The Search for Chaos and Nonlinearities in Swedish Stock Index Returns

Byström, Hans; Amilon, Henrik

1998

The Search for Chaos and Nonlinearities in Swedish Stock Index Returns

Henrik Amilon and Hans NE Byström*
Department of Economics
Lund University
P.O. Box 7082
S-220 07 Lund
Sweden
phone: +46-46-222 79 09
E-mail: Henrik.Amilon@nek.lu.se
E-mail: Hans.Bystrom@nek.lu.se

Abstract

Numerous empirical studies have shown evidence of nonlinearities in financial time series, which can be of both a deterministic and a stochastic nature. Chaos is an example of the former, and heteroscedasticity in the conditional variance an example of the latter. We apply a test, the BDS test, to Swedish Stock Index returns and detect large deviations from the IID-hypothesis. There is no evidence of chaos, and most of the nonlinearities are due to conditionally heteroscedastic error terms. We look at monthly, daily, and 15-minute return series, and find no sensitivity in the results to choice of sampling frequency. Different GARCH models often seem to explain the nonlinearities detected by the BDS test, which is particularly the case for GARCH models with t-distributed errors fitted to monthly and daily returns.

*The authors are grateful to Björn Hansson and Lars Nordén for providing the data, and to David Edgerton for useful comments on statistical issues.

1 Introduction

Common assumptions in financial economics are that financial variables like stock returns and exchange rates can be described by stochastic processes, and that economic systems are linear. From a theoretical point of view, nothing indicates that this must be the case, and numerous empirical studies have shown that different nonlinear dependences might exist in financial time series. The high frequency of crashes and booms that can be observed in stock markets is an example of behavior non-consistent with a linear model[1] with normally distributed disturbances.

Nonlinear models explaining this type of behavior can be both stochastic and deterministic. An example of a stochastic nonlinear model is Engle's (1982) autoregressive conditionally heteroscedastic (ARCH) model with extensions. These models have been successful in explaining volatility clustering in stock returns and in modeling interest rates. A chaotic process, on the other hand, is perfectly deterministic but its behavior is indistinguishable from pure randomness. Chaotic systems are always nonlinear and have been important in explaining seemingly random behavior in the sciences. A very important fact regarding chaotic models, which is different from the usual situation in economics, is that all movements are generated within the model. No external shocks need to be introduced.

The lack of explanative power in linear models has led numerous researchers to the study of chaos in financial economics, see e.g. Scheinkman and LeBaron (1989), Peters (1991), Larrain (1991), and Hsieh (1991). Hsieh (1989) instead focuses on nonlinear stochastic models, in particular GARCH models. Varson and Doran (1995) try to distinguish chaos from random nonlinearities with the Grassberger-Procaccia correlation dimension. The disadvantages of this method are that very long data series are needed, and that it does not constitute a statistic test. Therefore, Brock, Dechert, and Scheinkman (1987) proposed a related statistic test, the BDS test, based on the correlation integral, a measure of spatial correlation, and it can be used to detect deviations from the IID-hypotheses. Since the BDS test does not distinguish between different causes for rejections of IID, Hsieh (1991) first filters data with different linear and nonlinear filters in order to detect what kind of model can explain the nonlinearities.

In this chapter, we look at Swedish stock index returns, and try to detect the presence of nonlinearities in this market. Can nonlinear effects, for instance, help explaining the large jumps in stock prices that occur with fairly high frequency? Our data covers monthly returns from 1919 to 1996, daily data from 1977 to 1996, and intradaily data (15 minutes) from January 1992

[1] By a linear model, we mean a model with a linear mean and additive IID error terms.

to August 1993.

