NONLINEAR TIME SERIES, COMPLEXITY THEORY, AND FINANCE[1]

by

William A. Brock
Department of Economics
The University of Wisconsin

Pedro J. F. de Lima
Department of Economics
The Johns Hopkins University

First Draft, March, 1995, This Draft, September,1995

[1]A condensend version of this paper is forthcoming in G. Maddala, C. Rao, eds., *Handbook of Statistics Volume 14: Statistical Methods in Finance*, North Holland, New York. We would like to thank Michelle Barnes, Craig Hiemstra and Blake LeBaron for their helpful comments. The usual disclaimer applies.

Contents

NONLINEAR TIME SERIES, COMPLEXITY THEORY, AND FINANCE

William A. Brock
Department of Economics
The University of Wisconsin

Pedro J. F. de Lima
Department of Economics
The Johns Hopkins University

First Draft, March, 1995, This Draft, September,1995

1 Introduction

This article describes statistical aspects of a line of recent work in finance that is associated with the words "nonlinearity," "long term dependence," "fat tails,""chaos theory," and "complexity theory." We shall give a rather lengthy introduction in order to give the reader a road map through the issues taken up here. In the spirit of a road map, we shall indicate Section headings where each issue is discussed in detail or give references if the issue is not dealt with in this article. Before we begin let us give a brief overview of some recent "trendy" topics which shall play a role in this article.

Centers of research in complexity theory such as the Brussels School, the Stuttgart School, the Santa Fe Institute, and hosts of other related Centers and Institutes springing up around the world are turning to computer based methods as well as analytical methods to study phenomena that lie within the rubric of "complex systems."

Indeed highly publicized centers such as the Santa Fe Institute (SFI) place the computer and various types of "Adaptive Computation Methods" and "Artificial Life" at the center of their research strategies. In general the SFI methods blend together ideas from economics, evolutionary biology, computer science, interacting systems theory, and statistical mechanics.

A good statement of the SFI approach for economics and finance is in the SFI July, 1993 newsletter, edited by LeBaron. A good example of SFI style work in finance is the work on "artificial stock markets" by Arthur, Holland, LeBaron, Palmer, and Taylor (1993). In that work different species of trading strategies coevolve as they strive to maximize a measure of fitness, i.e., profits. The system is designed to run on a desktop computer and could be viewed as a form of "artificial economic life" in the SFI sense of that term. There are no analytical results available for the Arthur et al. system.

The book edited by Friedman and Rust (1993) contains somewhat related work along analytical, experimental, and empirical lines. There is a section on the design and experience of the SFI evolutionary tournament where trading strategies competed against each other in a setting reminiscent of Axelrod's famous work on evolutionary tournaments for prisoner dilemma games.

Finance works such as Brock (1993), Friggit (1994), Vaga (1994) fall into the interacting systems category. Vaga (1994) builds on his earlier works which apply statistical mechanics to build a stock market model that can exhibit phase transitions. Friggit (1994) uses statistical mechanics type methods to propose and study a theory of evolutive dynamics for high frequency foreign exchange markets. Brock (1993) builds a theory based on a unification of discrete choice theoretic modelling from econometrics, received asset pricing theories, and statistical mechanics. More will be said about this kind of theory in Section four below.

The field of statistics itself has been moving in a related direction. Simulation based methods such as the Bootstrap (P. Hall (1994), Jeong and Maddala (1993)) and Dynamic Method of Simulated Moments (Duffie and Singleton (1993) and references to McFadden (1989) and Pakes and Pollard (1989)) are pushing analytical methods such as asymptotic expansions (first and higher order) off the center of the stage.

We shall devote part of this article to an argument for a style of research in statistical finance where models inspired by direct theoretical arguments are estimated by computer assisted methods such as MSM and where model adequacy (specification testing) is done by bootstrapping financially relevant quantities under the null. That is to say, the quantities that are inputted into the specification tests are themselves motivated by the type of economic and financial behavior one is trying to study. For example, distributions of statistics gleaned off of trading strategies are bootstrapped under the null model being tested in Brock, Lakonishok, LeBaron (1992), and Levich and Thomas (1993). This approach to specification testing is described in Section four below.

1.1 Theoretic and statistical models

The theme of argument above contains the subtheme that a closer analytical study between theoretic models and the statistical models that are estimated would be useful. This goes beyond the usual complaint that theory and econometrics need to be closer.

Here is a simple example. Most asset pricing models such as those treated in the book by Altug and Labadie (1994) generate an equilibrium asset pricing function of the form $p_t = p(y_t)$ where y_t is a low dimensional state vector for the system. ARCH-type models are intended to model the innovations $\epsilon_t \equiv p_t - E_{t-1}p_t$. Consider the broad class of ARCH models $\epsilon_t = \sigma_t Z_t$ where $\{Z_t\}$ is IID with mean zero and variance one with a symmetric about zero distribution (e.g. normal) and σ_t^2 (the conditional variance of ϵ_t) is a function of past ϵ's and σ's. Call these ARCH processes, "symmetric ARCH processes." We shall show that $\{\epsilon_t\}$ symmetric ARCH almost implies $p(.)$ is essentially linear, i.e.,

$$E_{t-1}p(y_t) = p(E_{t-1}y_t) \text{ for all past y's.}$$

This may imply unpleasant restrictions on the primitives of asset pricing models like those used in Lucas (1978), Brock (1982), and Duffie and Singleton (1993). For example, in the context of the models of Brock (1982) and Duffie and Singleton (1993), this is close to requiring that the utility function be logarithmic and the production function be Cobb Douglas with multiplicative shocks. One may not wish to impose such structure on the primitives of the model. In any event the implication that $p(y_t)$ is linear in the state variable y_t is potentially unpleasant. We state

Proposition 1 *Assume y_t is one dimensional, $p_t \equiv p(y_t), p(.)$ increasing in y. Furthermore assume $\eta_t \equiv y_t - E_{t-1}y_t$ and $\epsilon_t \equiv p_t - E_{t-1}p_t$, are conditionally (on past y's) symmetrically distributed with mean zero and finite variance with unique conditional medians of zero. Then $p(E_{t-1}y_t) = E_{t-1}p(y_t)$ for all past y's.*

Proof: By assumption,

$$\begin{aligned}
&\text{Prob}\{\epsilon_t = p(E_{t-1}y_t + \eta_t) - Ep(E_{t-1}y_t + \eta_t) \le 0\} = 1/2 \\
&\qquad\qquad = \text{Prob}\{\eta_t \le p^{-1}[E_{t-1}p(E_{t-1}y_t + \eta_t)] - E_{t-1}y_t\}.
\end{aligned}$$

Now, by assumption η_t is conditionally symmetrically distributed about zero, so the conditional median of η_t is zero. Hence,

$$p^{-1}[E_{t-1}p(E_{t-1}y_t + \eta_t)] - E_{t-1}y_t = 0.$$

Thus, $E_{t-1}p(E_{t-1}y_t + h_t) = p(E_{t-1}y_t)$. Q.E.D.

This type of proposition can be generalized to the first component of $p(y_t, y_{t-1}, \ldots, y_{t-L})$ by following the above argument for the first component. While ARCH models can easily accomodate non symmetrically distributed innovations, empirical applications of ARCH models commonly assume symmetry of the innovations.[2]

Furthermore, the survey of Bollerslev, Engle, and Nelson (1994) contains no work which studies the "inverse mapping" between the statistical structure assumed in the ARCH-type model being estimated and the underlying structure imposed upon the utilities, production functions, and market institutions of the underlying asset pricing model that would give inspiration or motivation for the ARCH-type model being estimated. We gave a sample above of what such research might look like. Turn now to possible ways of "omnibus testing" data sets for the appropriateness of estimation of statistical models within a particular parametric class such as symmetric ARCH.

Define $S_t \equiv \text{sgn}(\epsilon_t)$, sgn(x) equals the sign of x. The large class of ARCH-type processes generate an IID process for $\{S_t\}$, regardless of the shape of the innovations' distribution. This suggests use of tests of IID for $\{\text{sgn}(\epsilon_t)\}$ as a general test for ARCH effects. Furthermore, one could test the symmetry of the innovations' distribution by testing the hypothesis that $\text{Prob}(S_t = 1) = 0.5$. Note that this last result holds even if $\{Z_t\}$ is a dependent process, as it is the case for some more general ARCH representations—the weak-ARCH structure presented in Bollerslev, Engle and Nelson (1994). To put it another way rejection of IID questions the general ARCH specification, whereas rejection of the hypothesis that $\text{Prob}(S_t = 1) = 0.5$ suggests that no member of the broad symmetric ARCH class describes the data. There are, however, some problems with this testing strategy. First, this assumes that

[2]Normal, Student-t, generalized Student-t, and generalized error distributions appear to be the more commonly used distributions. An exception is the semiparametric ARCH model of Engle and Rivera (1993).

are no ARCH effects in the conditional mean or, at least, that such effects are weak. Second, tests based upon signs are weak because too much information is lost by the sign transformation. But a testing procedure motivated by the bispectrum may be possible. This is a good point of departure for a brief digressive discussion of two closely related issues in "omnibus testing": (i) Testing for "nonlinearity"; (ii) Testing a class of models such as, for example, symmetric ARCH-type.

We shall confine the meaning of the word "nonlinearity" to methods or models that cannot be analyzed by reduction to linearity via a change of units or extension of analogues of linear methods to higher conditional moments beyond conditional means. We must define what we mean by "stochastic linearity."

Following Brock and Potter (1993, and references to Hall and Heyde, and Priestley) call a zero mean strictly stationary stochastic process $\{Y_t\}$ with enough regularity so that it possesses a one sided (causal) Wold representation, *IID (MDS) linear* if it has a representation $Y_t = \sum \alpha_j \epsilon_t$ where the $\{\epsilon_s\}$ are Independent and Identically Distributed (Martingale Difference Sequence). We say $\{\epsilon_s\}$ is an MDS (w.r.t. the σ- algebra generated by past $\epsilon' s$) if,

$$E\{\epsilon_s | \epsilon_{s-1}, \epsilon_{s-2}, \ldots\} = 0, \text{ all s.}$$

Note that GARCH models with zero conditional means are MDS-Linear. Also note that since the Wold representation is essentially unfalsifiable (unless one tested for nonstationarity itself), it is not useful to call a strictly stationary process $\{Y_t\}$ "linear" if it has a moving average representation with uncorrelated errors. For this reason the notions of IID (MDS) linearity are introduced. Note also that MDS linearity implies that the best Mean Squared Error predictor is the best Linear predictor. More about this will be said in Section 2 where we discuss testing for linearity.

We exclude extended discussion of variations of ARCH and its cousins simply to reduce the scope of this survey and to reduce overlap with excellent surveys such as Bollerslev, Engle, and Nelson (1994) which are readily available elsewhere. However, we shall spend some time discussing issues related to ARCH models.

For example, a major part of the work that applied "nonlinearity tests" to financial data can be viewed as "residual diagnostics" of linear models or variations of GARCH models. Rejections of linear null models by such general "portmanteau" diagnostic tests such as the Bispectral Skewness Test (Subba Rao and Gabr (1980) and Hinich (1982)) or BDS (Brock, Dechert, and Scheinkman (1987)) may be due to many reasons besides neglected nonlinearity. It may be due to neglected nonstationarity, moment condition failure, mispecification of the linear model class, power against some members of the linear model class, etc.

Moment condition failure is an important issue in finance. Here is an example of a promising strategy that must be modified to deal with moment condition failure. Note that zero mean GARCH-type models are martingale difference sequences, hence a special case of general MDS-linear processes. So, consider the use of the bispectrum to test whether the stationary stochastic process $\{X_t\}$ has the one sided MDS-Linear representation,

$$X_t = \sum_{u=0}^{\infty} g_u \epsilon_{t-u}$$

where $E\{\epsilon_s|\epsilon_{s-1},\ldots\} = 0$, $s = 1, 2, \ldots$ Assume, W.L.O.G., that $\mathrm{EX}_s = 0$, all s. Compute third order cumulants $\mu(s_1, s_2) \equiv \mathrm{E}X_t X_{t+s_1} X_{t+s_2}$ as in Priestley (1981). One is lead to examination of terms of the form $E\{\epsilon_s \epsilon_{s+k} \epsilon_{s+l}\}$. The MDS property of $\{\epsilon_s\}$ allows one to show that $E\{\epsilon_s \epsilon_{s+k} \epsilon_{s+l}\} = 0$ except for $k = l > 0$. At this point a version of the bispectral test could, perhaps, be designed to test the general MDS property by shutting off power against terms of the form $E\{\epsilon_s \epsilon_{s+k}^2\}, k > 0$. See Barnett et al. (1994, especially references to the work of Hinich and his co-authors) for discussion of bispectral tests. Returning to ARCH, note, however that $E\{\epsilon_s \epsilon_{s+k}^2\} = 0$, for $k > 0$ for a large class of ARCH-type processes. Consider the GARCH(p, q) class, $\epsilon_s = \sigma_s Z_s$ where,

$$\sigma_s^2 = \alpha_0 + \alpha_1 \epsilon_{s-1}^2 + \ldots + \alpha_p \epsilon_{s-p}^2 + \beta_1 \sigma_{s-1}^2 + \ldots + \beta_q \sigma_{s-q}^2; \;\; \{Z\} \text{ IIDN}(0, 1).$$

Compute to show $E\{\epsilon_s \epsilon_{s+k}^2\} = 0$, for all s, k, for the GARCH(p, q) class. Hence all third order cumulants are zero for GARCH(p, q)-driven linear processes. Hence the bispectrum is zero for such processes. To put it another way, the bispectrum is zero for any stationary process with a "Wold" type representation which is driven by GARCH(p, q) innovations. Since, in financial applications, the conditional mean of returns is small relative to the conditional variance, this suggests a potentially useful screening test for linear models driven by GARCH(p, q) innovations. However, there is a potential difficulty in carrying out this useful research strategy.

Innovations in models fitted to financial returns tend to have heavy tails. de Lima (1994a) shows that the bispectral test is badly sized for heavy tailed data. In particular he shows that the bispectral test requires finite sixth moments to be valid. Many financial datasets do not appear to have finite fourth moments much less finite sixth moments, and the bispectral test tends to reject a pareto IID null too often when its tail exponent is chosen compatible with that estimated for financial data sets. Hence this poses a potential practical problem to implementing the above "portmanteau test" for linear processes driven by GARCH(p, q) innovations. Nevertheless we believe that research into uses of variations on the bispectrum would be useful.

For example, one possible strategy to deal with de Lima's size problem is to bootstrap the bispectral skewness statistic under the null that the returns data under scrutiny lie in the GARCH(p, q) class. Of course this application of the bootstrap is well beyond the scope of the asymptotic theory that we have been able to find for the bootstrap (cf. Jeong and Maddala (1993), Vinod (1993), LePage and Billard (1992), Leger, Politis, and Romano (1992)). While there has been a lot of work on the "moving block bootstrap," work on bootstrapping the null distribution of interesting quantities (interesting to economists, at least) under parametric time series volatility models such as GARCH seems sparse. More will be said about this in Section four below. Let us now return to the initial theme of discussion of applications of "complexity theory" to finance.

1.2 Complexity theory in Finance

While "complexity theory" sometimes is taken to include chaos theory, we shall not spend much time on chaos theoretic applications to finance here. That topic has been covered by many reviews including, Abhyankar, Copeland, and Wong (1995), Brock, Hsieh, and LeBaron (1991), Creedy and Martin (1994), LeBaron (1994), and Scheinkman (1992).

"Complexity" theory is a rather vague term. We use it here to refer to the research practices of centers such as the Brussels School (e.g. Prigogine and Sanglier (1987)), the Stuttgart School (e.g. Weidlich (1991)), and the Santa Fe Institute. Indeed the notion of "complexity" is so hard to define that a recent SCIENTIFIC AMERICAN article on the subject by Horgan (1995) quotes the MIT physicist Seth Lloyd as having compiled a list of at least 31 different definitions of "complexity" that have been proposed. We shall take the strategy here of an "intellectual factor analysis." I.e. we extract a few broad themes that capture the bulk of the research practices of "complexity research" that we wish to cover for this particular article.

An important subset of these research practices includes building dynamical systems models of the form, $Y_t = h(X_t, \xi_t), X_t = F(X_{t-1}, \eta_t, \theta)$ where X_t is the state vector at date t, Y_t is the vector of observables emitted by the system, ξ_t is a stochastic shock that may hit the observer function h at time t, η_t is a stochastic shock that may hit the system's law of motion F, at time t and θ is a vector of "tuning" parameters or "slow changing" parameters. The long run behavior of the system for each fixed θ is studied by a mixture of analytic and computer-based methods. Then θ is varied to study how this long run behavior changes. These changes are associated with "emergent behavior" or "emergent structure."

There is also a major subtheme of this line of research which emphasizes how "simple rules" F can induce complicated behavior of the observables, Y. The hope of this subtheme of research is to use a combination of computer based and analytic based methods to catalogue "universality classes of F's" as mechanisms to generate different types of "complexity" and to use this research strategy to unearth a small number of universality classes of F's that generate the complex behavior we see in Nature.

Wide classes of systems are searched to catalogue similar "species" of emergent structure. "Routes to chaos" such as period doubling cascades of bifurcations are well known examples of this type of methodology. Descriptions of this type of research are in Allen and McGlade's study of fisheries (in Prigogine and Sanglier (1987)), Weidlich's survey of Stuttgart School research (Weidlich (1991)), and Krugman's discussion of uses of this style of research in international trade and economic geography (Krugman (1993)).

An interesting subtheme of "complexity" theory is the research into complicated systems whose inner mechanisms are so complex that they are studied by searching for "scaling laws" in observables emitted by the systems where "scaling laws" are broadly interpreted to include regularities of autocorrelations and cross correlations of asset returns, volatility of asset returns, and volume of trading across different assets such as different stocks, foreign exchange, etc.

