. Because of the long time correlations the efficient market hypothesis (EMH) is violated, although the violation will not likely help a small trader to beat the market is $H$ close enough to $1 / 2$. To have correlations strong enough to beat the market effectively, one needs an exponent $H$ considerably different than $1 / 2$.

The efficient market hypothesis (EMH) may obeyed by a good model to zeroth order: the EMH simply reflects the fact that the market is very hard, but not necessarily impossible, to beat. To zeroth order, there are no systematic patterns (correlations) in the market.
Here’s something entirely new: we can also generate fat tails $f(x,t) = x^{-\mu}$ for large $x$, for all possible exponents $\mu \geq 2$ (fig. 3), via a stochastic differential equation (9) where the tail exponent $\mu$ is uniquely determined by the nonlinearity parameter $\varepsilon$ in the returns diffusion coefficient (Alejandro-Quinones et al, 2004)

$$D(x, t) = d_+ (1 + \frac{|x - \delta|}{\sqrt{d_+ \Delta t}}) + \varepsilon \left( \frac{|x - \delta|}{\sqrt{d_+ \Delta t}} \right)^2$$

(14)

This is a surprising result. Many papers and some books have been written on nonstationary volatile stochastic processes, but few examples have been given that combine nonstationarity, volatility, and fat tails. We’ve combined all three.

Next, I will emphasize a point that’s central for extracting dynamical models from empirical data. There is nonuniqueness in the choice of time dependence of time dependence in equations (10) and (11) chosen to fit finance market data (fig. 1). Given the known nonuniqueness faced in extracting chaotic dynamics from data, this is not a surprise. One attempts to extract an infinite precision model from finite precision data. Newton didn’t face this problem because of the underlying space-time symmetry principles, but if you would try to extract Newton’s second law from a chaotic system like the three body
problem, then you’d have the same difficulty. In applying the new model defined by eqn. (10) to option pricing, we found (McCauley and Gunaratne, 2003) that we have the unwarranted luck that the nonuniqueness doesn’t matter on time scales much less than a hundred years. Normally, one should not expect such luck in modeling. Finite precision in empirical data always implies nonuniqueness in the inference of an infinite precision model. Without the underlying space-time symmetry principles used to pin down laws of motion in physics, the nonuniqueness can be severe, but the nonuniqueness involved in the nonempirical postulation of models is far, far worse. One cannot capture the essence of market behavior merely be imagining how agents might behave (as in ratex), one must instead ask the market directly.

The main aim of economic theory in our era should be to match the success of the empirical description of financial markets for at least one nonfinancial market. Toward that end, ideas of stability and equilibrium in economics should either be verified empirically or else completely abandoned as guiding theoretical principles. In particular, economics texts should stop teaching ‘rational expectations’ as if that model would bear any realistic relation to real markets. To continue to teach a completely falsified model is to mislead generations of students. Again, Newton’s first law and the law of gravity can be verified to high decimal precision in experiments on earth and on the moon, but no market has yet been found that even approximately reflects ratex.
Note that financial markets have been accurately described by very simple stochastic dynamics, so where’s the complexity? Complexity leads us into questions of computational limitations or intractability. The highest degree of computational complexity is that of a Turing machine (Feynman, 1996; Velupillai, 2000). We expect that markets are not merely stochastic (“random”) but are also complex. Can the empirically observed time series of a complex system be used to infer the underlying dynamics? We know now that Newton would have had serious problems were it necessary to discover the basic laws of physics by analyzing time series for a chaotic system like the 3-body problem, but complexity turns out to be much worse that deterministic chaos. In what follows, I assume that all functions that we use to define a deterministic dynamical model are Turing computable, and that computable numbers are used as control parameters and initial conditions in the model. By this restriction we avoid the trivial noncomputability of the measure one set of numbers that can be defined to ‘exist’ in the continuum, but cannot be generated algorithmically.

7. Complexity in Physics, Biology, and Markets

To date, we have no physically or biologically motivated definitions of complexity that are mathematically adequate, in spite of the fact that cell biology provides us with numerous examples of natural complexity. Our everyday computers are an example of complexity and can be described dynamically as
Newtonian electro-mechanical machines. Contrary to expectations in some quarters, scaling is not an example of complexity, nor is stochastic dynamics (‘randomness’). Moore has discussed the nature of maximal complexity in deterministic dynamics.

We can generate maximal computational complexity by using simple deterministic dynamics (Moore, 1990, 1991). Low dimensional iterated maps that are equivalent to Turing machines provide examples. These dynamical systems have no attractors, no symbolic dynamics/no generating partition, and so exhibit no scaling laws that would inform us of behavior at smaller length scales in terms of observed dynamics at larger length scales. Instead, ‘surprises’, new unforeseen behavior, are possible at all length scales. By length scales, I think here of the hierarchy of coarsegrainings defined by the generating partition in a chaotic system (McCauley, 1993), where one looks in finer and finer detail at the dynamics, increasing the magnification of the microscope, so to speak. Without symbolic dynamics and the corresponding generating partition, we have no way to deduce a Turing-equivalent dynamical system from time series. This is a serious drawback in anyone’s book.