Since the movements cannot be successfully described by linear models, we try to determine how to model the nonlinearities that must be introduced. Typically, nonlinearities can enter in the process governing returns (nonlinearities in the conditional mean) or in the process of the time-varying conditional variance (nonlinearities in the conditional variance). In order to have a chaotic returns process, there must exist nonlinearities in the conditional mean. To test this hypothesis, we investigate the predictability of linear versus nonlinear models fitted to data by using neural networks. The weak evidence of nonlinearities in the conditional mean makes us continue our search for causes of the non-IID stock returns. We turn to the BDS test, and try to separate the different causes of a rejection of the IID-hypothesis in our market; chaos, heteroscedastic conditional variance, nonstationarity, or simple linear autocorrelation. We reject the hypothesis of conditional mean changes or nonstationarities causing the BDS test to reject the IID hypothesis. Instead, we find strong evidence of conditional variance dependences in the stock index returns. We do not only find significant ARCH-effects but also the stronger result that ARCH-effects alone contribute to more or less all the non-IID behavior in our data.

In section 2, we give an introduction to chaos, nonlinearities, and the BDS test, in section 3 our data is described, in section 4 we present our results, and section 5 concludes the chapter.

2 Chaos and How to Detect it

What is chaos? A clear cut definition has not yet been generally accepted, but three fundamental properties must be included (see Strogatz (1996)):

- Chaotic motion must have an aperiodic long-term behavior. This means that there are trajectories[2] which do not settle down to fixed points, periodic orbits, or quasiperiodic orbits as $t \to \infty$.
- Chaotic motion exhibits sensitive dependence on initial conditions, which means that nearby trajectories diverge exponentially fast.
- Chaotic motion is purely deterministic. The irregular behavior stems from the nonlinearity of the system and not from random effects.

A dynamic system evolving in a chaotic way is also said to have a *strange attractor*. It is possible to quantify the sensitivity on the initial conditions above by defining the *Liapunov*

[2] The phase space is the space spanned by the state variables. A trajectory is the path of evolution of a point in the phase space.

exponents for chaotic systems in continuous time (flows), and discrete time (maps). For a flow in n dimensions, an infinitesimal sphere will be distorted into an infinitesimal ellipsoid during its evolution in phase space. If $\epsilon_k(t)$, $k = 1, ..., n$ denotes the length of the kth principal axes of the ellipsoid, then $\epsilon_k(t) \propto \epsilon_k(0)\, e^{\lambda_k t}$, where the λ_k's are the Liapunov exponents. The n different Liapunov exponents therefore describe the deformation of the system. For a strange attractor, at least one of the λ_k's must be positive because of the separation of neighboring trajectories. The sum of the λ_k's describes the contraction of volume, which must be negative since, by definition, an attractor (strange or otherwise) should attract all trajectories starting in a sufficiently small open set containing the attractor. Chaotic behavior demands a positive Liapunov exponent, which implies the existence of a time horizon beyond which prediction breaks down, since any discrepancy in the estimate of an initial state will grow exponentially.

From where does chaos come? In the "definition" of chaos, two properties seem hard to combine. How can trajectories on a strange attractor diverge endlessly and yet remain bounded? The answer is that strange attractors result from a stretching and folding process. To be more concrete, consider again the sphere in the phase space. A strange attractor is generated when the system contracts the sphere in some directions and stretches it in others. To remain bounded, the distorted sphere must be folded back onto itself. After a large number of iterations, the sphere is spread throughout a bounded region in the phase plane. The mechanisms involved reflect the volume contraction, the sensitivity on initial conditions and the boundedness of the attractor.

2.1 The Correlation Integral

How do we test for the presence of chaos? One way would be to calculate the largest Liapunov exponent and see if it is positive, but this is not easily done in real world situations. Moreover, the other Liapunov exponents are even harder to estimate so this procedure cannot distinguish between different types of strange attractors. Grassberger and Procaccia (1983) therefore proposed a different procedure based on spatial correlation.

Consider a set $\{\mathbf{x}_t\}_1^T$ of points on an attractor. Define the *correlation integral* as $C(l) = \lim_{T\to\infty} C^T(l)$,

$$C^T(l) = \frac{1}{T^2}\sum_{s,t}\theta(l - |\mathbf{x}_t - \mathbf{x}_s|),$$

where $\theta(a) = 1$ if $a > 0$, and 0 otherwise. The quantity $C(l)$ measures the fraction of the total number of pairs $(\mathbf{x}_t, \mathbf{x}_s)$ whose distance $|\mathbf{x}_t - \mathbf{x}_s|$ is less than l. Grassberger and Procaccia found that $C(l)$, and its sample estimate $C^T(l)$, are proportional to l^d where d is called the *correlation dimension.* In practice, d is estimated as the slope in a plot of $\ln C^T(l)$ versus $\ln(l)$.