The intent in searching for these "scaling laws" is the hope that they will be robust to the details of a particular complex system and that they will be approximately the same across broad classes of complex systems. Note the similarity to the objective of finding broad classes of dynamical systems where the "emergent behavior" as θ changes is the same within that class.

The hope that there are "universal scaling laws" across widely disparate complex systems is one of the well springs that drives this style of research. A drawback of this "universal scaling laws" style of research is that most of the scaling laws are unconditional statistical objects, whereas, in finance at least, we are much more interested in *conditional probabilities*. In many cases the set of data generating stochastic processes consistent with a given "scaling law" may just be too large to have much interest in finance. As an extreme example consider the set of stochastic processes $\{X_t\}$ consistent with the Central Limit Theorem. That particular kind of "universal scaling law" is of limited usefulness in discriminating among alternative data generating mechanisms in finance. Let us explain what we mean in more detail.

Much of statistics and econometrics centers around "Root N Central Limit Theorem" scaling:

$$n^{-1/2}\sum_{i=1}^{n}(X_i - \mathrm{E}X_i) \longrightarrow N(0,V), n \to \infty.$$

Here $\longrightarrow$ is weak convergence; $N(0,V)$ denotes normal with mean zero and variance V; and $\{X_t\}$ is a stochastic process with enough regularity so that the CLT is valid. For example $\{X_t\}$ could be weakly nonstationary and weakly dependent and the CLT would still be valid. But, while, such scaling is used in other ways, such as hypothesis testing, it is not very useful as a discriminator across the class of potential data generating mechanisms.

We shall be concerned in this article with mechanisms that lead to scaling that is *not Root N*. While such scaling still suffers from being a crude discriminator across the class of potential data generating processes, the hope is that the different scaling than CLT will lead to useful insights into what classes of data generating processes can generate such non root N scaling.

A good example of this particular style of research is Bak and Chen (1991) who attempt to show that a particular class of probabilistic cellular automata, called "sandpile" models, are good abstractions for a broad variety of complex systems encountered in nature. Think of a real sandpile sitting on a table with sand being dropped upon it from above and think of "sandslides," i.e., "avalanches," of various sizes being triggered by this falling sand when the sandpile reaches "criticality." Furthermore they argue that sandpile models exhibit "power law scaling" of observables such as the distribution of avalanche size and such power law scaling such as "$1/f$" noise is widely observed in nature. They argue that the robustness of the power law scaling to the details of particular sandpile automata is a "universal" property which makes the sandpile automaton a particularly useful metaphor for mechanisms that lead to power law scaling.

In economics Scheinkman and Woodford (1994) have argued that local interactions and strong nonlinearities can combine through forward and backward linkages to create a breakdown of the Root N central limit theorem in a model of inventory dynamics which is built along the lines of the sandpile model. "Final demand" plays the role of the driving force of falling sand in Scheinkman and Woodford.

A similar theme shows up with scaling near "phase transitions" in interacting systems models where it is argued that the form of this scaling is surprisingly robust to the details of the particular model under scrutiny (Ellis (1985, pp. 178-9)).

The common criticism that interacting systems models require "tuning" of an exogenous parameter to generate non Root N scaling can be blunted by reformulation within the context of discrete choice random utility theory where the intensity of choice (the "tuning parameter") becomes endogenous along the lines of Brock (1993). All that needs to be done to endogenize the intensity of choice is to make it a function of the difference between the utilities of the choice alternatives.

This can be motivated by modelling the tradeoff between costly choice effort and the gain in utility to expending such effort. One tractable way to do this is to set up a two stage problem where the first stage chooses $\{p_i\}$ to maximize entropy $E = -\sum p_i \ln(p_i)$ s.t. $\sum p_i U_i = \bar{U}(e), \sum p_i = 1$, and the second stage chooses effort, e, to maximize $\sum \bar{p}_i(e) U_i - c(e)$, where $\bar{p}_i(e)$ is the probability of choice i from the first stage.

This can be viewed as an adaptation of the ideas of E.T. Jaynes into an economic tradeoff where $\bar{U}(e)$ represents the average amount of utility garnered from random choice when effort level, e, is put into it. See Brock (1993) for references to Jaynes and more on the relationship between maximum entropy, discrete choice, and statistical mechanics.

In any event, whatever one's opinion on the need for "tuning" an outside parameter to "criticality," interactions models that generate non Root N scaling may play a role in understanding financial forces that lead to long term dependence and apparent non Root N scaling in the empirical work described in Section 3. Turn now to discussion of structural and empirical modelling by frequency. This is motivated by the belief that the economic forces differ by frequency.

1.3 Frequency based study: structural and empirical

It is useful to organize discussion of structural theoretic based models and empirical/statistical models in finance by frequencies. At the highest frequencies, tic by tic for example, the market microstructural institutions surely matter. Phenomena such as bid/ask bounce and nonsynchronous trading surely loom large. See the work of Grossman, Miller, Froot, Schwartz and others in the Smith Report (1990) and the work of Domowitz and his co-authors in Friedman and Rust (1993) for discussion of institutional rules, their impact on price discovery and volatility, as well as time series properties of returns at high frequency. Domowitz shows that an institutional quantity which he calls, "the length of the order book" plays a key role in inducing time series properties of returns, the bid/ask spread, and volatility at the very high frequencies.

For another example, the reader should examine the work of Froot, Gammill, and Perold for the Smith Report (1990) in order to see how the auto-correlation function at the 15 minute frequency for the S&P 500 cash index has moved dramatically closer to zero over the period 1983-1989 and the possible explanations given there. They argue that reduction in transactions costs coupled with new trading practices such as portfolio and futures trading have acted to impound new information into prices much more rapidly than before. The possibilities that changes in bid/ask bounce or changes in non trading effects explain the drop in predictability are discounted by Froot et al.

Experimental and theoretical work on auction theory and market microstructural institutions (for example see the discussion in Friedman and Rust (1993) and the Roger Smith Report (1990)) has documented differences in performance of different "auction" systems.

We put the noisy rational expectations models discussed in Grossman's book (1989) into the Medium to High frequency class. The Highest frequency class contains market microstructure models like those surveyed by Goodhart and O'Hara (1994) as well as the differences in auction institutions discussed above. Recent surveys at the Very High frequency are Goodhart and O'Hara (1994) and Guillaume et al. (1994). Methods designed to analyze financial phenomena at approximately weekly frequencies go into the Medium Frequency class.

In order to organize our discussion in this article we shall view the market micro structure as operating at the highest frequency (from tic by tic to the 15 minute frequency, perhaps), whereas the information arrival process and price discovery itself taking place at the next highest frequency (15 minute frequency to daily frequency, perhaps). We shall view the "discovered" prices themselves as moving at the next highest frequency. We shall also view phenomena such as bid/ask bounce and nonsynchronous trading as occurring at, perhaps, a slightly higher frequency than price discovery itself.

It is well known that there are "daily seasonalities" in volume of trade and volatility of returns associated with the open and the close. This intraday seasonality causes problems for time series analysis. Andersen and Bollerslev (1994), show that application of "traditional time series methods [...] to raw high frequency returns may give rise to erroneous inference about the return volatility dynamics. ...Moreover, de-seasonalization appears critical in uncovering the complex link between the short- and long-run return components, which may help explain the apparent conflict between the long-memory volatility characteristic observed in interday data and the rapid short-run decay associated with news arrivals in intraday data."

For example Brock and Kleidon (1992) propose a model that "explains" bid/ask spreads over the trading day from open to close and discuss evidence interpreted in the context of their model versus alternative models. We view this type of phenomena as, possibly, taking place at a higher frequency than the frequency of asymmetric information based theories such as Grossman (1989), but, taking place at, a lower frequency than phenomena induced by Domowitz's trading institutions in Friedman and Rust (1993).

At the other extreme, the lowest frequency is the growth frequency studied by Mehra (1991), for example. At this frequency long run movements in (i) technical change, (ii) institutional change in the private sector, (iii) institutional change in the government sector, (iv) the age distribution of the population, etc. play the major role.

We shall put methods designed to analyze monthly and lower frequencies into the Low Frequency class. One should think of these frequencies as business cycle frequencies or lower. For example, we put the Euler Equation and Consumption Based Capital Asset Pricing Model (CCAMP) based methods surveyed by Altug and Labadie (1994), Campbell, Lo, and MacKinlay (1993), and Singleton (1990), as well as models which focus on finance constraints and mean reversion such as Jog and Schaller (1994) into the low frequency class. We also put the structural exchange rate models based on explicit modelling of the demand for money which are surveyed by Altug and Labadie (1994) into the low frequency class. Of course some of these phenomena may operate at a higher frequency. The boundary we are trying to draw here is very vague.

In this article we wish to exposit some work that lies at the boundary of theory based structural approaches and "econometric" approaches. We also wish to put forward and motivate a view of specification testing in financial econometrics that may be somewhat controversial. Before we get formal let us try to explain in plain English what we mean.

In the work surveyed by Singleton (1990) and the recent related line of work by Duffie and Singleton (1993) an explicit theoretical economic model forms the basic launch point of the statistical analysis. In Singleton (1990) much of the analysis flows from the Lucas (1978) pure exchange asset pricing model and its relatives. Singleton (1990) concludes that "comovements in consumption and various asset returns are not well described by a wide variety of representative agent models of price determination." de Fontnouvelle (1995) surveys studies, including his own, which take transactions costs into account. Potentially realistic transactions costs appear to reduce some of the conflict with data.

In Duffie and Singleton (1993) the production based asset pricing models of Brock (1982) and Michener (1984) serve as the launch point for the statistical analysis which itself is a dynamic extension of the Simulated Method of Moments of McFadden (1989) and Pakes and Pollard (1989).

Contrast this approach with the GARCH literature which constructs statistical models of asset returns and estimates them with few attempts to *directly* derive such models from an underlying theoretic structure. Here the pure economic theory like that which serves as the foundation of the type of work surveyed by Singleton (1990) lies in the background at best in the "purely statistical" work discussed in surveys like Bollerslev, Engle, and

Nelson (1994).

At the daily frequency it is typical to associate movements in returns, volume, and volatility with the arrival of information. Development of a structural based approach that parallels the research in Singleton (1990) is still a topic for future research. For example, Lamoureux and Lastrapes (1994) quote Gallant, Rossi, and Tauchen (1992, p. 202) as saying the following about theoretically based models: "...they have not evolved sufficiently to guide the specification of an empirical model of daily stock market data." Lamoureux and Lastrapes (1994) go on to develop a "statistical" model of daily stock returns and daily volume. In section four below, we briefly describe some structural modelling that attempts to go part way towards an empirical model of daily stock market.

We believe that this gap between the structure of the theoretic models that inspire the econometric models and the structure of the econometrics models which are actually estimated will vanish as developments in extensions of the bootstrap to financial time series problems and developments in extension of dynamic methods of simulated moments proceed. Computational advances such as those techniques discussed in Judd's forthcoming book (1995) will play a key role.

This review is organized as follows. Section 1 contains the introduction. Section two discusses several tests of nonlinearity including the bispectral skewness test of Subba Rao and Gabr (1980), Hinich (1982), and the BDS test of Brock, Dechert, and Scheinkman (1987). It is pointed out that these tests are inconsistent. There are departures from linearity that these tests cannot detect. A discussion of some consistent tests follows. However, rejections of linearity of asset returns are common when these tests are used. The main issue in finance does not seem to be the *inability* to detect departures from linearity because rejections of linearity are so frequent. The main issue is to find reasons for the rejections. The discussion turns to the possibility that fat tailed returns distributions may be responsible for the rejections. This motivates methods of estimation of tail thickness. The discussion will present evidence that some of the tests reject the null too frequently under moment conditions appropriate for use on the heavy tailed data common in financial applications.

Section three explores possible nonstationarities and long term dependencies, such as long memory, in asset returns. In view of the recent interest in long memory processes both in academic finance and in more popular writings in finance, we provide a fairly complete discussion of long memory both in returns and volatility of returns. Topics covered include Fractionally Integrated Generalized AutoRegressive Conditional Heteroskedasticity (FIGARCH), a cousin (FIEGARCH), Stochastic Volatility Models, Hurst Exponents, and Rescaled Range statistics. It is shown that the rescaled range test for long term dependence can be fooled by short term dependent Markov switching stochastic processes such as Hamilton and Susmel's (1994) SWARCH models. But the Hurst Exponent itself is more robust against this form of short term dependence. Finally, Section four gives a brief discussion of bootstrap-based specification tests of null models such as parametric representations of the Efficient Market Hypothesis. These tests are based upon statistics gleaned off of trading strategies. We also give a brief discussion of the use of asymmetric information theory to generate potential explanations for the stylized features of autocorrelations and cross correlations among returns, volatility of returns, and trading volume. Furthermore we show how a modification of received asymmetric information theory can serve as a potential explanation of abrupt changes in returns, volatility of returns, and trading volume that seem inexplicable by changes in "news."

2 Nonlinearity in stock returns

2.1 Lagrange multiplier and portmanteau tests of nonlinearity

A wide variety of tests for nonlinearity is available in the literature. We can broadly divide these tests into two categories, namely, tests designed with an alternative in mind—as the Lagrange multiplier class of tests (Rao's score test)—and portmanteau tests. Granger and Teräsvirta (1993) show that many of the available tests of nonlinearity have a Lagrange multiplier (LM) type interpretation. This class includes the Tsay (1986) test, the RESET tests of Ramsey (1969) and Thursby and Schmidt (1977), the neural network test of Lee, White and Granger (1993), White's (1987) dynamic information test, LM tests against ARCH effects (Engle (1982) and McLeod and Li (1984)), the LM tests of Saikkonen and Luukkonen (1988) against bilinear alternatives and exponential autoregressive models, and the LM test of Luukkonen, Saikkonen and Teräsvirta (1988) against smooth transition autoregressive models.

Two portmanteau tests of linearity are the bispectrum test of Subba Rao and Gabr (1980) and Hinich (1982) and the BDS test (1987). These two tests are among the few nonlinearity tests that do not have a Lagrange multiplier type-test interpretation and both tests are known to have power against a wide variety of nonlinear alternatives. This last characteristic has made these two tests quite popular among practitioners.

The bispectrum test is based on the fact that for a zero-mean linear process y_t the *skewness function*

$$\frac{|B(\omega_1,\omega_2)|^2}{S(\omega_1)S(\omega_2)S(\omega_1+\omega_2)} \tag{2.1}$$

is constant for all pairs of frequencies (ω_1,ω_2). $B(\omega_1,\omega_2)$ is the power bispectrum—the Fourier transform of the third-order cumulant $E[y_t y_{t+h} y_{t+k}]$—and $S(\omega)$ is the power spectrum—the Fourier transform of $E[y_t y_{t+k}]$. Hinich's (1982) test of linearity looks at the dispersion of estimates of the skewness function at different frequencies.

The BDS test is a function of the Grassberger-Procaccia correlation integral, $C_{\epsilon,m} = \binom{N}{2}^{-1} \sum\sum_{1\le s<t\le N} \chi_\epsilon(\|Y_t^m - Y_s^m\|)$, for N observations of the time series y_t, where $Y_t^m \equiv (y_t, y_{t+1}, \ldots, y_{t+m-1})$, $\|.\|$ is the max-norm, and $\chi_\epsilon(.)$ is the symmetric indicator kernel with $\chi_\epsilon(x) = 1$ if $|x| < \epsilon$ and 0 otherwise. BDS (1987) show that if y_t is an independent and identically distributed (iid) sequence, then $C_{\epsilon,m} = (C_{\epsilon,1})^m$ as $N \to \infty$, and the statistic

$$\mathrm{BDS}_{\epsilon,m} \;=\; \sqrt{N}\,\frac{C_{\epsilon,m} - (C_{\epsilon,1})^m}{s_{\epsilon,m}} \tag{2.2}$$

converges in distribution to a standard normal distribution, for $\epsilon > 0$ and $m = 2, 3, \ldots . s_{\epsilon,m}$ is an estimate of the asymptotic standard deviation of $\sqrt{N}\left(C_{\epsilon,m} - C_{\epsilon,1}^m\right)$ under the null of iid. A simple interpretation of the test can be given by noting that $C_{\epsilon,m}$ is an estimator of $Pr\{\|Y_t^m - Y_s^m\| < \varepsilon\}$, while $C_{\epsilon,1}$ is an estimator of estimates $Pr\{\|y_t - y_s\| < \varepsilon\}$. Under the null of iid

$$\begin{aligned} &Pr\{\|Y_t^m - Y_s^m\| < \varepsilon\} = \\ &\quad Pr\{|y_t - y_s| < \varepsilon|, \ldots, | < \varepsilon, |y_{t+m-1} - y_{s+m-1}| < \varepsilon\} \simeq \left(Pr\{|y_t - y_s| < \varepsilon\}\right)^m \end{aligned}$$

that is, the BDS test estimates the difference between the joint distribution and the product of the marginal distributions in the appropriate intervals. Note that this analogy is not complete because there might be some overlap between y_{t+i} and y_{s+j}.

The BDS test becomes a portmanteau test of linearity if applied to the estimated residuals of a linear model. The null distribution of the test is not affected by this procedure, provided that $\sqrt{N}$-consistent estimation of the parameters of the null model is possible.[3] Proofs of this result are available in the original BDS (1987) paper, as well as Brock, Hsieh and LeBaron (1991) and de Lima (1995). The first two papers derive their result using continuous approximations to the indicator kernel $\chi_\epsilon(.)$. The approach taken by de Lima (1995) generalizes results by Randles (1982) to deal with $\chi_\epsilon(.)$ directly. In particular, these results show that if the data generating process is an ARMA(p,q) model driven by iid innovations with finite second moments, the estimation of the parameters of the ARMA process does not affect the null distribution of the BDS test. Furthermore, this statement remains valid if the linear process has an autoregressive representation driven by iid innovations whose distribution is a member of the family of stable distributions, that is, the nuisance parameter-free property of the BDS statistic applies to a large class of linear processes with infinite variances—see de Lima (1995).