Mutations of viruses and bacteria to new, unexpected forms provide an example of the surprises characteristic of complexity. Such surprises now occur on very short time
scales, time scales shorter than the time required to discover new antibiotics, e.g.

In continuous time dynamics, at the shortest time scales there is no way to distinguish complexity from simplicity in deterministic dynamics. This assertion can be extended analytically to slightly larger time scales. Every deterministic dynamical system \( dp/dt = \varepsilon(p) \), even a chaotic or complex one, has a unique, well-defined solution (is globally solvable) so long as the velocity field \( \varepsilon(p) \) satisfies a Lipshitz condition with respect to the \( n \) variables \( p_i \). If, in addition, the velocity field \( \varepsilon(p) \) is analytic in those variables then the power series locally defining the time evolution operator \( U(t) = e^{Lt} \),

\[
p_i(t) = p_i^0 + t(Lp_i)_0 + \frac{1}{2} t^2 (L^2 p_i)_0 + ... \tag{15}
\]

has a nonvanishing radius of convergence, so that the solution of the dynamical system can in principle be defined by power series combined with analytic continuation for all finite times (Poincaré, 1993). \( L \) is the infinitesimal generator of the flow and is determined by \( \varepsilon(p) \). The radius of convergence of (15) is typically small and unknown. Unless one can determine the singularities of (15) in the complex time plane, one does not know when and where to continue analytically. Therefore, in practice, we cannot expect to solve nonintegrable dynamical systems more than locally, for only very short time intervals. This is a restriction on predictability that precedes any
computability limitations that may arise in deterministic dynamics. This limit on predictability is ignored by economists who claim that they can make reliable global predictions.

In deterministic iterated maps, the surprises arise *internally* from the system’s dynamics. In order to imagine more clearly how surprises could appear in a finance market in the short run, we can consider the market modeled by fluctuating asset price (described to zeroth order by (9)) and the liquidity bath, which Brownian motion theory assumes to remain unchanged. The diffusion term in (9) assumes implicitly that the liquidity bath is there, that you can make small trades without affecting the market, to zeroth order. The analogy of the liquidity bath with the heat bath for a Brownian particle is described in McCauley (2004). In a financial market, the occurrence of a surprise may cause the liquidity bath to dry up suddenly (market crash). In that case, (8) and (9) do not apply: a liquidity drought is not a Wiener, lognormal, Levy, exponential, or any other continuous time stochastic process. *It is more approximately the complete absence of the noise traders* (meaning that $D(x,t)\approx 0$). In order to try to include surprises mathematically, one could try to model the interacting system of agents trying to set prices in the absence of ‘value’, avoiding assuming the liquidity bath/Brownian motion approximation explicitly and then try to derive (9) from the model under a liquidity bath approximation.
Summarizing, for a deterministic dynamical system with universal computational capability, nothing can be said in advance about the future, either statistically or otherwise: the future is computationally undecideable. This maximum degree of computational complexity occurs in low dimensional nonintegrable conservative Newtonian dynamics. In particular, billiard ball dynamics exhibit positive Liapunov exponents and provide us with an example of a chaotic system that is mixing (Cvitanovic et al’, 2003). But billiard balls can also be used to compute reversibly and universally (Fredkin and Toffoli, 1982). Such a method of computation would be impractical because the positive Liapunov exponents magnify errors in initial conditions of the billiard balls, messing up the computation.

Molecular biology is largely about complexity at the molecular (DNA-protein) level. E.g., the thick, impressive, and heavy text by Alberts et al (2002) is an encyclopedia of cell biology, but displays no equations. Again, with no equations as an aid, Weinberg (1999) describes the 5-6 independent mutations required to produce a metastasizing tumor. All these impressive biological phenomena remind us more of the results of a complicated computer program than of a dynamical system, and have all been discovered reductively by standard isolation of cause and effect in controlled, repeatable experiments.
Many economists and econophysicists would like to use a biological analogy in economics, but the stumbling block is the complete absence of a falsifiable dynamical description of biological evolution. Instead of simple equations, we have simple objects (genes) that behave more like symbols in a complicated computer program. Complex adaptable mathematical models notwithstanding, there exists no mathematical description of evolution that is empirically correct at the macroscopic or microscopic level. Schrödinger (1944), following the track initiated by Mendel that eventually led to the identification of the molecular structure of DNA and the genetic code, explained quite clearly why evolution can only be understood mutation by mutation at the molecular level of genes. Mendelism provides us with a clear picture of Darwinism at the cellular level. The only precise definition of biological evolution relies on mutations, there is no falsifiable model of Darwinism at the macroscopic level. That is, we can understand how DNA mutates to a new form but we do not have a model showing falsifiably how a fish evolves into a bird. That’s not to say that it didn’t happen, only that we don’t have, and probably never will have, a model that helps us to picture how it happened.