When dealing with finite data sets, the power law, $C^T(l) \propto l^d$, only holds over an intermediate range of l, since at large l, all points will be within a distance of l, while on the other hand, at extremely small l, no pair of points will be within l. Typically, strange attractors have a fractal structure and therefore, a non-integer correlation dimension.

2.2 Attractor Reconstruction

A time series, $\{\mathbf{x}_t\}$, of all state variables $\mathbf{x} \in R^n$ is often not accessible. In many cases, it is not even possible to specify either the relevant components of $\mathbf{x}$ or its dimension, which might be high. Nevertheless, such systems might have low-dimensional attractors. Fortunately, Packard et al. (1980) and Takens (1981) show that the dynamics in the full phase space can be reconstructed from measurements of a single-variable time series $\{\{x_t\}; x_t \in R, x \subset \mathbf{x}\}$. The main idea is to construct m-dimensional vectors

$$\boldsymbol{\xi}_t = (x_t, x_{t+\tau}, ..., x_{t+(m-1)\tau})$$

for some delay $\tau > 0$. If $\{\mathbf{x}_t\}$ possesses an attractor that can be embedded in an n-dimensional $\mathbf{x}$-space, then the topological structure of the attractor remains unchanged when embedded in $\boldsymbol{\xi}$-space, provided that the *embedding dimension* m is large enough. A necessary condition is $m \geq n$ and a sufficient condition is $m \geq 2n+1$. In other words, it is possible to make some kind of "variable substitution" between the unobservable variables in $\mathbf{x}$ and the lagged observables.

Since $\{\mathbf{x}_t\}$ and $\{\boldsymbol{\xi}_t\}$ are observationally equivalent, they also have the same correlation dimension, provided that m is large enough. Thus, $C_m^T(l)$ can be calculated, now with a subscript m, using $\boldsymbol{\xi}_t$ instead of $\mathbf{x}_t$. If $\{\boldsymbol{\xi}_t\}$ is obtained from a chaotic time series, then the computed correlation dimension will level off at its true value when the embedding dimension is large enough, so that there is enough room for the attractor to unfold. The point to be made is that if $\{\boldsymbol{\xi}_t\}$ is obtained from a purely random (IID) sequence, the correlation dimension keeps increasing with m, since noise always fills up the space in which it is embedded. In principle, estimates of d can thus be used to distinguish chaos from noise.

The delay τ should not be chosen too small since then $x_t \approx x_{t+\tau} \approx ...$, and the attractor might have problems in disentangling. If, on the other hand, τ is chosen too large, problems arise because the spatial correlation is low for distant values in the time series.

Unfortunately, the Grassberger-Procaccia method has some drawbacks. First, it must be emphasized that the method breaks down when the embedding dimension is too large due to the sparsity of data, which causes statistical sampling problems. Second, Grassberger and Procaccia based their results on time series of 10000-30000 points, although they argued that only a few thousand points would be necessary to obtain reasonable estimates. Ramsey and Yuan (1989) have shown that there is a tendency to underestimate the slope in data sets with as many as 2000 points, thus indicating chaos when none is present. Third, the graphical procedure is not put on firm statistical ground and it might be difficult to interpret when noise is added to the system.

2.3 Influence of Noise

In real life, systems can seldom be isolated from the unpredictable environment. The observed dynamics will then consist of two parts; the intrinsic deterministic dynamics of the system and the influence of random noise. One common way of modeling the randomness is as *dynamical noise*:

$$x_t = f(x_{t-1}, ...) + \varepsilon_t, \tag{1}$$

where ε_t is IID, x_t is the observable, and f is a deterministic, possibly nonlinear, function. If f is chaotic, much of what has been said above about chaos is still valid when noise is added to the system, *provided the noise level is not too high.* Otherwise, the trajectories could jump out of the basin of attraction, or jump between different parts of the attractor in such a way that it becomes impossible to separate the behavior from true noise; the attractor has lost its structure. An acceptable noise level is hard to determine since different chaotic systems exhibit very different sensitivities to noise, but generally, only small amounts of noise can be introduced.