While the local power properties of most LM-type tests of nonlinearity are relatively easy to characterize—see for example Granger and Teräsvirta (1993), the distributions of the bispectrum and BDS tests are not known under the alternative hypothesis. For that reason, a considerable number of papers have studied the power properties of these tests by means of Monte Carlo simulations, c.f. Brock, Hsieh and LeBaron (1991), Lee, White and Granger (1993), and Barnett et al (1994). As expected, the corresponding LM-tests seem to dominate for alternatives that are local to the null hy-

[3]This nuisance parameter-free property of the BDS test remains valid if the test is aplied to data generating processes that are additive in the error term, $y_t = G(X_t, \beta) + U_t$, where β is a vector of parameters and X_t is a (vector of) time series, satisfying a mixing property. Moreover, this property carries through to some multiplicative models of the type $y_t = G(X_t, \beta)\ U_t$, provided that the test is applied to $\ln(\hat{U}_t^2)$, where $\hat{U}_t$ are the estimated residuals. This last result shows that by means of an appropriate transformation of the residuals, the null asymptotic distribution of the BDS test is not affected by the use of estimated residuals from GARCH and EGARCH processes. See Brock and Potter (1993) and de Lima (1995) for both analytical and simulation results.

pothesis. However, these tests are usually not very powerful against other departures of the null, while the BDS test appears quite powerful for almost every departure of the null—for example, as documented by Brock, Hsieh and LeBaron (1991), the power of the BDS test against ARCH alternative is close to Engle's (1982) LM test. This is true for both nonlinear stochastic processes and nonlinear deterministic, chaotic alternatives.

2.2 Consistent tests of linearity

It should be noted that neither the BDS test nor the bispectrum test are consistent tests of nonlinearity, that is, there are known departures from linearity for which these tests have zero power. Dechert (1988) presents an example of a dependent process that the BDS test has no power to detect. Also, there are nonlinear processes that exhibit a flat skewness function, such as GARCH processes. The asymptotic power of the bispectrum test of linearity against GARCH processes is zero, because of the tests' failure to recognize the nonlinearity behind the flat skewness function.[4]

Bierens (1990) presents a consistent conditional moment test. The test is closely related to the neural network test described in Lee, White and Granger (1993) and it can be used as a consistent test of linearity in the mean. The null hypothesis for the test is defined as $E[y|X] = X'\beta$, almost surely, where (y, X) is a vector of iid random variables defined on $\mathbb{R} \times \mathbb{R}^k$ and β is $ak \times 1$ vector of parameters. Alternatively, one could define the random variable $u = y - E[y|X]$ and test the hypothesis that $E[u|X] = 0$. The mean independence between u and X implies that $E[u \ \Psi(X)] = 0$, for any function $\Psi(X)$. Bierens shows that the choice $\Psi(X) = \exp(s'\phi(X))$ generates a consistent conditional moment test. Here Φ is an arbitrary bounded one to one mapping from $\mathbb{R}^k$ into $\mathbb{R}^k$, and $s \in S$, where S is some subset of $\mathbb{R}^k$. de Jong (1992) extends Bierens' results into a framework that allows for data dependence and for the fact that the conditional expectation of y_t might depend on an infinite number of random variables, that is, $y_t = E[y_t|z_{t-1}, z_{t-2}, \ldots] + u_t$ where $z_t = (y_t, X_t)$. In other words, under the null of linearity, the disturbance terms u_t are a martingale difference sequence.

[4]It has been suggested that for such type of processes nonlinearity can be detected using higher-order polyspectrum based tests. The sample size requirements of such tests appear exceedingly demanding—see Barnett et al. (1994).

The practical implementation of this consistent conditional moment test faces some difficulties. First, not much is known about the size and power properties of this test. In particular, different mappings Φ are likely to have a significant impact on the small sample properties of the test. However, from a distributional point of view, the choice of s is a more delicate issue. Consistency is achieved by considering some functional of the process

$$M(s) = N^{-1/2} \sum_{t=1}^{N} \Big((y_t - X_t'\hat{\beta}) \exp(s'\phi(X) \Big)$$

with $M(s)$ viewed as a random element of the space of continuous functions on a compact subset of $\mathbb{R}^k$. Bierens presents two alternative approaches to construct a consistent test from the empirical process $M(s)$. The first one (Bierens 1990, theorem 3, p. 1450) gives rise to a statistic with an asymptotic distribution function that depends on the distribution of the data. Therefore, critical values for the test statistic have to be simulated each time the test is applied to a different data set. The second approach (Bierens 1990, theorem 4, p. 1451) produces a tractable null distribution but the resulting test statistic is discontinuous in sample size.

A few alternatives to this conditional moment test have been proposed in the literature. Wooldridge (1992) proposes a test that compares least squares estimation of the null model with a sieve estimator—e.g. White and Wooldridge (1991)—of a compact approximation to the alternative model. Note that the alternative hypothesis defines an infinite dimensional set. Therefore, as the sample size grows, the sieve estimator must be defined on an increasingly larger dimensional space. Similarly, de Jong and Bierens (1994) consider a consistent chi-square test where the (possibly) misspecified conditional mean function is approximated by means of series expansions.

Bradley and McClelland (1994a and 1994b) propose a modification of the Bierens test that provides a (asymptotically) most powerful test among the class of consistent conditional moment tests. Let $\hat{u}$ be the estimated residuals from least squares estimation of the model $y_i = X_i'\beta + u_i$, where the observations $\{(y_i, X_i) : i = 1, 2, \ldots, N\}$ are a random sample from a distribution function $F(y, x)$, such that $E[y|X] = \Theta' X$. Bradley and McClelland (1994a) show that $\Psi(X) = E[\hat{u}|X]$ is the function that maximizes $E[\hat{u}\Psi(X)]$ among the set of bounded functions. This guarantees consistency—$E[\hat{u}E[\hat{u}X]]$ is different from zero whenever $E[\hat{u}\exp(s'\phi(X))]$ is non zero. $E[\hat{u}|X]$ is estimated by nonparametric kernel methods with bandwidth selection determined by cross-validation. To avoid overfitting problems associated with this procedure—which would result in size distortions,—Bradley and McClelland apply resampling techniques to the estimated residuals. This may be a potential problem for time series applications, namely if the conditional variance is not constant over time. Also, the nonparametric kernel method used to estimate $E[\hat{u}|X]$ under the alternative may not be appropriate in a time series context as the misspecified conditional mean function might involve an infinite number of variables.

2.3 Nonlinearities and fat-tailed distributions

The derivation of the (asymptotic) null distribution of statistical tests requires technical assumptions on the nature of the distribution that generates the data. In particular, some moment conditions are usually imposed so that a central limit theorem can be applied to the test statistic under study. A simple test of the hypothesis that the mean of a random variable X is μ_0 illustrates the problem quite clearly. Two type types of auxiliary assumptions are brought in: the type of temporal dependence in the data and a moment condition. If random sampling can be assumed, a finite second moment guarantees that the Lindberg-Levy central limit theorem can be used to approximate the distribution of the sample mean.

The same type of auxiliary moment condition assumptions need to be made in the derivation of the asymptotic distribution of nonlinearity tests. All the tests summarized in section 3 assume that the data is generated by distributions with at least finite fourth-order moments. The only exception is the BDS test—see de Lima (1994a). This is a consequence of the fact that the moment conditions required for convergence of the BDS statistic to a normal random variable apply to the indicator kernel $\chi_\epsilon(.)$.[5] Because $\chi_\epsilon(.)$ is a binary variable all its moments are finite. However, some moment conditions need still to be imposed because the BDS test is applied to the estimated residuals of an ARMA(p, q) and this estimating process should involve $\sqrt{N}$-consistent estimation techniques. As mentioned previously, iid innovations with finite variances are sufficient for $\sqrt{N}$-consistent estimation

[5]This is a direct application of results by Denker and Keller (1983) on central limit theorems for U-statistics.

of the parameters of an ARMA(p, q) model.

The robustness of nonlinearity tests to moment condition failure is of particular relevance for financial time series. The fatness of the tails of the distribution of stock and other financial asset returns is a well established stylized fact. Financial time series exhibit excess kurtosis. Furthermore, Mandelbrot (1963) provides evidence that unconditional second moments might not exist for commodity price changes. This has lead him to suggest the family of stable distributions as an alternative to the gaussian model. It should be noted that although the normal distribution is itself a stable distribution, it is the only member of the stable family that has finite second moment (and all other higher-order moments).

Random variables that belong to the family of stable distributions have some nice theoretical properties. For example, they are the only family of distributions with domains of attraction and closed under addition.[6] Their usefulness as a model for financial time series has been strongly contested, though. Alternative characterizations of the marginal distribution of stock returns have been proposed—e.g. the t-student distribution of Blattberg and Gonedes (1974), and the mixture model of Clark (1974). Hsu, Miller and Wichern (1974) provide evidence that nonstationarities in the variance may bias Mandelbrot's statistical methods in favor of the stable model. Comparisons of these different approaches as well as discussions of the efficiency of the statistical methods involved in the estimation of the distributions are described in, among others, Fielitz and Rozelle (1983), Akigary and Booth (1988), Akigary and Lamoureux (1989), for stock returns, and Boothe and Glassman (1987) and Koedijk, Schafgans, and de Vries (1990) for exchange rates.

More recently, Jansen and de Vries (1991) and Loretan and Phillips (1993) take a more direct approach to the problem of determining the existence of moments. Instead of trying to characterize the entire distribution, these two papers concentrate on the tails of the distribution, because the existence of moments is ultimately determined by rate of decay of the tails of the density function. Loretan and Phillips (1993) present estimates of the maximal moment exponent, $\alpha = \sup_{q>0} E \mid X^q \mid < \infty$, for a group of stock market and exchange returns. The parameter α is estimated using the procedure

[6]See Zolatarev (1986) for an extensive survey and Samorodnitsky and Taqqu (1994) for some recent developments.

developed by Hill (1975) and Hall (1982). Let $X_1, X_2, \ldots, X_N$ be a sample of independent observations on a distribution with (asymptotically) Pareto-type tails. Let $X_{N,1}, X_{N,2}, \ldots, X_{N,N}$ represent the ordered sample values. The maximal moment exponent can then be consistently estimated by

$$\hat{\alpha}_s = \left(s^{-1} \sum_{j=1}^{s} \ln X_{N,N-j+1} - \ln X_{N,N-s} \right)^{-1}$$

for some positive integer s. Letting s grow with the sample size (although at a smaller rate), Hall (1982) shows that $s^{1/2}(\hat{\alpha}_s - \alpha)$ converges to a $N(0, \alpha^2)$ random variable. Loretan and Phillips (1993) estimates suggest that variances are finite but fourth moments may not exist. In other words, these results provide strong evidence against gaussianity but also show little support for the stable model.

Mittnik and Rachev (1993) and Pagan (1995) argue, however, that the estimator used by Loretan and Phillips is not a very reliable measure of the shape of the tails of the unconditional distribution of asset returns. First, different choices for s—the number of order statistics—appear to produce significantly different estimates of the maximal moment exponent α, especially when the number of observations is not very large. However, for reasonably sized samples, Loretan's (1991) simulations indicate that $\hat{\alpha}_s$ is a robust estimator of α if s does not exceed 10% of the sample size. This rule of thumb was first suggested by DuMouchel (1983). Second, the convergence results provided by Hall (1982) assume random sampling. Pagan (1995) reports simulation results showing that the standard deviation of $\hat{\alpha}_s$ can be significantly larger than predicted from the iid case if the data is generated from a GARCH process. Note that GARCH-type processes generate heavy-tailed distributed data: de Haan, Resnik, Rootzen, and de Vries (1989) show that the unconditional distribution of ARCH variates has Pareto tails and de Vries (1991) presents a GARCH-type model where the unconditional distribution is stable. Furthermore, estimation of GARCH models for high frequency stock returns data usually produces parameter estimates that imply that fourth moments do not exist. Nelson (1990) shows that an IGARCH(1,1) model, although strictly stationary, does not have a finite variance.

The consequences of using nonlinearity tests when moment condition failure is an issue are investigated de Lima (1994a). From the point of view of asymptotic theory it is shown that the distribution of the tests becomes non-standard. As an example, for iid sequences that do not have finite fourth moments, it is shown that the normalization of the sum of the squares of the first h autocorrelations of the process by the number of observations does not provide convergence to a nondegenerate random variable (see de Lima 1994a, Proposition 1.) In other words, for this type of processes the McLeod-Li statistic collapses asymptotically to zero.[7]

Simulation experiments presented in de Lima (1994a) show that most nonlinearity tests behave as predicted by the asymptotic result derived for the McLeod$-Li$ test. In particular, the sampling distributions of those tests exhibit a pole around the origin. This would suggest that under moment condition failure and without the appropriate scaling of the tests' statistics, the empirical sizes would always be below the tests' nominal sizes. However, the simulation experiments also reveal that the variance of the tests can be extremely large, giving rise to a significant number of large values for the tests' statistics. This effect is especially more pronounced for extreme cases of moment failure. Further, tests that are designed to have maximal power against misspecification of the conditional variance as well as the bispectrum test seem to be especially sensitive to the non-existence of moments.[8] Overall, the only test that appears robust to moment condition failure—in both the asymptotic and the sampling distributions—is the BDS statistic.

de Lima (1994a) presents a study of the relationship between nonlinearities and moment condition failure in a sample of 2165 individual stock returns listed in the 1991 Daily Stock files of the CRSP tapes. The median value of $\hat{\alpha}_s$ in the sample is 2.8 with more than 95% of the estimates above 2 (finite variance) and less than 2% above four. The application of the nonlinear-

[7]Note that an appropriately scaled version of the McLeod$-Li$ statistic converges to a well defined random variable, although the limiting random variable does not have a chisquare distribution and the rate of convergence is slower than for the standard case.

[8]The bispectrum test appears particularly sensitive to the problem of moment condition failure. Simulation experiments reported in de Lima (1994a) show that for iid sequences generated from the Pareto family of distributions satisfying

$$(P_\alpha) \qquad \begin{cases} P(X > x) & = \ .5(x+1)^{-\alpha}, \quad x < 0 \\ P(X < -x) & = \ .5(x+1)^{-\alpha}, \quad x > 0 \end{cases}$$

with $\alpha = 1.5$ (the maximal moment exponent) and 5000 observations, the 1%-sized test rejects the null of iid in 60% of the cases. Similarly large type-I errors are found for values of α between 2 and 6.

ity tests to randomly shuffled series shows a remarkable resemblance to the simulation experiments. This empirical study also shows that evidence of nonlinearity in stock returns can not all be attributed to the non-robustness of nonlinearity tests to moment condition failure. However, it shows that some of those tests are not very trustworthy in testing situations involving heavy-tailed data.

2.4 Other topics in nonlinearity testing

2.4.1 Nonlinearities and nonstationarities

Constancy of the moments of the unconditional distribution of asset returns is a typical assumption of many time series models, including volatility processes such as GARCH. However, given the rate at which new financial and technological tools have been introduced in financial markets, the case for existence of structural changes (and thus for lack of stationarity) seems quite strong, especially when relatively large periods of time are considered. For example, Pagan and Schwert (1990) and Loretan and Phillips (1993) reject the hypothesis that stock returns are covariance-stationarity. Therefore, it is of particular interest to determine whether findings of nonlinearity might be due to nonstationarities in the data.

In terms of GARCH models, Diebold (1986) and Lamoureux and Lastrapes (1990) suggest that shifts in the unconditional variance could explain common findings of persistence in the conditional variance. Simonato (1992) applies a GARCH process with changes in regime—using Goldfeld and Quandt (1973) switching-regression method—to a group of European exchange rates and finds that consideration of structural breaks greatly reduces evidence for GARCH effects. Another model that tries to capture the idea that several volatility periods are present in the data is Cai (1994) and Hamilton and Susmel (1994) Markov switching ARCH (SWARCH) model.[9]

A characterization of stock returns as nonstationary processes with discrete shifts in the unconditional variance can be traced back to Hsu, Miller and Wichern (1974). Hinich and Patterson (1985) challenge this view, supporting the alternative hypothesis that stock prices are realizations of nonlinear stochastic processes. They argue that nonstationarities would bias the

[9] See section 3 for a more general discussion of variance persistence and the SWARCH model.

bispectrum test used in their analysis toward acceptance of linearity. Given that their tests statistics clearly reject this hypothesis, they discard the existence of nonstationarities in daily stock returns during the period July 1962 through December 1977. Using the BDS test, Hsieh (1991) rejects the hypothesis that structural breaks are responsible for the rejection of linearity by means of subsample analysis and by looking at data with different (higher) frequencies. Because the BDS test rejects the null hypothesis for all different subsamples and frequencies, Hsieh concludes that "[...] it is unlikely that infrequent structural changes are causing the rejection of IID [...]". The distinction between nonlinearity and nonstationarity is also central to Inclán (1993), who presents a nonparametric approach to distinguish between shifts in the unconditional variance and a time-varying conditional variance.

de Lima (1994b) uses a generalization of the BDS test to investigate whether rejections of linearity for stock market returns are due to nonstationarities in the data. This paper uses the fact that normalized partial sums of the BDS statistic converge to standard Brownian motion and analyses common stock returns indexes between January 1980 and December 1990. It is shown that the period that goes from October 15, 1987 and November 20, 1987 assumes an extremely influential role in the rejection of nonlinearity provided by the BDS statistic for the entire period: for any subsample period starting in January 1980 and ending before October 15, 1987 the BDS test would not reject the null of linearity. Note that Diebold and Lopez (1995), using the autocorrelation function of the squared returns, conclude that evidence for GARCH effects in stock returns during the eighties is also small. However, de Lima (1994b) results also indicate that nonlinearities seem to play an active role in the dynamics of stock indexes after October 1987.