The terms ‘emergence’ and ‘self-organization’ are not precisely defined. They mean different things to different people. It’s not

---

1 Mendel was trained in the Galilean method: he studied and taught physics in Vienna. He did not get an academic position, and so retreated to Brn in what is now Slovakia, and studied peas. The idea of a ‘code script’ in chromosomes was suggested by Schrödinger (1944).
clear what writers could have in mind, other than symmetry-breaking and pattern formation at a bifurcation in nonlinear dynamics, when they claim that a system ‘self-organizes’\(^2\). Some researchers who study complex models mathematically expect to discover new, ‘emergent’ dynamics for complex systems, but so far no one has produced an empirically relevant or even theoretically clear example. See Lee (2004) for a survey of some of the usual ideas of self-organization and emergence. Crutchfield and Young (1990), Crutchfield\(^3\) (1994) and others have partly developed the interesting idea of nontrivial computational capability appearing spontaneously within a dynamical system due to bifurcations. This doesn’t present us with new dynamics, it’s about an increase in complexity in already existing dynamics due to a bifurcation. Crutchfield assumes a generating partition and symbolic dynamics, but Moore has shown that we have to give up those ideas for dynamics with Turing-equivalent complexity. Another weakness in Crutchfield’s program is his restriction of noise to stationary processes. That won’t work for market data, or for realistic market models either. Can the program be extended and then applied to teach us anything new or useful about economic or biologic data?

\(^2\)Hermann Haken (1983), at the Landau-Ginzburg level of nonequilibrium statistical physics, provided examples of bifurcations to pattern formation via symmetry breaking. All subsequent writers have used ‘self-organized’ as if the term would be self-explanatory, even when there is no apparent symmetry breaking. Is a deterministic or noisy stable equilibrium point or limit cycle (or other invariant set without escape) an example of self-organization? If so, then maybe we don’t need the phrase at all.

\(^3\)My Galilean approach is completely contrary to the postmodernist philosophical outlook expressed, especially in part I, of Crutchfield’s 1994 paper.
I now offer an observation to try to clarify ‘emergence’: whatever length and time scales one studies, one first needs to discover approximately invariant objects before one can hope to discover new dynamics. The ‘emergent dynamics’, if such dynamics can be discovered, will be the dynamics of those objects. Now, what many complexity theorists hope and expect is that new dynamical laws beyond physics will somehow emerge statistically-observationally, or can be postulated, at larger than molecular length and time scales, laws that cannot be derived systematically from phenomena at smaller length scales. A good example is that many Darwinists would like to be able to ignore physics and chemistry altogether and try to understand biological evolution macroscopically, independently of the mass of details of genetics, which have emerged from controlled experiments and data analysis.

Consider specifically cell biology, where the emergent invariant objects are genes. Genes constitute a four-letter alphabet used to make three letter words. From the perspective of quantum physics, genes and the genetic code are a clear example of emergent phenomena. With the genetic code, we arrive at the basis for computational complexity in biology. Both DNA and RNA are known to have nontrivial computational capability (Adelman, 1994; Bennett, 1982; Lipton, 1989). One can think of the genes as ‘emergent’ objects on long, helical molecules, DNA and RNA. But just because genes and the code of life have emerged on an approximately
one-dimensional tape, we do not yet know any corresponding new dynamical equations that describe genetics, cell biology, or cancer. So far, one can only use quantum or classical mechanics, or chemical kinetics, in various different approximations to try to calculate some aspects of cell biology.

My main conclusion is that ‘emergence’ does not guarantee the appearance of new laws of motion. Apparently, invariant objects can emerge without the existence of any new laws of motion to describe those objects. Genes obey simple rules and form four letter words but that, taken alone, doesn’t tell us much about the consequences of genetics, which reflect the most important possible example in nature of computational complexity: the evolution from molecules to cells and human life.

Finally, dreams of holism are pure illusion. Every mathematical model that can be written down represents some kind of attempt at reductionism. The only question is: does the attempt succeed or fail? Here are some examples. The renormalization group method in statistical physics, valid at order-disorder transitions, reduces phenomena at a critical point approximately to symmetry and dimension. Quantum theory, the law of nature at very small length scales explains chemistry via electrons, protons, atoms and molecules. Cell biology successfully reduces observed phenomena to very large, complicated molecules, to genes, DNA, proteins, and cells. Proponents of self-organized criticality try to reduce the
important features of nature to the equivalent of sand grains and sand piles via the hope for an underlying universality principle (Bak, 1996). Network enthusiasts likewise hope to reduce many interesting phenomena to nodes and links (Barabasi, 2002). The worst weakness in the latter two programs is that there are no known universality principles for driven-dissipative systems far from thermal equilibrium, except at the transition to chaos.