2.4.2 Identification of nonlinear alternatives

Despite their usefulness as general tests for nonlinearity, a rejection of the null by any of the two portmanteau tests described above gives the applied researcher little or no guidance on the actual nature of nonlinearities that might be causing the rejection of the null hypothesis. A test closely related to the BDS test, due to Savit and Green (1991) and Wu, Savit and Brock (1993) is of particular interest in this regard. Instead of relying on estimates of unconditional probabilities, these two papers propose a test that uses correlation integral type estimators of the sequence of conditional probability statements,

$$\begin{aligned}
\text{Prob}\,\{A_{t,s}|A_{t-1,s-1}\} &= \text{Prob}\,\{A_{t,s}\} \\
\text{Prob}\,\{A_{t,s}|A_{t-1,s-1}, A_{t-2,s-2}\} &= \text{Prob}\,\{A_{t,s}|A_{t-1,s-1}\} \qquad (2.3) \\
\vdots \qquad\qquad & \qquad\qquad \vdots \\
\text{Prob}\,\{A_{t,s}|A_{t-1,s-1}, \ldots, A_{t-k,s-k}\} &= \text{Prob}\,\{A_{t,s}|A_{t-1,s-1}, \ldots, A_{t-k+1,s-k+1}\}
\end{aligned}$$

where $A_{t,s} = \{(y_t, y_s) : |y_t - y_s| < \epsilon\}$. These equalities hold under iid and using the definition of conditional probabilities it can be shown that they can be estimated by correlation integral-type quantities. Savit and Green's (1991) insight is that, under the alternative, these conditional probabilities can be used to detect at which lag temporal dependence is strongest. This type of analysis is of particular interest to Markov processes, commonly used in nonparametric time series analysis—e.g. Robinson (1983) and Gallant, Rossi and Tauchen (1993). Alternative approaches to the identification of nonlinear time series processes include the nonparametric version of the final prediction error criterion of Auestad and Tjöstheim (1990) and Tjöstheim and Auestad (1994) and Granger and Lin (1994) mutual information coefficient (relative entropy),

$$\delta(f, f_x f_y) = \iint f(x,y) \log \left\{ \frac{f(x,y)}{f_x(x) f_y(y)} \right\} dx\, dy$$

where (x, y) is a pair of random variables with joint density function $f(x, y)$ and marginals $f_x(x)$ and $f_y(y)$. See also Granger and Teräsvirta (1993) for a general discussion of the use of nonparametric techniques in nonlinear modeling.

2.4.3 Multivariate extensions

Conditional probability statements of the type described in (2.3) can also be used to detect whether there are nonlinear causal relations between variables. Baek and Brock (1992a) define nonlinear Granger causality in the following terms: a time series $\{y_t\}$ does not cause $\{x_t\}$ if

$$\text{Prob}\left\{A_{t,s}(X^m)|A_{t-h,s-h}(X^h), A_{t-k,s-k}(Y^k)\right\} = \text{Prob}\left\{A_{t,s}(X^m)|A_{t-h,s-h}(X^h)\right\} \qquad (2.4)$$

where $A_{t,s}(W^m) = \{(W_t^m, W_s^m) \| W_t^m - W_s^m \| < \epsilon\}$, for $W = X, Y$. This means that the random variable y has no predictive power for x. Rewriting expression (2.4) in terms of ratios of unconditional probabilities and estimating the corresponding terms by correlation integral type statistics, Baek and Brock (1992a) show that (a normalized version of) the resulting statistic converges to a normal random variable, under the null hypothesis of non-causality from y to x. Baek and Brock (1992a) and Hiemstra and Jones (1994a) present alternative estimators of the asymptotic variance under different assumptions about the dependence properties for y and x. As for the univariate testing procedures involving the BDS statistic, the tests for nonlinear Granger causality are applied to estimated residuals of linear models. In the present case, nonlinear predictive power consists of any remaining predictive power that is left in the series after the data is filtered by a vector autoregressive model.

Hiemstra and Jones (1994*a*) apply this testing strategy to daily stock returns and percentage changes in trading volume. Their work provides evidence of nonlinear Granger causality in both directions. However, note that the nonlinear impulse response analysis of Gallant, Rossi and Tauchen (1993), while supporting the idea that returns Granger-cause trading volume, does not detect a significant feedback mechanism from volume to prices.

Correlation integral based methods have also been employed to detect general nonlinearities in multivariate setups. Baek and Brock (1992b) generalize the BDS test for the null hypothesis that a vector of time series is temporal and cross sectional independent.

3 Long memory in stock returns

3.1 Long memory in the mean

The random walk hypothesis has dominated the empirical work on the characterization of the long run behavior of asset prices. The methods used to test this hypothesis include autoregressions of multiperiod returns—Fama and French (1988)—and variance ratio tests—Lo and MacKinlay (1988) and Poterba and Summers (1988). These two methods are closely related—see, for example, Kim, Nelson and Startz (1991)—and their application reflects a concern with the power of traditional tests to detect interesting alternatives to the null hypothesis of market efficiency.

One commonly studied alternative is the mean-reverting behavior of stock prices, corresponding to the idea that a given change in prices will be followed, in long time horizons, by predictable changes with opposite sign. This hypothesis describes stock prices—p_t—as the sum of a random walk$---p_t^*$— and a stationary component—u_t. Summers (1986) argues that the transitory component is a slowly decaying process, namely an AR(1) process $u_t = \rho u_{t-1} + \epsilon_t$, where ϵ_t is a white noise process and ρ is close to but less than one.

Lo and MacKinlay (1988) and Poterba and Summers (1988) report variance ratio statistics that give some support to the hypothesis that stock prices are mean reverting. In particular, variance ratios appear to be greater than one for lags shorter than a year and below unity for longer lags. As the variance ratio statistic at lag q is a weighted sum of the first q autocorrelations of stock returns—Cochrane (1988) and Lo and MacKinlay (1988), the observed pattern of variance ratios implies that stock returns are positively correlated over short time horizons, and negatively correlated over longer intervals. Note that this predictability of long-horizon returns is consistent with models where (some) agents behave irrationally (noise traders) as well as with efficient markets with time-varying equilibrium expected returns.

Kim, Nelson and Startz (1991) and Richardson (1993), among others, have presented evidence that the tests used to detect mean reverting behavior might produce spurious results.

A new disaggregated approach to the study of mean reversion which uses data on individual firms, and stresses the structural role of variation of financing constraints across different classes of firms is in Jog and Schaller (1994). Differential variation of finance constraints across different classes of firms (such as different size classes) appears to be a promising way to explain the well known variations in mean reversion across periods of financial stress such as the Great Depression as well as a promising way to respect scale economies in raising funds that can be exploited by larger firms. Also one expects the impact of central bank policy to vary across different classes of firms.

Lo (1991) takes a somewhat different approach than the rest of the aggregative literature, to present a simple alternative model that generates a similar pattern for the variance ratio statistics. Lo's (1991) example assumes that stock returns are the sum of an $AR(1)$ and a long memory process.

Long memory stationary processes are characterized by the slow (hyperbolic) decay of their autocorrelation function, as opposed to short memory processes (such as ARMA) whose autocorrelation function exhibits geometric decay. Alternatively, a long memory process can be characterized by the behavior of its spectral density function at the origin.[10] Long memory processes can generate non-periodical cyclical patterns as the ones observed by Hurst (1951) for the Nile River, where long periods of dryness are followed by long flood periods. Mandelbrot and Wallis (1968) coined this phenomenon as the Joseph or Hurst effect.

The first paper that discusses the importance of long memory processes in asset markets is Mandelbrot (1971). Mandelbrot shows that under long range dependence perfect arbitraging is not possible. Mandelbrot has raised an important point here which has been expanded upon by Hodges (1995) to show that Fractal Brownian Motion is not a promising model for stock returns unless the market is grossly inefficient. He calculates that "for a market with a Hurst Exponent outside the range 0.4 to 0.6 less than 300 transactions would be required" to obtain "essentially riskless profits." He provides a useful table which relates Hurst exponent values, Sharpe Ratios, and numbers of transactions needed to capture profits under options strategies.

Hodges has cast a lot of doubt on the plausibility of "long memory in mean" with Hurst exponents that deviate very far from 1/2. This is so because it is very easy to manufacture the profits and control the risks in a mean/variance setting if the returns data are truly generated by a Fractal Brownian Motion with Hurst exponent very far from 1/2.

Whatever the surface plausibility of long memory, because traditional methods of financial economics rely heavily on the possibility of arbitraging, the detection of long memory in stock returns has emerged as a relevant empirical question. Greene and Fielitz (1977) is the first empirical investigation of the long memory hypothesis for stock returns. Their analysis relies heavily

[10]The autocorrelation function—$\rho(k)$—of a long memory process satisfies $\rho(k) \sim Ck^{2H-2}, C > 0$, for $0 < H < 1$. For $H > 1/2, \sum \rho(k) = \infty$, whereas for $H < 1/2, \sum |\rho(k)| < \infty$ and $\sum \rho(k) = 0$. Correspondingly, the spectral density $f(\omega) = \sum e^{-i\omega}\rho(k)/2\pi$ diverges at the origin for $H > 1/2$ and tends to zero as $|\omega| \to 0$. Some authors reserve the term long memory for the first type of processes, and label the second type as "intermediate" memory or anti-persistent. See Beran (1994) and Brockwell and Davis (1991). For a survey of long memory processes and their application to Economics, see Baillie (1995).

on the rescaled range (R/S) statistic first proposed by Hurst (1951). For a time series X_t and any arbitrary time interval of width s and starting point t, the sample sequential range $R(t,s)$ is defined as

$$R(t,s) = \max_{0\leq k\leq s}\left\{X^*_{t+k} - \left(X^*_t + \frac{k}{s}\left[X^*_{t+s} - X^*_t\right]\right)\right\} \\ - \min_{0\leq k\leq s}\left\{X^*_{t+k} - \left(X^*_t + \frac{k}{s}\left[X^*_{t+s} - X^*_t\right]\right)\right\}$$

where X^*_t is the cumulative sum of X_t over the interval from 0 to t, that is, $X^*_t = \sum_{u=1}^{t} X_u$, with $X^*_0 = 0$, for convenience. The sample range is usually normalized by the standard deviation for the lag s,

$$S(t,s) = \left(\frac{1}{s}\sum_{k=1}^{s} X^2_{t+k} - \frac{1}{s^2}\left[\sum_{k=1}^{s} X_{t+k}\right]^2\right)^{1/2}$$

and the resulting ratio is known as the rescaled range R/S. In a series of papers, Mandelbrot and some of his co-workers have shown that the rescaled range statistic can distinguish between short and long memory processes, in the sense that for a stationary process with short range dependence the R/S statistic converges to a non degenerate random variable at rate $s^{1/2}$, whereas for processes that exhibit long range dependence the R/S statistic converges to a non degenerate random variable at rate s^H, where H, the Hurst coefficient, is different from 1/2—see Mandelbrot (1975). Moreover, theorem 6 in Mandelbrot (1975) establishes that the rate of convergence is also $s^{1/2}$ for iid sequences in the domain of attraction of stable distributions with infinite variance.

In practical terms, the plot of the logarithm of the R/S statistic against the logarithm of s, for different values of s, should reveal whether the data was generated by a short-range or long-range dependent process: the different points should be spread around a straight line with slope 1/2 for short-range dependent processes and slope $H \neq 1/2$ for long-range dependent processes. Wallis and Matalas (1970) present a Monte Carlo simulation of two alternative procedures for selecting lags and starting points, known as F Hurst and G Hurst. In both cases, the estimate of H, the exponent of long-range dependence, is the slope of the least squares regression of $\log(R/S)$ on a constant and on $\log(s)$. Greene and Fielitz (1977) conduct such analysis on the daily returns to 200 common stocks listed in the New York Stock Exchange, concluding that long-term dependence characterizes a significant percentage of the sample. More recently, Peters (1994) also uses R/S analysis and provides evidence of the Hurst effect in the returns to some common financial assets.

These findings of long memory in stock returns have been disputed on the grounds that classical R/S analysis is biased by the presence of short-term dependence, a fact already discussed by Wallis and Matalas (1970) and further studied by Davies and Harte (1987). Aydogan and Booth (1988) suggest that the Greene and Fielitz (1977) results might indeed be the outcome of the non-robustness of classical R/S analysis to serial dependency and non-stationarities. To correct for the bias induced by serial correlation, Peters (1994) applies classical R/S analysis to the estimated residuals of first order autoregressive processes. Furthermore, he compares the values of the R/S statistics obtained for different lag lengths with the expected value of the R/S statistic. This expected value was computed by Anis and Lloyd (1976) for white noise processes. The value used by Peters (1994) reflects a correction term determined by simulation. However, note that Peters (1994) method still does not allow for formal hypothesis testing and his working assumption that an AR(1) filter removes short-term serial dependence for all series under test is highly questionable.

Lo (1991) presents a refinement of R/S methods that allows formal statistical testing and is robust to serial correlation and some forms of non-stationarity. Under the null of short-memory,[11] Lo shows that the statistic $Q(n) \equiv R(1,n)/\tilde{S}(1,n)$ converges weakly to the range of a Brownian bridge on the unit interval, a random variable with mean $\sqrt{\pi/2}$ and variance $\pi^2/6 - \pi/2$ and whose distribution function is positively skewed. The main innovation of Lo's procedure is the use of the Newey-West heteroskedasticity and autocorrelation consistent estimator,

$$\tilde{S}(1,n)^2 = \frac{1}{n}\sum_{k=1}^{n}(X_k - \bar{X})^2 + \frac{2}{n}\sum_{j=1}^{q}\omega_j(q)\left\{\sum_{k=j+1}^{n}(X_k - \bar{X})(X_{k-j} - \bar{X}\right\}$$

in place of of $S(1,n)^2$. Furthermore, Lo's test does not have to rely on

[11] A short-memory process is defined by Lo as a strong mixing process whose mixing coefficients decay sufficiently fast to zero.

subsample analysis as the classical R/S analysis. Lo (1991) applies the $Q(n)$ statistic to daily and monthly stock returns indexes (the equally and the value weighted indexes on the CRSP files) and concludes that Greene and Fielitz (1977) methods overstate the existence of long memory in stock returns.

The $Q(n)$ statistic—also known as the modified R/S statistic—has been applied by several researchers to other financial data sets—see Bailley, namely Cheung and Lai (1993) to gold market returns, Cheung, Lai, and Lai (1993) and Crato (1994) to international stock markets, Goetzmann (1993) to historical stock returns series, Hiemstra and Jones (1994b) to a panel of stock returns and Mills (1993) to monthly UK stock returns. The evidence produced by these papers is largely concurrent with Lo's (1991) results, with the transformed R/S statistic finding little evidence of long memory in the returns to those financial assets. However, Pagan (1995) stresses that the choice of q, the number of autocorrelations included in the Newey-West estimator $\tilde{S}(1,n)^2$ is critical in terms of the results, with a small q usually providing evidence favorable to the alternative (as in the traditional Greene and Fieltiz application where q is set to zero), and a large q supporting the null. Andrews (1991) provides an automatic selection rule for q(also used by Lo (1991) in his application. However, this rule has optimal properties only for AR(1) processes.

One additional problem with the $Q(n)$ statistic appears to be its sensitivity to moment condition failure. Hiemstra and Jones (1994b) uncover a positive relation between maximal moment estimates and the probability of a left-tail rejection by the R/S test in their sample of stock returns. The relationship appears reversed for right-tail rejections. Note that, as mentioned previously, Mandelbrot (1975) and Mandelbrot and Taqqu (1979) show that the classical R/S analysis provides an almost surely consistent estimator of the Hurst coefficient even for iid data generated by infinite variance processes. However, these two papers provide no characterization of the limiting distribution of the R/S statistic. Furthermore, Lo (1991) proves convergence to the range of a Brownian bridge under the assumption that the first $4+\delta$ ($\delta > 0$) moments of the distribution of the data are finite.

A simple simulation study, reported in table 1, appears to confirm that while heavy-tailed data do not seem to affect the properties of the R/S estimator of the Hurst coefficient, the sampling distribution of the test is shifted to the left relatively to the asymptotic distribution, as observed by Hiemstra and Jones (1994b). Table 1 reports the results of computing the R/S statistic

Table 1: Estimated Sizes of the Rescaled Range (R/S) test under moment condition failure

Nominal Size		$\alpha = 1.5$					$\alpha = 4$				
		0	20	40	60	A	0	20	40	60	A
Left	0.01	0.023	0.020	0.017	0.014	0.023	0.018	0.013	0.010	0.007	0.018
Tail	0.05	0.098	0.092	0.081	0.075	0.093	0.071	0.062	0.051	0.044	0.072
	0.10	0.174	0.172	0.166	0.156	0.177	0.136	0.124	0.110	0.099	0.130
Right	0.10	0.030	0.024	0.022	0.016	0.030	0.092	0.086	0.078	0.080	0.093
Tail	0.05	0.009	0.000	0.005	0.006	0.009	0.038	0.039	0.035	0.031	0.036
	0.01	0.000	0.000	0.000	0.000	0.000	0.009	0.007	0.008	0.007	0.008
	Mean	0.510	0.510	0.510	0.511	0.510	0.521	0.521	0.521	0.522	0.521
	Std	0.028	0.027	0.027	0.026	0.028	0.031	0.030	0.030	0.029	0.031

The data were generated from a Pareto distribution with $\alpha = 1.5$ and $\alpha = 4$, respectively. Each of the 1000 series had $N = 5000$ observations. The rows labeled as Mean and Std report the mean estimate of the Hurst coefficient and its standard error in the simulations. Each column reports the empirical size of the R/S test for a different number of autocorrelations q included in the estimator $\tilde{S}(1, n)$. The column labeled as A corresponds to Andrews (1991) optimal rule.

over 1000 series with 5000 observations generated from the family of Pareto distributions—see expression (P_α) in section 2—with parameters $\alpha = 1.5$ and $\alpha = 4$. The average estimate of the Hurst coefficient is close to 0.5 as determined by Mandelbrot. However, rejection rates on the left tail are above the nominal sizes given by the asymptotic distribution, whereas rejection rates on the right tail are below the nominal sizes given by the asymptotic distribution.[12] The shift in the empirical distribution is more pronounced for data generated with $\alpha = 1.5$ than for data generated with $\alpha = 4$. This fact should not come as a surprise given the moment assumptions made by Lo (1991).

Other tests of the long memory hypothesis are available in the literature. This set includes Geweke and Porter-Hudak (1983)—hereafter GPH, the locally optimal and beta-optimal tests of Davies and Harte (1987), the Lagrange multiplier tests developed by Robinson (1991a) and Agiakloglou,

[12] Left-tail rejections correspond to rejection of the null hypothesis $H = 1/2$ against the alternative $H < 1/2$ (anti-persistent long memory) while right-tail rejections correspond to rejection of the null hypothesis $H = 1/2$ against the alternative $H > 1/2$ (persistent long memory).

Newbold and Woahr (1994), and the locally best invariant test of Wu (1992), closely related to the goodness of fit statistic of Beran (1992).[13] In opposition to the modified R/S statistic, all these tests assume a parametric form for the alternative hypothesis, although the GPH test only requires a parametric specification of the long run dynamics of the alternative process. For this reason, the GPH test is sometimes designated as a semiparametric test.

The dominant parametric discrete-time model that exhibits hyperbolic decay of its autocorrelation function is the fractional integrated autoregressive moving average model (ARFIMA) introduced independently by Granger and Joyeux (1980) and Hosking (1981)—see Viano, Deniau, and Oppenheim (1994) for a continuous time version. For $-0.5 < d < .5$, X_t is said to follow an ARFIMA(p,d,q) model if it is the unique stationary solution to the equation

$$(1-B)^d \phi(B) X_t = \theta(B)\eta_t, \quad \eta_t \sim \text{iid} N(0, \sigma_\eta^2)$$

where B is the backshift operator $(B^j \ X_t = X_{t-j}, j = 0, \pm 1, \pm 2, \ldots)$, $\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \ldots - \phi_p z^p$ and $\theta(z) = 1 - \theta_1 z - \theta_2 z^2 - \ldots - \theta_q z^q$. Furthermore, the fractional differencing operator is defined through the expansion

$$(1-B)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} B^j. \tag{3.1}$$

See Brockwell and Davis (1991) for a detailed treatment of this model. The spectral density of an ARFIMA(p, d, q) model is proportional to $C \mid \lambda \mid^{-2d}$ as $\mid \lambda \mid \to 0$, for $C > 0$. The Geweke and Porter-Hudak (1983) test for long memory is based on this fact: regress the logarithm of the periodogram at low frequencies on some function of those frequencies and estimate d by the slope of this least squares regression.[14] GPH argued that the resulting estimator of d could capture the long-memory behavior without being contaminated by the short-memory behavior of the process. Robinson (1992) showed that this argument is asymptotically correct if, besides truncation of the higher periodogram frequencies, an additional truncation of the very first ordinates is performed. The usual t-test of the hypothesis that $d = 0$ against $d \neq 0$ is a

[13] Cheung (1993a) provides a Monte Carlo investigation of the small sample properties of some of the more popular tests of the long memory hypothesis.

[14] For alternative estimation procedures of this regression equation see Beran (1993), Robinson (1993), and Jensen (1994).

test of the null hypothesis of short-memory against long-memory alternatives. It should be noted that the small sample properties of both the GPH and Lo's rescaled-range test can be very sensitive to large autoregressive and moving average effects—see Cheung (1993a).

Using the GPH approach, Cheung (1993b) finds some evidence of long memory in a set of nominal exchange rates and Cheung and Lai (1993) show that some linear combinations of foreign and domestic prices are long range dependent, that is, foreign and domestic prices are fractionally cointegrated. In their cross-section of stock returns, Hiemstra and Jones (1994b) find a close relationship between rejections of the short-memory null using the R/S statistic and the GPH test.

3.2 Long memory in volatilities

One of the more active research areas in long memory models is their application to volatility processes. This follows the analysis of conditional variance models started with Engle's (1982) seminal paper on autoregressive conditional heteroskedasticity (ARCH) models. ARCH models are defined as $y_t = \sigma_t Z_t$, where Z_t is usually taken to be an independent, identically distributed process, with $E(Z_t) = 0$ and $\text{Var}(Z_t) = 1$. The variable σ_t^2 is a positive, $\mathcal{F}_{t-1}$-measurable function, where $\mathcal{F}_{t-1}$ is the sigma-algebra generated by $(Z_{t-1}, Z_{t-2}, \ldots)$. Therefore, σ_t^2 is the conditional variance of the process y_t.

Typically, the sample autcorrelation function of stock returns series resembles the autocorrelation function of a white noise process. However, the sample autocorrelation function of measures of volatility, such as the squared returns, the absolute returns, or the logarithm of squared returns, is positive with very slow decay. This fact explains why many applications of GARCH-type models involving high-frequency data indicate the presence of an approximate unit root in the univariate representation for volatility. This feature is present in the original Engle (1982) paper, and it has motivated some of the extensions of Engle's original work, namely Bollerslev (1986) generalized ARCH (GARCH) and Engle and Bollerslev (1986) integrated GARCH (IGARCH). Furthermore, applications of Nelson's (1991) exponential GARCH (EGARCH) model usually find roots to the autoregressive polynomial close to the unit circle. That is, high-frequency stock market data displays highly persistent volatility.

The very slow decay of the autocorrelation function of the squared residuals motivated Crato and de Lima (1994) to apply the modified R/S and the GPH test to the squared residuals of various filtered U.S. stock returns indexes. The hypothesis that volatilities are short memory processes is clearly rejected for high frequency series. The rationale for applying long memory tests to the squared series comes from the fact that the conditional variance σ_t^2 of a GARCH(p,q) process can be written as an infinite-dimensional ARCH(∞), as in Bollerslev (1986). Therefore, this testing procedure parallels the Lagrange multiplier tests for GARCH effects, which are also performed on the squared series. Ding, Granger and Engle (1993) also study the decay of the autocorrelations of fractional moments of returns series. For returns (y_t) on the $SP500$ index, they construct the series $\mid y_t \mid^\nu$ for different positive values of ν and find very slow decaying autocorrelations. This has lead them to introduce a new class of ARCH models, the asymmetric power-ARCH, where ν becomes a parameter to be estimated. However, this model is still finitely parameterized, making it a short-memory model.

Two class of models have been proposed to capture the slow decay of the autocorrelation function of volatility series. One such class includes the fractional integrated GARCH (FIGARCH) and the fractionally integrated EGARCH models of Baillie, Bollerslev and Mikkelson (1993) and Bollerslev and Mikkelson (1994), and it is the natural extension of the ARCH class of models that allows a hyperbolic rate of decay for lagged squared innovations. The second class of long memory volatility models are the stochastic volatility models of Harvey (1993) and Breidt, Crato and de Lima (1994).

The FIGARCH(p,d,q) model is defined as

$$(1-B)^d\phi(B)(\sigma_t^2-\mu)=\theta(B)(y_t^2-\mu)$$

where $\phi(z)$ and $\theta(z)$ are p-th and q-th order polynomials, respectively and $(1-B)^d$ is defined as in (3.1). Like IGARCH processes, the FIGARCH process is strictly stationary but not covariance stationary, because the variance is not finite. Consequently, the autocovariance function of y_t^2 is not defined and the use of spectral and autocovariance methods is not directly possible. Furthermore, the asymptotic properties of the (quasi)-maximum likelihood estimators discussed by Baillie, Bollerslev and Mikkelson (1993) rely on verification of a set of conditions put forward by Bollerslev and Wooldridge (1992). At this point, it is not yet known whether those conditions are satisfied for FIGARCH processes.

The FIEGARCH(p,d,q) model, defined as

$$\log \sigma_t^2 = \mu_t + \theta(B)\phi(B)^{-1}(1-B)^{-d} g(Z_{t-1})$$

defines a strictly stationary and ergodic process. Moreover, $(\log \sigma_t^2 - \mu_t)$ is a covariance stationary process if $d < 0.5$. Note that the function

$$g(Z_t) = \delta_1 Z_t + \delta_2(\mid Z_t \mid - E \mid Z_t \mid)$$

was introduced by Nelson (1991) to capture the fact that stock price changes tend to be negatively correlated with changes in stock volatility, the so-called leverage effect. The asymptotic properties of the maximum likelihood estimator of the parameters of the FIEGARCH model are also dependent on verification of the same set of conditions put forward by Bollerslev and Wooldridge (1992).

Simulation experiments in Baillie, Bollerslev, and Mikkelson (1993) show that if a GARCH process is fitted to data generated by a FIGARCH model, the estimates obtained for the autoregressive polynomial imply roots that are very close to the unit circle, as it is typical in financial data. Moreover, in their application of the FIGARCH model to the exchange rate between US dollars and the German mark, the hypothesis of IGARCH behavior against fractionally integrated behavior is clearly rejected. Similar results are obtained by Bollerslev and Mikkelson (1993) in their application of the FIEGARCH model to daily stock returns on the Standard and Poor's 500 stock index.

The second class of models that allows for long memory in volatilities is the stochastic volatility class of models of Harvey (1993) and Breidt, Crato and de Lima (1994). A stochastic volatility model is an unobserved components model obtained as the product of two stochastic processes, say $y_t = \sigma_t Z_t$, where Z_t can be defined as for the ARCH model case, but σ_t^2 is no longer an $\mathcal{F}_{t-1}$-measurable process. Taylor (1986) assumes that the volatility logarithm $\ln(\sigma_t)$ follows a stationary, Gaussian AR(1) process. Note that stochastic volatility processes can be seen as the Euler approximation to the continuous time models used in theoretical finance, where the asset price $P(t)$ and the volatility $\sigma(t)$ each follow a diffusion process. Taylor (1994) presents a recent survey of the alternative specifications assumed for the volatility process. Breidt, Crato and de Lima (1994) propose a stochastic volatility model that captures the slow decay of the autocorrelation function of the

Table 2: Rejections of short-memory in a sample of stock returns using the Geweke and Porter-Hudak (GPH) and the Rescaled Range (R/S) tests

	GPH		R/S	
	X	X^2	X	X^2
10% Test	16.4%	72.1%	12.0%	51.8%
5% Test	10.5%	65.3%	4.0%	41.6%
Mean	0.511	0.794	0.524	0.563
Std	0.165	0.188	0.031	0.034

The rows labeled as Mean and Std report the mean estimate of the Hurst coefficient and its standard error across return series for the GPH and R/S methods. The GPH test was computed for frequencies between $N^{0.1}$ and $N^{0.5}$. The number of autocorrelations considered in the R/S test follows Andrews (1991).

(logarithm of the) squared returns through an ARFIMA process for a function of the volatility process. Specifically, it is assumed that $\sigma_t = \sigma \exp(v_t/2)$, where v_t is a long memory process independent of Z_t. It is straightforward to show that both y_t and y_t^2 are covariance and strictly stationary. After some transformations the model can be written as $x_t = \mu + v_t + \epsilon_t$, where $x_t = \log y_t^2, \mu = \log\ \sigma^2 + E(\log Z_t^2)$ and ϵ_t is iid with mean zero and variance $\pi^2/2$, under the assumption that Z_t is Gaussian. x_t inherits the long memory properties from v_t.

In their application of the long memory stochastic volatility model to stock returns, v_t is an ARFIMA$(1, d, 0)$ model with an estimated value for the differencing parameter d of 0.444. A standard t-statistic test clearly rejects the hypothesis that a short-memory process generated the data. The model is estimated by maximizing the Whittle's frequency-domain approximation to the Gaussian likelihood of the model. It is shown that this procedure gives consistent estimators of the parameters of the model.

As with many other parameterizations concerning volatility processes, the robustness of the findings of long memory in the variance of stock returns processes remains yet to be addressed. In the first place, there are not many economic arguments available to support these statistical findings. Bollerslev and Mikkelson (1994) suggest that long memory in the volatilities of stock market indexes is a consequence of aggregation, because individual returns appear to have less persistent volatility. Granger (1980) shows that the sum of $AR(1)$ processes with coefficients drawn randomly from a suitable distribution approaches a long memory process, as the number of terms in the sum increases. The same result can be derived in the context of short-memory stochastic volatility models, with aggregation generating the observed long memory in the market index. However, a simple application of long memory tests to a sample of 2165 returns extracted from the CRSP tapes, seems to contradict this hypothesis. The results presented in table 2 for the level series are consistent with the results obtained by Hiemstra and Jones (1994b) for a similar sample, displaying little evidence of long memory in the means. However, both long memory tests indicate that a large percentage of the series exhibits some evidence of long memory in volatilities. However, these results should be taken with extreme care because, as shown in Crato and de Lima (1994) short memory volatility processes such as GARCH can lead to rejections of the short memory null by any of the tests of long memory considered in table 3.

Models of conditional heteroskedasticity are likely to be misspecified. One way of comparing alternative specifications is by concentrating on the ability of the models to track some of the sample features. Breidt, Crato and de Lima (1994) show that the autocorrelation function of the logarithm of the squared process estimated from their long memory stochastic volatility model fits the sample autocorrelation quite closely. In particular, the model can replicate the slow decay of the sample autocorrelation function, a feature that a short-memory process like Nelson's (1991) EGARCH can not match. The traditional GARCH(1,1) and IGARCH(1,1) models also show problems in generating this type of autocorrelation function. However, it is well known that the presence of nonstationarities can generate spurious evidence of extremely persistent features in the data. As mentioned in section 2, nonstationarities have been suggested as explanation for the findings of persistence in the variance. Simulation results in Cheung (1993a) show that the R/S and the GPH test have robustness problems with shifts in the level of the series, which in terms of testing long memory in volatilities would mean that these two tests might have robustness problems to shifts in the variance.

In this regard, a particularly interesting model is the Hamilton and Susmel (1994) switching ARCH (SWARCH) model. In this model, there are a finite number of volatility states (s_t) and the state variable is governed by a Markov-chain with transition probabilities

$$\text{Prob}(s_t = j \mid s_{t-1} = i, s_{t-2} = k, \ldots, y_{t-1}, y_{t-2}, \ldots) = \text{Prob}(s_t = j \mid s_{t-1} = i) = p_{ij}$$

The return process is then defined as $y_t = g(s_t)^{1/2} u_t$, where $g(s_t)^{1/2}$ is constant at each different regime s_t and u_t is an ARCH-type model. Hamilton and Susmel (1994) consider several alternative ARCH specifications for u_t, including the Glosten, Jagannathan, and Runkle (1994) parameterization that incorporates leverage effects into the ARCH framework. In this particular parameterization—designated by SWARCH-L(p,q), where L stands for leverage effects$- - -\sigma_t^2$ is given by

$$\sigma_t^2 = \omega + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \ldots + \alpha_q u_{t-q}^2 + \delta d_{t-1} u_{t-1}^2$$

where d_{t-1} is a dummy variable that discriminates between positive and negative values of u_{t-1}^2. In the particular class of SWARCH models considered by Hamilton and Susmel (1994), the scale of the process changes with the regime but the parameters of the u_t are independent of the volatility state. Hamilton and Susmel (1994) fit SWARCH models to weekly stock returns.

This class of models presents slightly better one-period ahead forecasts (depending on the loss function considered) than more conventional GARCH models. For example, a SWARCH model with four volatility states is a conditional heteroskedastic model that has smaller mean squared error than a model with constant variance. It should be noted that this last model is a two parameter model (mean and variance) whereas the SWARCH model involves the estimation of fifteen different parameters. Some SWARCH type models may also lead to multimodal *unconditional* distributions which may be counterfactual.

To address the question of whether data generated from a SWARCH model would appear like a long memory volatility process to the R/S and GPH tests, we ran a small Monte Carlo simulation experiment. We took the student t SWARCH-L(3,2) estimated by Hamilton and Susmel (1994) and generated two sets of 1000 series, the first one with 1024 observations and the second one with 2048 observations. We computed the two tests on the level

Table 3: Estimated sizes of the Geweke and Porter-Hudak (GPH) and Rescaled Range (R/S) tests for SWARCH models

	GPH		R/S	
Size	X	X^2	X	X^2
0.10	0.162	0.327	0.067	0.466
0.05	0.099	0.236	0.029	0.353
Mean	0.501	0.658	0.525	0.565
Std	0.173	0.168	0.035	0.042

The data were generated from the Student-t SWARCH-L(3,2) model reported in Hamilton and Susmel (1994). Each of the 1000 series had N=1024 observations. The rows labeled as Mean and Std report the mean estimate of the Hurst coefficient and its standard error across return series for the GPH and R/S methods. The GPH test was computed for frequencies between $N^{0.1}$ and $N^{0.5}$. The number of autocorrelations considered in the R/S test follows Andrews (1991).

series, and again on the squared series. As expected, when applied to the levels, the tests indicate no evidence of long memory. However, table 3 shows that when applied to the squares of the series the tests spuriously detect evidence of long memory. Furthermore, the percentage of rejections is likely to increase if the data were generated from a SWARCH-L model estimated with higher frequency—e.g. daily—data. Similar results are reported by Crato and de Lima (1994) for data generated by gaussian GARCH and IGARCH models, where it is shown that the generated data tends to produce larger values of the long memory test statistics than the ones actually observed in the data.