I end by suggesting an biological analogy for economics. The creation of new markets depends on new inventions and their exploitation for profit. Mathematical invention has been described psychologically by Hadamard (1945). Conventional ideas of psychology completely fail to describe the solitary mental act of invention, whether in mathematical discovery or as in the invention of the steam engine or the sequential computer. Every breakthrough that leads to a new invention is an example of a 'surprise', of something emerging from within the system (the system includes human brains and human actions) that was not foreseen. A completely new product, like the gasoline engine or the PC, is based on an invention. The creation of a successful new market, based on a new product, is partly analogous to an epidemic: the disease spreads seemingly uncontrollably at first, and then eventually meets limited or negative growth. The simplest mathematical model of creation that I can think of would be described by the growth of a 'tree', where new branches (inventions or breakthroughs) appear
suddenly without warning. This is not like a search tree in a known computer program. Growth of any kind is a form of instability, and mathematical trees reflecting instability do appear in nature, in the turbulent eddy cascade e.g., but in that case the element of ‘surprise’ is missing.

Summarizing, I’ve discussed the use of the Galilean method in finance and have suggested that it be applied in economics. Empirically motivated models are necessary beforehand if mathematics is to be made effective in general economics, as it has become in finance theory. Worries about complexity are premature before adequate empirical market models have been deduced. Market time series and histograms are, of course, of limited value in predicting the future: they reflect in some coarsegrained fashion how we’ve been behaving economically. The future in socio-economic phenomena is to some unknown degree undecidable and can’t be known in advance, not even statistically. Using market statistics as a basis for prediction assumes that tomorrow will be statistically like yesterday. If we’ve modeled carefully, as in finance, then this assumption may not get us into hot water so long as there are no surprises. Insurance companies make money by assuming that the future will be like the past statistically (they take in to account fat tails but hope for stationarity), and lose money when it isn’t.

Of course, one can also make nonempirically based mathematical or even nonmathematical models, and assert that
if we assume this and that, then we expect that thus and such will happen. That sort of modeling activity is not necessarily completely vacuous, because socio-economic beliefs can be made into reality by acting strongly enough on wishes or expectations, there are self-fulfilling prophecies that go beyond the realm of science: e.g., a model can be enforced or legislated. Both communism (implemented via bloody dictatorships) and globalization (implemented via massive deregulation and privatization, big financial transfers, and supragovernmental\textsuperscript{4} edict) provide examples. Neo-classical economics/’rational expectations’ is a mathematized ideology that encourages unlimited deregulation. The construction of competing models based on real market statistics will be useful for confronting the ‘best of all possible worlds’ claims of the ideologues and other true believers with the reality of continually evolving markets and economies. Instability and surprises are the main aspects of market reality in our era.

Acknowledgement

I thank Massimo Salzano for inviting me to give this paper at the stimulating and useful Windows 2004 meeting in Salerno, and for making it possible for me to attend. I’m grateful to Vela Velupillai for both strong encouragement and very useful

\textsuperscript{4} Examples of powerful supragovernmental organizations are the IMF, the World Bank, the World trade Organization, and the European Union. One might try to argue roughly that the U.S. Federal Reserve Bank has somewhat comparable influence.
criticism. I’m also very grateful to my always helpful, because very critical and clear thinking, home editor, Cornelia Küffner, for reading and suggesting improvements in the earliest and latest versions of the manuscript. Her suggestions were built into the text. Finally, I’m grateful to my good friend Yi-Cheng Zhang for sharing his strategy with me several years ago. That strategy was imperfectly implemented both in this paper and in my lecture.

References


Osborne, M.F.M., in Cootner (1964).


**Figure Captions**

1. The histogram for the distribution of relative price increments for US Bonds for a period of 600 days. The horizontal axis is the variable \( x = \ln(p(t+\Delta t)/p(t)) \), and the
vertical axis is the logarithm of the frequency of its occurrence ($\Delta t=4$ hours). The piecewise linearity of the plot implies that the distribution of returns $x$ is exponential.

2. Histogram of USD/DM hourly returns, and Gaussian returns (dashed line). Figure courtesy of Michel Dacorogna.

3. The exponential distribution $F(u)=f(x,t)$ developes fat tails in returns $x$ when a quadratic term $O((x-r\Delta t)/\Delta t^{1/2})^2$ is included in the diffusion coefficient $D(x,t)$. Here, $u=(x-r\Delta t)/\sqrt{\Delta t}$. 