However, the estimate of the Hurst coefficient provided by the R/S analysis might provide some useful information in discriminating between spurious rejections of the hypothesis that volatility processes are short memory processes against the alternative that they have long memory characteristics. Breidt, Crato and de Lima (1994) provide some Monte Carlo evidence that while the R/S statistic itself tends to over-reject the short-memory null, the estimate of the Hurst coefficient provided by the R/S statistic is close to its theoretical value of 0.5, when the number of autocorrelations included in the estimator $\tilde{S}(1, N)$ is given by Andrews (1991) optimal rule. Note that the mean of the estimated Hurst coefficients reported in table 3 using the R/S method is 0.525 for the level series and 0.565 for the squared series, with little variation in the simulation results. Using the same estimator, Breidt, Crato and de Lima (1994) report estimated values of the Hurst coefficient above 0.65 for the squared returns from the value weighted and equally weighted CRSP daily series. This point seems worthy of further investigation. Turn now to Section 4 where we shall give a brief discussion of bootstrap-based methods in statistical finance.

4 Bootstrap based specification tests and related methods

In this section we give a brief discussion of bootstrap-based specification testing and the use of asymmetric information models to explain stylized features of returns at the high to medium frequencies which include the daily frequency.

A recent surge of interest has developed around bootstrap-based specification tests or "goodness-of-fit" tests of parameterizations of the Efficient Market Hypothesis. See, for example, Brock, Lakonishok, LeBaron (1992), hereafter referred to as "BLL," Levich and Thomas (1993), Antoniewicz (1992), (1993), and Vaga (1994). We shall only give a brief discussion here which is based upon BLL.

BLL explore operationalizing the idea that various fundamentalist or technical trading rules have evolved over the decades (centuries?) as statistical tests of the null hypothesis of a particular parameterization of the Efficient Market Hypothesis, for example, the random walk model. They view the traders as acting like they are "as-if" Neyman Pearson statisticians designing a test of the EMH null against a set of alternatives against which they have a greater incentive, as well as a greater skill, to discover than academics.

BLL argue that the economic self interest of traders, as well as competitive pressures of survival, pushes them to find alternatives to the EMH null and to design tests of the null with maximal power against those alternatives.

Of course the scientist cannot observe much of this complicated process. However, the scientist can infer trading rules by examination of literature that caters to professional traders. Simply scan this literature and pick out trading rules that loom large in this literature.

In finance it is natural to argue that departures from natural attempts to parameterize the Efficient Market Hypothesis such as GARCH, EGARCH, et al., models and their cousins are likely to be so subtle that only traders whose lives depend upon it, will be able to design tests with good power against alternatives. The motivation of BLL is to use recent advances in bootstrap theory and computer technology to mimic these traders but do this in such a way to give statistical precision to the results. Explanation of the method of BLL of using trading rules to study the conditional distribution of stock returns will require some notation.

Let τ be a trading rule. The intuitive idea follows. We want to test whether a null class, $\mathbf{C}_o$(for example, a class of parameterized versions of the EMH) , of parametric models of the conditional probability distribution is adequate to describe a finite sample, $\{R_t, t = 1, 2, \ldots, N\}$ of stock returns. Here $R_t \equiv \log(P_t) - \log(P_{t-1})$, where P_t is the price of the stock at date t.

The idea is to estimate, by a good quality method such as a root N consistent method, a member M_o of this null class, $\mathbf{C}_o$, of parametric models of the conditional distribution. If the class adequately describes the data $\{R_t\}$ then the (at least root N consistently estimated) member, M_o, of $\mathbf{C}_o$ should fit the data well. We want the measure of "goodness-of-fit" to have economic meaning for at least two reasons. First, "statistical significance" may have little to do with "economic significance." Second, any model is, at best, a rather crude description of reality. Hence, we expect any good test to reject the null model under large samples, especially if it is asymptotically consistent against the set of alternatives to the null. When the test rejects, we want an indication of how wrong the null model is, as well as an economically meaningful instruction how to fix it. BLL used five "goodness-of-fit" measures: Conditional mean and variance following buy and sell signals and trading profits. To fix ideas let us consider the bootstrap distribution of trading profits under $M_o, \mathbf{C}_o$.

Take a trading rule, τ, and then use the bootstrap to estimate the "$(\tau)p$-value" of that null class, $\mathbf{C}_o$, by bootstrapping the distribution of trading profits using τ under the null of $\mathbf{C}_o$. The trading rule is to be heuristically viewed as a test of approximately optimal power against some alternative $\mathbf{C}_a$

that the traders who use trading rule τ think is out there in reality.

This strategy raises some points of departure from the mainstream of finance articles that study profits from trading rules. BLL are *not* interested in the issue of whether risk adjusted transactions cost adjusted profits exist for a trading strategy under scrutiny. The methodological posture of BLL is to use the trading strategy, together with bootstrap technology to create a specification test of the null model that is based on *financial* self interested behavior which the null model is attempting to model in the first place. This "self consistency" of the specification testing methodology with the behavior being modelled seems natural—especially in a subject like finance which studies sophisticated forward looking strategic behavior of traders.

Let us state the issues another way for emphasis. First, taking into account of transactions costs may actually dull the ability of the BLL procedure to detect departures from the null, $\mathbf{C}_o$. Hence transactions costs are ignored. Second, the issue whether profits remain after adjustment for risk and transactions costs relative to buy-and-hold is not directly relevant to the efficacy of the method. Third implementation of the method by use of simple "stone age technical trading tools" such as moving averages may be useful. Fourth, if τ is chosen after looking at the data the size of the resulting test will be distorted.

This point is closely related to the data snooping problem studied by Lakonishok and Schmidt (1988) and Lo and MacKinlay (1990). We attempt to deal with this problem by testing subsamples and by attempting to select rules that traders have been observed using. The latter, of course, makes an implicit assumption that actual traders are not locking onto worthless rules after snooping the data. This assumption may not be valid.

BLL use two classes of trading rules, i.e., moving averages and trading range break, to test the adequacy of the following four classes of univariate models using daily Dow returns, 1897-1986. The models are (i) Random walk with IID innovations, (ii) AutoRegressive models with IID innovations, (iii) Generalized Auto Regressive Conditionally Heteroskedastic in Mean (GARCH$-M$) with IID stochastic driver, (iv) Nelson's EGARCH with IID stochastic driver. The reader is free to use favorite trading rules, which may be mixtures of fundamentalist and technical trading strategies. Turn now to explanation of use of the bootstrap in this context.

4.1 Use of the bootstrap to estimate confidence intervals for technical trading profits under null models.

BLL compute the null distributions of four conditional moments (conditional mean and variance of h-period returns following buy and sell signals) and trading profits under four popular null models for returns $R_t \equiv \log(P_t/P_{t-1})$. All of these null models can be written in the returns form

$$R_t = H_1(R_{t-1}, \ldots, R_{t-K_1}; Z_t, \ldots, Z_{t-K_2}; \sigma_t^2, \ldots, \sigma_{t-K_3}^2, \theta_1), \quad (4.1)$$
$$\sigma_t^2 = G_1(R_{t-1}, \ldots, R_{t-L_1}; Z_t, \ldots, Z_{t-L_2}; \sigma_{t-1}^2, \ldots, \sigma_{t-L_3}^2 \theta_2), \quad (4.2)$$

where $\{Z_t\}$ is an IID process of ultimate "drivers" and θ_1, θ_2 are vectors of parameters to be estimated. If one takes σ_t^2 to be the conditional variance of returns, one can see that GARCH-M and EGARCH models and most of their cousins can be written in the form (4.1), (4.2).

The BLL procedure works as follows. Impose enough regularity conditions so that (4.1), (4.2) has a unique, ergodic stationary distribution (See Duffie and Singleton (1993), for example). Estimate (4.1), (4.2) and let $\hat{\theta}, \{\hat{Z}_t\}$ denote the estimated parameter vector and the estimated residuals. Since BLL were not explicitly interested in standard errors of $\hat{\theta}$, they conducted a "conditional" bootstrap of trading rule quantities as follows. Keep $\hat{\theta}$ fixed and conduct a Monte Carlo simulation as follows.

Let $\hat{F}_N$ denote the empirical distribution of $\{\hat{Z}_t\}$ which places mass $1/N$ on each $\hat{Z}_t, t = 1, 2, \ldots, N$. Fix $\max\{K_1, L_1\}$ initial R's, $\max\{K_2, L_2\}$ initial Z's, and $\max\{K_3, L_3\}$ initial σ^2's. Now resample $\hat{F}_N, B$ times by drawing N times from $\hat{F}_N$ with replacement. For each of the B, N-resamplings from $\hat{F}_N$, generate "fake" price data, $\{P_t^b, \sigma_t^{2b}; t = 1, 2, \ldots, N\}$ and compute the vector of quantities of interest, call it $T(\{Z^b\}_N, \hat{F}_N)$, where $\{Z^b\}_N \equiv \{Z_1^b, \ldots, Z_N^b\}$. In our case the quantities of interest are conditional means and variances of h-period returns following buy and sell signals of trading rule τ and profits generated by τ.

The bootstrap estimates

$$\text{Prob}\{T(\{Z\}_N, F) \epsilon A\} = J(F; A) \quad (4.3)$$

by

$$(1/B)\sum_{b=1}^{B} 1\left\{T(\{Z^b\}_N, F_N) \in A\right\}$$

Note that

$$(1/B)\sum_{b=1}^{B} 1\left\{T(\{Z^b\}_N, F_N) \in A\right\} \longrightarrow J(F_N; A), \text{ as } B \longrightarrow \infty.$$

Here Prob$\{.\}$ denotes the probability of event $\{.\}$ and $1\{T(\{Z^b\}_N, F_N)\epsilon A\}$ is the indicator function which is 1 if $T(\{Z^b\}_N, F_N)\epsilon A$ and is 0 otherwise. Note that application of the bootstrap requires us to show that (i) the quantity we wish to calculate can be written, in population, in the form (4.3), (ii) F_N is a good approximation to F, it(iii) as $N \to \infty, J(F_N; A) \to J(F; A)$, (iv) under the null hypothesis that a member of model class (4.1), (4.2) actually generated the returns data how close is $J(\hat{F}_N; A)$ to $J(F; A)$? (v) how good are tests based upon $J(\hat{F}_N; A)$ of the null hypothesis in terms of standard performance characteristics such as size and power?

The BLL approach was to "bootstrap" the empirical distribution of $T(\{Z\}_N, F_N)$ by,

$$\text{attach mass } 1/B \text{ to each } T(\{Z^b\}_N, \hat{F}_N), b = 1, 2, \ldots, B,$$

where each $\{Z^b\}_N$ represents N IID draws from $\hat{F}_N$, holding $\hat{\theta}$ *fixed*. Notice that the *true* Z's are not available. Only the *estimated* $\hat{Z}$'s are available. Hence these must be estimated consistently, under the null, as $N \to \infty$.

Note that step (iii) also presents a potential problem. Equations (4.1) and (4.2) are dynamics of *returns* not price *levels*. BLL show that the BUY and SELL events, as well as the h-period returns can be written as a time *stationary* function of lagged returns and other lagged state variables for a *finite* number of lags independent of N. Note that even though price levels are not stationary, *returns* are stationary and all quantities that are bootstrapped by BLL are shown to be stationary functions of returns.

The stage is now set for application of a version of bootstrap to estimate $J(F; A)$ and set up a specification test of the null by comparing the estimate $J(F_N; A)$ with the value of the same vector of quantities of interest for the actual data.

Since BLL hold $\hat{\theta}$ *fixed* rather than re-estimating it through each pass through the bootstrap algorithm, we shall call this a *conditional* bootstrap. Since no asymptotic theory seems to be available for bootstrapping BLL-type statistics under the GARCH-type time series nulls, one could conduct "quality evaluation" experiments by generating a long fake returns series from the estimated GARCH-type model, and giving "fake econometricians" samples of various lengths $N_1, N_2, \ldots, N_K$, upon which to estimate the GARCH-type model and to bootstrap the null distribution of profit-based statistics. Convergence behavior of the null distribution may be studied as N is increased. for very large values of B. Strictly speaking one would have to study the convergence behavior as both N and B are increased, but let us think of B being taken approximately infinity by the availability of cheap computer time for each N. These experiments indicated that convergence quality was usuable as soon as B was larger than 500.

Of course, we would much rather have theory available for convergence of bootstrap distributions under weak dependent parameterized time series null models like those studied by BLL. The presentation of computer evidence of convergence quality seems useful until theory is available.

4.2 Structural models and stylized features of stock returns

Turn now to a brief discussion of recent efforts to bring asymmetric information models closer to explanation of empirical features of high to medium frequency asset market data.

Recent works by Sargent (1993), Wang (1993), (1994), Brock and Lebaron (1995), de Fontnouvelle (1995), and references to the works of Admati, Campbell, Grossman, Hellwig, Lang, Litzenberger, Madrigal, Pfleiderer, Singleton, Stiglitz and others, have pushed the theory of asymmetric information models closer to an empirical model capable of explaining features of market data at higher frequencies than the business cycle frequencies stressed in the macrofinance works surveyed by Singleton (1990), and Altug and Labadie (1994). Without getting into formal detail let us attempt to give a description of some of this work and the stylized features of market activity that we wish the models to reproduce.

Here are the stylized features: (i) The autocorrelation function of returns on individual assets is approximately zero at all leads and lags. This is a stylized statement of a version of the Efficient Markets Hypothesis. (ii) The autocorrelation function of a measure of volatility such as squared returns or absolute value of returns is positive with a slowly decaying tail (slower decay for indices). Feature two is a stylized version of the "ARCH" type phenomenon which has stimulated a voluminous "statistical" literature (cf. Bollerslev, Engle, and Nelson (1994)). Evidence for the slow decay of the autocorrelation function of volatility was discussed in Section three of this article. (iii) The autocorrelation function of trading volume has a similar shape to that of volatility. We shall call features (ii) and (iii) volatility and volume "persistence."

(iv) The cross correlation function of volume and volatility is positive for volatility with current volume and falls off rapidly to zero for leads and lags. There may be some asymmetry in the falling off in leads versus lags (e.g. Antoniewicz $(1992a, b)).(v)$ Short term predictability in the near-future increases when near-past volatility falls (LeBaron (1992)). (vi) Abrupt changes in returns, volatility, and trading volume occur which are hard to attach to "news." Turn now to an informal description of asymmetric information models.

At each point in time risk averse traders receive signals on components of the actual future value of assets that are being traded today. Signals are random variables which are equal to the component of future value plus noise. Precision is the ratio of the component variance to the signal noise variance. A background level of trading volume is generated by different realizations of signals even though the precision is the same. Trading volume is also generated by disparity in the precisions of signals across traders.

If the structure of the model is common knowledge and traders are rationally conditioning on price and signals then the famous no-trade theorems of Milgrom, Stokey, and Tirole (cf. Sargent (1993) for a nice exposition) assert that volume will dry up unless a source of randomness is added so that traders are forced to "signal process."

Wang's papers (1993), (1994) give elegant closed form solutions to a class of dynamic heterogenous agent asymmetric information models which reproduce some of the stylized features of market data. However, no work except that of Brock and LeBaron (1995) and de Fontnouvelle (1995) both endogenizes the information structure and calibrates the resulting models to see how closely they replicate the features (i)-(vi) above.

Brock and LeBaron (1995) build an asymmetric information model with short lived assets and short lived traders where traders decide whether to spend resources on purchase of a precise signal to sharpen their conditional expectation on the end-of-period value of the asset or spend nothing and get a publically available crude conditional expectation. Call the actual end-of-period value of the asset, the "fundamental." The fundamental is a random variable which the market is pricing. The information purchase decision is based upon a discrete choice random utility model where the deterministic part of the utility is based upon a distributed lag measure of trading profits. The trading profits are calculated along an equilibrium path.

de Fontnouvelle (1995) develops a much more sophisticated model along the same lines, but with infinite lived assets. He shows how persistence in the profit measure that governs the choice of signal purchase generates persistence in volatility and volume. It appears that if his profit measure decays slowly enough his model may be able to produce slowly decaying autocorrelation functions for volatility and volume. This may shed some light on the slow decay of volatility autocorrelations documented in Section three.

de Fontnouvelle "solves" his model by developing an expansion around a known solution. Both models discussed here reproduce features (i)-(iv) with some limited success. Hence, since it has infinite lived assets, the de Fontnouvelle (1995) model may be a candidate for estimation on high frequency returns and volume data somewhat along the lines of Duffie and Singleton (1993).

If one "backs off" from "ultra" rationality and does not allow traders to condition on the equilibrium price function then this kind of model generates trading volume which is persistent provided the heterogeneity of traders is persistent. The trader heterogeneity can be made persistent in the Brock and Lebaron (1995) model provided that the decision whether or not to purchase the signal is made on a slower time scale than the time scale of data observation. Infinite lived assets, together with slow decay of the distributed lags in the profit measure allow de Fontnouvelle to produce persistence without introduction of a slower time scale for information purchase decisions.

Volatility of price changes (or returns) depends upon the average precision of the market. The average precision of the market is defined to be the weighted average of precision of each trader type with the fraction of traders in that type. Volatility of price change is higher when market precision is high because the market is closely "tracking" the random end-of-period value which it is attempting to price. When precision is lowest, price change is proportional to the change in publically available conditional expectations. If the publically available information is very "coarse" this price change could be small.

This observation contains a lesson, which is, perhaps, obvious to academics, but maybe not to commentators in the press: Observed market volatility can not be automatically associated with problematic "excess" volatility.

It can be shown that volatility persistence may be magnified provided that the precision purchase decision is made on a slower time scale than the data time scale. The precision purchase decision might be considered as a metaphor for the "style" of the traders, i.e. whether they are "short term", "medium term," or "long term" traders. This is so because at least part of the cost of signal precision is the opportunity cost of traders in maintaining their trading expertise and information base. Hence, for high frequency data, it may be plausibly realistic that the "style" of the traders does not change as fast as the data is collected.

It is of interest to ask whether volatility persistence is inherent in the fundamental which the market is attempting to price, i.e. "estimate," or whether the market pricing process itself adds the volatility persistence. If traders are risk averse, volatility persistence in the fundamental can make them timid in their trading so that the contemporaneous correlation between volume and volatility damps enough to conflict with stylized feature (iv). Brock and LeBaron (1995) and de Fontnouvelle (1995) discuss this potential conflict with the stylized features unless the volatility persistence is being added by the market pricing process itself. Even though the above argument suggests the possibility that the market pricing process itself may be adding volatility persistence over and above the volatility persistence which is in the fundamental, the jury is still out on this issue.

Now consider the impact of adding "outside" shares which the trading community as a whole must hold in equilibrium. This creates risk which the community as a whole cannot avoid. The trading community must be compensated to hold this risk. This effect creates a risk premium which discounts the equilibrium stock price.

Randomness in the net supply of these "outside" shares is introduced in much of the asymmetric information literature in order to prevent common knowledge and price conditioning from drying up volume in equilibrium (See Sargent (1993)).

If changes in the net supply of outside shares is positively correlated then the LeBaron effect (v) can be explained within the context of the Brock LeBaron model. Here is why. Near-past volatility increases when near-past market precision increases. When market precision is infinite, autocorrelation in outside share supply has zero effect on autocorrelation of price change. This is so because the depressing effect upon equilibrium price caused by these outside shares is caused by the risk that the community must bear in holding these outside shares. But when market precision is infinite this risk is zero.

Autocorrelation of near-future price changes with current price changes is a ratio of covariance to the product of standard deviations. A rise in market precision increases the standard deviations for the reasons we gave above. The covariance would be zero because of fact (i). But it is positive in the Brock LeBaron model when the covariance in net supply of outside shares is positive and the market precision is finite. If the market precision increases this covariance is decreased for the reasons given above. We have an explanation for fact (v) within the context of this model. It remains to be seen whether this corresponds to any reason found in reality. However, de Fontnouvelle (1995) is able to produce the LeBaron (1992) effect in his more realistic model.

Let us discuss fact (i). In the Brock and LeBaron (1995) model, observed market price is a predictor of the fundamental. Hence price differences represent differences in predictors which makes it fairly easy for the model to reproduce the stylized fact (i) provided that the fundamental is a random walk. de Fontnouvelle's model (1995) can do a better job of reproducing this feature because the intertemporal forces that act to produce low autocorrelation at higher frequencies are better captured by his model.

Finally let us discuss the last fact (vi). Brock and LeBaron briefly discuss embedding their model in the general asset pricing framework with social interactions developed by Brock (1993). This framework grafts social interactions in the choice decision of whether to buy more precise information onto conventional asset pricing models and generates asset pricing formulae that can display abrupt changes in equilibrium asset values provided the social interactions are strong enough. This is due to the interactions causing a breakdown of the cross sectional central limit theorem as the large economy limit is taken.

In the Brock and LeBaron model all that is needed for the breakdown of the cross sectional central limit theorem, in the large economy limit, is that the product of the intensity of choice with the strength of the social interactions be large enough. In other words high intensity of choice, i.e., a lot of "rationality" can combine with a small amount of "sociology" to produce large responses to small changes in the environment. If the intensity of choice is parameterized as a function of the difference in profit measures from buying the signal versus not buying the signal, then this kind of model can not only endogenize "jumps" in market data but also lead to "phases" in the market where volatility and "excess returns" differ. In high precision phases volatility is high because the market is "tracking" well, but "excess" returns are not high because very little risk is being borne by holding the outside shares. This can be viewed as an integration of Vaga's (1994) "Coherent Market Hypothesis" with more conventional asset pricing theories.

This kind of modelling can produce behavior which looks more like the "Markov switching" models of Hamilton and Susmel (1994) which are discussed in Section three. The different regimes correspond to the different phases when most traders are well informed and when most traders are poorly informed. The social interactions magnify the coherence of traders decisions so that the trading group acts more like a "clump" rather than a group of independent random variables. This clumping can generate behavior that looks more like Markov switching. Section three shows how Markov switching models can produce "spurious" long term dependence in volatility.

Of course, we do not wish to imply that social interactions are the only realistic forces that may produce abrupt changes in market data. See Jacklin, Kleidon, and Pfleiderer (1992) for a discussion of the role of other forces such as portfolio insurance, stale prices, trading institutions, etc., in producing abrupt changes such as the October crashes.

In this section we have discussed very recent work on calibration of "structural" models to reproduce common features of financial data at relatively high frequencies. Furthermore these kinds of models appear tractable enough to estimate on returns and volume data with computer intensive methods, like those of Duffie and Singleton (1993). It may be possible to use bootstrap-based specification tests along the lines discussed in this section to judge the models, provided that advances in computer technology continue to drive computation costs down. Specification tests based upon quantities of direct financial interest like trading profits may give us better information than conventional specification testing on how to fix the model if it is rejected by the specification test. Turn now to some brief closing remarks.

5 Concluding remarks

This article has given a highly selective survey of some recent work in finance. The survey has given a brief discussion of: (i) "complexity theory" and its possible role in generating "fat tailed" returns data in finance, (ii) phenomena by frequency, (iii) nonlinearity testing, (iv) testing for long memory, (v) cautions raised by moment condition failure of popular tests, (vi) problems raised by testing for existence of moments, (vii) bootstrap-based specification testing based upon quantities of interest in finance such as trading profits, (viii) some recent efforts in asymmetric information structural modeling with calibration.

In view of the challenges posed to conventional analytics by this type of work, we believe that future progress will make use of computer intensive methods such as Judd and Bernardo (1993), Judd (1994), and Rust (1994). Computer intensive methods will allow a closer dialogue between features of the data, structural modeling, and specification testing which uses financially relevant quantities such as trading profits.

References

Abhyankar, A., L. Copeland, and W. Wang (1995), "Nonlinear dynamics in real-time equity market indices: Evidence from the UK," *Economic Journal* (forthcoming).

Abhyankar, A., (1994), "Linear and Nonlinear Granger Causality: Evidence from the $FT-SE100$ Index Futures and Cash Markets," Department of Accountacy and Finance, University of Stirling, Scotland.

Abu-Mostafa, Y., Chm., (1994), *Proceedings of Neural Networks in the Capital Markets: NNCM '94*, California Institute of Technology.

Agiakloglou, C., P. Newbold, and M. Woahr (1994), "Lagrange multiplier tests for fractional difference," *Journal of Time Series Analysis* **15**, 253-262.

Akgiray, V. and G. C. Booth (1988), "The stable-law model of stock returns," *Journal of Business and Economic Statistics* **6**, 51-57.

Akgiray, V. and C. Lamoureux (1989), "Estimation of stable parameters: a comparative study," *Journal of Business and Economic Statistics* **7**, 85-93.

Altug, S., and P. Labadie (1994), *Dynamic Choice and Asset Markets*, New York: Academic Press.

Andersen, T. (1995),"Return volatility and trading volume: an information flow intepretation of stochastic volatility," Department of Finance, Kellogg School of Management, Northwestern University W.P. #170.

Andersen, T., and T. Bollerslev (1994),"Intraday Seasonality and Volatility Persistence in Foreign Exchange and Equity Markets," Department of Finance, Kellogg School of Management, Northwestern University, W.P. #186.

Andrews, D. (1991), "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica* **59**, 817-858.

Anis, A. and E. Lloyd (1976),"The expected values of the adjusted rescaled Hurst range of independent normal summands," *Biometrika* **63**, 111-116.

Antoniewicz, R. (1992),"A Causal Relationship Between Stock Returns and Volume," Board of Governors, Federal Reserve System, Washington, D.C.

Antoniewicz, R. (1993), "Relative Volume and Subsequent Stock Price Movements," Board of Governors, Federal Reserve System, Washington, D.C.

Arthur, B., J. Holland, B. LeBaron, R. Palmer, and P. Tayler (1993), "Artificial Economic Life: A simple model of a stockmarket," Santa Fe Institute Working Paper.

Aydogan, K. and G. Booth (1988), "Are there long cycles in common stock returns?," *Southern Economic Journal* **55**, 141-149.

Auestad, B. and D. Tj́ostheim (1990),"Identification of nonlinear time series: 1st order characterization and order determination," *Biometrika* **77**, 669-687.

Baek, E. and W. Brock (1992a), "A general test for nonlinear Granger causality: Bivariate model," mimeo, Department of Economics, University of Wisconsin-Madison.

Baek, E. and W. Brock (1992b),"A nonparametric test for temporal dependence in a vector of time series, *Statistica Sinica* **2**, 137-156.

Baillie, R. (1995), "Long memory processes and fractional integration in Econometrics," *Journal of Econometrics* (forthcoming).

Baillie, R., T. Bollerslev, and H. Mikkelsen, (1993),"Fractionally integrated generalized autoregressive conditional heteroskedasticity," Working Paper No. 168, Department of Finance, Northwestern University.

Bak, R., and D. Chen (1991), "Self-Organized Criticality," *Scientific American*, January.

Barnett, W., R. Gallant, M. Hinich, J. Jungeilges, D. Kaplan and M. Jensen (1994),"A single-blind controlled competition between tests for nonlinearity and chaos," Working Paper No. 190, Department of Economics, Washington University in St. Louis.

Beran, J. (1992),"A goodness of fit test for time series with slowly decaying serial correlations," *Journal of the Royal Statistical Society, Series B* **54**, 749-760.

Beran, J. (1993), "Fitting long-memory models by generalized linear regression," *Biometrika* **80**, 817-822.

Beran, J. (1994), *Statistics for Long-Memory Processes*, Chapman and Hall, New York.

Bierens, H. (1990),"A Consistent Conditional Moment Test of Functional Form," *Econometrica* **58**, 1443-1458.

Blattberg, R. C. and N. J. Gonedes (1974),"A comparison of the stable and student distributions as statistical models for stock prices," *Journal of Business* **47**, 244-280.

Bollerslev, T. (1986), "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics* **31**, 307-327.

Bollerslev, T., R. Engle, and D. Nelson (1994),"ARCH models," in R. Engle and D. McFadden, eds., *The Handbook of Econometrics*, Vol. IV, North-Holland, Amsterdam.

Bollerslev, T. and Mikkelsen, H. (1993),"Modeling and pricing long-memory in stock market volatility," Working Paper No. 134, Department of Finance, Northwestern University.

Bollerslev, T. and J. Wooldridge (1992),"Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances," *Econometric Reviews* **11**, 143-172.

Boothe, P. and D. Glassman (1987),"The statistical distribution of exchange rates: empirical evidence and economic implications," *Journal of International Economics* **22**, 297-320.

Bradley, R. and R. McClelland (1994a),"An improved nonparametric test for misspecification of functional form," mimeo, Bureau of Labor Statistics.

Bradley, R. and R. McClelland (1994b), "A kernel test for neglected nonlinearity," mimeo, Bureau of Labor Statitics.

Breidt, J., N. Crato, and P. de Lima (1994),"Modeling long memory stochastic volatility," Working Papers in Economics No. 323, Department of Economics, The Johns Hopkins University.

Brock, W., (1982),"Asset Prices in a Production Economy," in *The Economics of Uncertainty*, ed. by J.J. McCall, Chicago: University of Chicago Press.

Brock, W., (1993), "Pathways to Randomness in the Economy: Emergent Nonlinearity and Chaos in Economics and Finance," *Estudios Economicos* **8**, El Colegio de Mexico, Enero-junio, 3-55.

Brock, W. A., W. D. Dechert, and J. Scheinkman (1987), "A test for independence based on the correlation dimension," Department of Economics, University of Wisconsin, University of Houston and University of Chicago. (Revised Version, 1991: Brock, W. A., W. D. Dechert, J. Scheinkman, and B. D. LeBaron), *Econometric Reviews* (forthcoming).

Brock, W., D. Hsieh, and B. LeBaron (1991), *A Test of Nonlinear Dynamics, Chaos and Instability: Theory and Evidence*, M.I.T Press, Cambridge.

Brock, W., and A. Kleidon (1992),"Periodic Market Closure and Trading Volume: A Model of Intraday Bids and Asks," *Journal of Economic Dynamics and Control* **16**, 451-489.

Brock, W., J. Lakonishok, and B. LeBaron (1992), "Simple Technical Trading Rules and the Stochastic Properties of Stock Returns," *Journal of Finance* **47**, 1731-1764.

Brock, W., and B. LeBaron (1995),"A dynamic structural model for stock return volatility and trading volume," *Review of Economics and Statistics* (forthcoming), NBER W.P. #4988.

Brock, W. A. and S. M. Potter (1993), "Nonlinear time series and macroeconometrics," in G. Maddala, C. Rao, and H. Vinod, eds., *Handbook of Statistics Volume 11: Econometrics*, North Holland, New York.

Brockwell, P. and R. Davis (1991), *Time Series: Theory and Models*, Springer-Verlag, New York.

Cai, J. (1994),"A Markov model of unconditional variance in ARCH," *Journal of Business and Economic Statistics* **12**, 309-316.

Campbell, J., S. Grossman, and J. Wang (1993),"Trading Volume and Serial Correlation in Serial Returns," *Quarterly Journal of Economics* **108**, 905-939.

Campbell, J., A. Lo, and C. MacKinlay (1993), *The Econometrics of Financial Markets*, Princeton University Press (forthcoming).

Cheung, Y. (1993a), "Tests for fractional integration: a Monte Carlo investigation," *Journal of Time Series Analysis* **14**, 331-345.

Cheung, Y. (1993b),"Long memory in foreign exchange rates," *Journal of Business and Economic Statistics* **11**, 93-101.

Cheung, Y. and K. Lai (1993), "Do gold markets have long-memory?," *The Financial Review* **28**, 181-202.

Cheung, Y., K. Lai and M. Lai (1993),"Are there long cycles in foreign stock returns?," *Journal of International Financial Markets, Institutions and Money* **3**, 33-47.

Clark, P. (1973), "A subordinated stochastic process model with finite variance for speculative prices," *Econometrica* **41**, 135-155.

Cochrane, J. (1988),"How big is the random walk in GNP?," *Journal of Political Economy* **96**, 893-920.

Crato, N. (1994),"Some international evidence regarding the stochastic memory of stock returns," *Applied Financial Economics* **4**, 33-39.

Crato, N. and P. de Lima (1994),"Long range dependence in the conditional variance of stock returns," *Economics Letters* **45**, 281-285.

Creedy, J., and V. Martin, (1994) *Chaos and Non-linear Models in Economics: Theory and Applications*, Brookfield, Vermont: Edward Elgar.

Davies, R. and D. Harte (1987), "Tests for Hurst effect," *Biometrika* **74**, 95-102.

de Fontnouvelle, P. (1995), *Three Models of Stock Trading*, PhD Thesis, Department of Economics, The University of Wisconsin, Madison.

de Haan, L., S. Resnik, H. Rootzen and C. de Vries (1989),"Extremal behavior of solutions to a stochastic difference equation with applications to ARCH-processes," *Stochastic Processes and their Applications* **32**, 213-224.

de Jong, R. (1992),"The Bierens test under data dependence," mimeo, Free University of Amsterdam.

de Jong, R. and H. Bierens (1994),"On the limit behavior of a chi-square type test if the number of conditional moments tested approaches infinity," *Econometric Theory* **9**, 70-90.

de Lima, P. (1994b),"On the robustness of nonlinearity tests to moment condition failure," Working Papers in Economics No. 336, Department of Economics, The Johns Hopkins University.

de Lima, P. (1994c),"Nonlinearities and nonstationarities in stock returns," mimeo, Department of Economics, The Johns Hopkins University.

de Lima, P. (1995), "Nuisance parameter free properties of correlation integral based statistics," *Econometric Reviews* (forthcoming).

de Vries, C. (1991),"On the relation between GARCH and stable processes," *Journal of Econometrics* **48**, 313-324.

Dechert, W. D. (1988),"A Characterization of Independence for a Gaussian Process in Terms of The Correlation Integral," University of Wisconsin SSRI W.P. 8812.

Denker, M., and G. Keller (1983),"On U-Statistics and v. Mises' Statistics for Weakly Dependent Processes," *Zeitschrift f́ur Wahrscheinlichkeitstheorie und Verwandte Gebiete* **64**, 617-656.

Diebold, F. (1986), "Modeling the persistence of conditional variances: comment," *Econometric Reviews* **5**, 51-56.

Diebold, F. and J. Lopez (1995),"Modeling volatility dynamics," in K. Hoover, ed., *Macroeconometrics: Developments, Tensions and Prospects*, Kluwer Publishing Co.

Ding, Z., C. Granger, and R. Engle (1993),"A long memory property of stock market returns and a new model.," *Journal of Empirical Finance* **1**, 83-106.

Duffie, D., and K. Singleton (1993),"Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica* **61**, 929-952.

DuMouchel, W. (1983),"Estimating the stable index α in order to measure the tail thickness: a critique," *Annals of Statistics* **11**, 1019-1031.

Efron, B., and R. Tibshirani (1986),"Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy," *Statistical Science* **1**, 54-77.

Ellis, R. (1985), *Entropy, Large Deviations and Statistical Mechanics*, New York, Springer-Verlag.

Engle, R. F. (1982), "Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation," *Econometrica* **50**, 987-1007.

Engle, R. and T. Bollerslev (1986),"Modelling the persistence of conditional variances," *Econometric Reviews* **5**, 1-50.

Fama, E. and K. French (1988),"Permanent and Temporary Components of Stock Prices," *Journal of Political Economy* **96**, 246-273.

Fielitz B. and J. Rozelle (1983),"Stable distributions and the mixtures of distributions hypothesis for common stock returns," *Journal of the American Statistical Association* **78**, 28-36.

Friedman, D., and J. Rust, eds., (1993), *The Double Auction Market: Institutions, Theories, and Evidence*, Addison-Wesley, Redwood City, California.

Friggit, J. (1995),"Statistical Mechanics of Evolutive Financial Markets: Application to Short Term FOREX Dynamics," Essec Business School, near Paris, France.

Gallant, R., P. Rossi, and G. Tauchen (1992), "Stock Prices and Volume," *Review of Financial Studies* **5**, 199-242.

Gallant, R., P. Rossi, and G. Tauchen (1993),"Nonlinear dynamic structures," *Econometrica* **61**, 871-907.

Geweke, J. and S. Porter-Hudak (1983),"The estimation and application of long memory time series models," *Journal of Time Series Analysis* **4**, 221-238.

Glosten, L., R. Jagannathan and D. Runkle (1994), "Reltionship between the expected value and the volatility of the nominal excess return on stocks," *Journal of Finance* **48**, 1779-1802.

Goetzman, W. (1993),"Patterns in three centuries of stock market prices," *Journal of Business* **66**, 249-270.

Goldfeld, S. and R. Quandt (1973),"A Markov model for switching regressions," *Journal of Econometrics* **1**, 3-15.

Goodhart, C., and M. O'Hara (1995),"High Frequency Data in Financial Markets: Issues and Applications," London School of Economics and Johnson Graduate School of Management, Cornell University.

Granger, C. (1980),"Long memory relationships and the aggregation of dynamic models," *Journal of Econometrics* **14**, 227-238.

Granger, C. and R. Joyeux (1980),"An introduction to long-range time series models and fractional differencing," *Journal of Time Series Analysis* **1**, 15-30.

Granger, C. and J. Lin (1994),"Using the mutual information coefficient to identify lags in nonlinear models," *Journal of Time Series Analysis* **15**, 371-384.

Granger, C. and T. Teŕasvirta, (1993), *Modeling Nonlinear Economic Relationships*, Oxford University Press, Oxford.

Greene M. and B. Fielitz (1977), "Long-term dependence in common stock returns," *Journal of Financial Economics* **4**, 339-349.

Grossman, S. (1989), *The Informational Role of Prices*, Cambridge, MA.: MIT Press.

Guillaume, D., M. Dacorogna, R. Dave', U. Muller, R. Olsen, and O. Pictet (1994),"From the Bird's Eye to the Microscope: A Survey of new stylized facts of the intra-daily foreign exchange markets," Olsen and Associates, Zurich, Switzerland.

Hall, P. (1982), "On some simple estimates of an exponent of regular variation," *Journal of the Royal Statistical Society* **44**, 37-42.

Hall, P. (1994),"Methodology and Theory for the Bootstrap," in R. Engle and D. McFadden, eds., *The Handbook of Econometrics*, Vol. IV, North-Holland, Amsterdam.

Hamilton, J. D. and R. Susmel (1994), "Autoregressive conditional heteroskedasticity and changes in regime," *Journal of Econometrics* **64**, 307-333.

Harvey, A.C. (1993),"Long memory in stochastic volatility," Mimeo, London School of Economics.

Hiemstra, C. and J. Jones (1994a),"Testing for linear and nonlinear Granger causality in the stock-volume retlation," *Journal of Finance* **49**, 1639-1664.

Hiemstra, C. and J. Jones (1994b),"Another look at long memory in common stock returns," Discussion Paper 94/077, University of Strathclyde.

Hiemstra, C., and C. Kramer (1994), "Nonlinearity and Endogeneity in Macro-Asset Pricing," Department of Finance, University of Strathclyde, Scotland.

Hill, B. (1975),"A simple general approach to inference about the tail of a distribution," *Annals of Mathematical Statistics* **3**, 1163-1174.

Hinich, M. (1982),"Testing for gaussianity and linearity of a stationary time series," *Journal of Time Series Analysis* **3**, 443-451.

Hinich, M. and D. Patterson (1985), "Evidence of nonlinearity in stock returns," *Journal of Business and Economic Statistics* **3**, 69-77.

Hodges, S. (1995), "Arbitrage in a Fractal Brownian Motion Market," Financial Options Research Centre, University of Warwick, March.

Horgan, J. (1995), "From Complexity to Perplexity: Can science achieve a unified theory of complex systems? Even at the Santa Fe Institute, some researchers have their doubts," *Scientific American* **276**, 104-109.

Horowitz, J. (1995), "Lecture notes on bootstrap," Lecture notes prepared for World Congress of the Econometric Society, Tokyo, Japan, 1995.

Hosking, J. (1981), "Fractional differencing," *Biometrika* **68**, 165-176.

Hsieh, D. A. (1991), "Chaos and nonlinear dynamics: application to financial markets," *The Journal of Finance* **46**, 1839-1877.

Hsu, D., R. Miller, and D. Wichern (1974), "On the stable paretian behavior of stock-market prices," *Journal of the American Statistical Association* **69**, 108-113.

Hurst, H. (1951), "Long-term storage capacity of reservoirs," *Transactions of the American Socitey of Civil Engineers* **116**, 770-799.

Inclán, C. (1993), "GARCH or sudden changes in variance? An empirical study," mimeo, Georgetown University.

Jacklin, C. A. Kleidon, and P. Pfleiderer (1992), "Underestimation of Portfolio Insurance and the Crash of October 1987," *Review of Financial Studies* **5**, 35-63.

Jaditz, T., and C. Sayers, (1993), "Is chaos generic in economic data?" *International Journal of Bifurcations and Chaos*, 745-755.

Jansen, D. and C. de Vries (1991), "On the frequency of large stock returns: putting booms and busts into perspective," *Review of Economics and Statistics* **73**, 18-24.

Jensen, M. (1994), "Wavelet analysis of fractionally integrated processes," mimeo, Department of Economics, Washington University.

Jeong, J., and G. Maddala (1993), "A Perspective on Application of Bootstrap Methods in Econometrics," in G. Maddala, C. Rao, H. Vinod, eds., *Handbook of Statistics 11: Econometrics*, 573-610.

Jog, V., and H. Schaller (1994), "Finance constraints and asset pricing: evidence on mean reversion," *Journal of Empirical Finance* **1**, 193-209.

Judd, K., (1994), *Numerical Methods in Economics*, forthcoming book, Hoover Institute.

Judd, K., A. Bernardo (1993), "Asset market equilibrium with general securities, tastes, returns, and information, asymmetries," working paper, Hoover Institution.

Kim, M., C. Nelson and R. Startz (1991),"Mean reversion in stock prices? A reapraisal of the empirical evidence," *Review of Economic Studies* **58**, 515-528.

Koedijk, K., M. Schafgans and C. de Vries (1990),"The tail index of exchange rate returns," *Journal of International Economics* **29**, 93-108.

Kramer, C. (1994),"Macroeconomic Seasonality and the January Effect," *Journal of Finance* **49**, 1883-1891.

Krugman, P. (1993),"Complexity and Emergent Structure in the International Economy," Department of Economics, Stanford University.

Lamoureux, C. and W. Lastrapes (1990),"Persistence in variance, structural change, and the GARCH model, *Journal of Business and Economic Statistics* **8**, 225-234.

Lamoureux, C., and W. Lastrapes (1994), "Endogenous Trading Volume and Momentum in Stock Return Volatility," *Journal of Business and Economic Statistics* **12**, 253-260.

LeBaron, B. (1992), Some relations between volatility and serial correlations in stock returns," *Journal of Business* **65**, 199-219.

LeBaron, B., (1993), *Emergent Structures: a Newsletter of the Economics Research Program at the Santa Fe Institute*, July, 1993.

LeBaron, B., (1994),"Chaos and nonlinear forecastiblity in economics and finance," *Philosophical Transactions of the Royal Society of London, Series A* **348**, 397-404.

Lee, T., H. White, and C. Granger (1993),"Testing for neglected nonlinearity in time series models. a comparison of neural network methods and alternative tests," *Journal of Econometrics* **56**, 269-290

Leger, C., D. Politis, and J. Romano (1992), "Bootstrap Technology and Applications," *Technometrics* **34**, 378-398.

LePage, R., and L. Billard (1992), *Exploring the Limits of Bootstrap*, John Wiley and Sons: New York.

Levich, R., and L. Thomas (1993),"The significance of technical trading-rule profits in the foreign exchange market: A bootstrap approach," *Journal of International Money and Finance* **12**, 451-474.

Lo, A. (1991),"Long-term memory in stock market prices," *Econometrica* **59**, 1279-1313.

Lo, A. and C. MacKinlay (1988),"Stock markets do not follow random walks: evidence from a simple specification test," *Review of Financial Studies* **1**, 41-66.

Loretan, M. (1991), Testing covariance stationarity of heavy-tailed economic time series, Ch. 3, Ph. D. Dissertation, Yale University.

Loretan, M. and P. C. B. Phillips, (1993), "Testing the covariance stationarity of heavy-tailed time series: an overview of the theory with applications to several financial datasets," *Journal of Empirical Finance* **1**, 211-248.

Lucas, R. (1978), "Asset Prices in an Exchange Economy," *Econometrica* **46**, 1429-1445.

Luukkonen, R., P. Saikkonen, and T. Ter´asvirta (1988), "Testing linearity against smooth transition autoregressions," *Biometrika* **75**, 491-499.

Mandelbrot, B. (1963), "The variation of certain speculative prices," *Journal of Business* **36**, 394-419.

Mandelbrot, B. (1971), "When can price be arbitraged efficiently? A limit to the validity of the random walk and martingale models," *Review of Economics and Statistics* **53**, 543-553.

Mandelbrot, B. (1975), "Limit theorems of the self-normalized range for weakly and strongly dependent processes," *Zeitschrift f´ur Wahrscheinlichkeitstheorie und Verwandte Gebiete* **31**, 271-285.

Mandelbrot, B. and M. Taqqu (1979), "Robust R/S analysis of long run serial correlation," 42nd session of the International Statistical Institute, Manila, Book 2, 69-99.

Mandelbrot, B. and J. Wallis (1968), "Noah, Joseph, and operational hydrology," *Water Resources Research* **4**, 967-988.

McLeod, A. and W. Li (1983), "Diagnostic checking ARMA time series models using squared-residual autocorrelations," *Journal of Time Series Analysis* **4**, 269-273.

McFadden, D. (1989), "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration," *Econometrica* **57**, 995-1026.

Mehra, R. (1991), "On the Volatility of Stock Market Prices," Working Paper, Department of Economics, The University of California, Santa Barbara, *Journal of Empirical Finance* (forthcoming).

Michener, R. (1984), "Permanent Income in General Equilibrium," *Journal of Monetary Economics* **14**, 297-305.

Mills, T. (1993), "Is there long-term memory in UK stock returns?," *Applied Financial Economics* **3**, 293-302.

Mittnik, S. and S. Rachev (1993), "Modeling asset returns with altrenative stable distributions," *Econometric Reviews* **12**, 261-330.

Nelson, D. B. (1990), "Stationarity and persistence in the GARCH(1,1) model," *Econometric Theory* **6**, 318-334.

Nelson, D. B. (1991), "Conditional heteroskedasticity in asset returns: a new approach," *Econometrica* **59**, 347-370.

Pagan, A. (1995), "The econometrics of financial markets," mimeo, The Australian National University and The University of Rochester.

Pagan, A. and G. Schwert (1990), "Testing for Covariance Stationarity in Stock Market Data," *Economics Letters* **33**, 165-170.

Pakes, A. and D. Pollard (1989), "Simulation and Asymptotics of Optimization Estimators," *Econometrica* **57**, 1027-1057.

Peters, E. (1994), *Fractal Market Analysis*, John Wiley & Sons, New York.

Poterba, J. and L. Summers (1988), "Mean reversion in stock returns: evidence and implications," *Journal of Financial Economics* **22**, 27-60.

Prigogine, I., and M. Sanglier (1987), eds., *Laws of Nature and Human Conduct: Specificties and Unifying Themes,* published by G.O.R.D.E.S. Task Force of Research Information and Study on Science, Bruxelles, Belgium.

Priestly, M. (1988), *Non-linear and Non-stationary Time Series Analysis*, Academic Press, New York.

Ramsey, J. B. (1969), "Tests for specification errors in classical linear least-squares regression analysis," *Journal of the Royal Statistical Society* **31**, 350-371.

Randles, R. (1982), "On the asymptotic normality of statistics with estimated parameters," *Annals of Statistics* **10**, 462-474.

Richardson, M. (1993), "Temporary components of stock prices: a skeptic's View," *Journal of Business and Economic Statistics* **11**, 199-207.

Robinson, P. (1983), "Nonparametric estimators for time series," *Journal of Time Series Analysis* **4**, 185-207.

Robinson, P. (1991a), "Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression," *Journal of Econometrics* **47**, 67-84.

Robinson, P., (1991b), "Consistent nonparametric entropy-based testing," *Review of Economic Studies* **58**, 437-453.

Robinson, P. (1993), "Log-periodogram regression for time series with long range-dependence," mimeo, London School of Economics.

Rosen, S., K. Murphy, and J. Scheinkman (1994), "Cattle Cycles," *Journal of Political Economy* **102**, 468-492.

Rust, J. (1994), "Structural Estimation of Markov Decision Processes," in R. Engle and D. McFadden, eds., *The Handbook of Econometrics*, Vol. IV, North-Holland, Amsterdam.

Saikkonen, P. and R. Luukkonen (1988), "Lagrange multiplier tests for testing nonlinearities in time series models," *Scandinavian Journal of Statistics* **15**, 55-58.

Samorodnitsky, G. and M. Taqqu (1994), *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Chapman and Hall, New York.

Sargent, T. (1993), *Bounded Rationality in Macroeconomics*, Oxford: Clarendon Press.

Sargent, T. (1995), "Adaptation of Macro Theory to Rational Expectations," Working Paper, Department of Economics, University of Chicago and Hoover Institution.

Savit, R. and M. Green (1991), "Time series and dependent variables," *Physica D* **50**, 521-544.

Scheinkman, J. (1992), "Stock Returns and Nonlinearities," in P. Newman, M. Milgate, and J. Eatwell, *The New Palgrave Dictionary of Money and Finance*, London: MacMillan, 591-593.

Scheinkman, J., and M. Woodford, (1994), "Self-Organized Criticality and Economic Fluctuations," *American Economic Review Papers and Proceedings*, May, 417-421.

Simonato, J. G. (1992), "Estimation of GARCH processes in the presence of structural change,", *Economics Letters* **40**, 155-158.

Singleton, K. (1990), "Specification and Estimation of Intertemporal Asset Pricing Models," in B. Friedman and F. Hahn, eds., *Handbook of Monetary Economics: I*, North Holland, Amsterdam.

Smith, R., Chm. (1990), *Market Volatility and Investor Confidence: Report of the Board of Directors of the New York Stock Exchange, Inc*, New York Stock Exchange: New York.

Subba Rao, T. and M. Gabr (1980), "A test for linearity of stationary time series," *Journal of Time Series Analysis* **1**, 145-158.

Summers, L. (1986), "Does the stock market rationally reflect fundamental values?," *The Journal of Finance* **44**, 1115-1153.

Taylor, S. (1994), "Modeling stochastic volatilty: a review and comparative study," *Mathematical Finance* **4**, 183-204.

Thursby, J. G. and P. Schmidt (1977), "Some properties of tests for specification error in a linear regression model," *Journal of the American Statistical Association* **72**, 635-641.

Tjöstheim, D. and B. Auestad (1994), "Nonparametric identification of nonlinear time series: selecting significant lags," *Journal of the American Statistical Association* **89**, 1410-1419.

Tsay, R. (1986), "Nonlinearity tests for time series," *Biometrika* **73**, 461-466.

Vaga, T. (1994), *Profiting from Chaos: Using Chaos Theory for Market Timing, Stock Selection, and Option Valuation*, New York: McGraw-Hill.

Viano, M., C. Deniau and G. Oppenheim (1994), "Continuous-time fractional ARMA processes," *Statistics & Probability Letters* **21**, 323-336.

Vinod, H. (1993), "Bootstrap Methods: Applications in Econometrics," in G. Maddala, C. Rao, H. Vinod, eds., *Handbook of Statistics 11: Econometrics*, 629-661.

Wallis, J. and N. Matalas (1970), "Small sample properties of H and K, estimators of the Hurst coefficient h," *Water Resources Research* **6**, 332.

Wang, J. (1993), "A model of intertemporal asset prices under asymmetric information," *Review of Economic Studies* **6**, 405-434.

Wang, J. (1994),"A model of comeptitive stock trading volume," *Journal of Political Economy* **102**, 127-168.

Weidlich, W. (1991), "Physics and Social Science: The Approach of Synergetics," *Physics Reports* **204**, 1-163.

West, K., H. Edison, D. Cho (1993),"A Utility-based comparison of some models of exchange rate volatility," *Journal of International Economics* **35**, 23-45.

White, H. (1987),"Specification testing in dynamic models," in Bewley T., ed., Advances in Econometrics, Fifth World Congress, Volume 1, Cambridge University Press, Cambridge.

White, H. and J. Wooldridge (1991),"Some results on sieve estimation with dependent observations," in W. Barnett, J. Powel and G. Tauchen, eds., *Semiparametric and Nonparametric Methods in Economics and Statistics*, Cambridge University Press, New York.

Wooldridge, J. (1992),"A test for functional form against nonparametric alternatives," *Econometric Theory* **8**, 452-475.

Wu, P. (1992),"Testing fractionally integrated time series," mimeo, Victoria University of Wellington.

Wu, K., R. Savit, and W. Brock (1993),"Statistical tests for deterministic effects in broad band time series," *Physica D* **69**, 172-188.

Zolatarev, V. (1986), *One-dimensional Stable Distributions*, Vol. 65 of "Translations of mathematical monographs," American Mathematical Society. Translation from the original 1983 Russian edition.