The Grassberger-Procaccia method must be modified when dealing with noise (see Ben-Mizrachi et al. (1984)). For length scales, l, below those where the random component blurs the structure, the slope of $\ln C_m^T(l)$ versus $\ln(l)$ is proportional to the embedding dimension, while for length scales above, the slope is equal to the correlation dimension of the deterministic system ($C_m^T(l) \propto l^d$).

Equation (1) can also be used to describe different stochastic processes. If, for instance, f is a nonlinear (or linear) map with a stable attracting fix point, then the output corresponds to passing random noise through a nonlinear (linear) recursive filter, thereby obtaining stationary aperiodic behavior. In principle, the Grassberger-Procaccia method could be used to identify the properties of f by calculating the slope at different l and m, and comparing it to the slope of an IID process of the same length and with the same moments, i.e. the time series *randomly permuted* or *scrambled.*

Figure 1 shows $\ln(C(l))$ versus $\ln(l)$ for the Hénon (1976) map

$$x_{k+1} = 1 + y_k - a x_k^2, \tag{2}$$

$$y_{k+1} = b x_k, \tag{3}$$

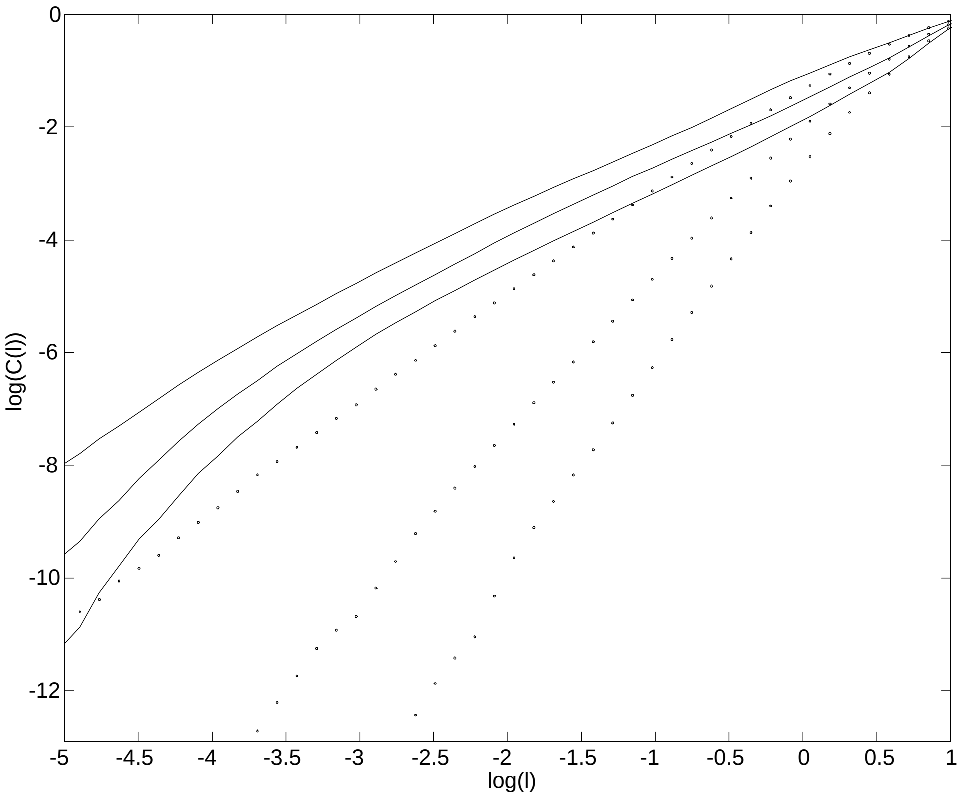

Figure 1: Plot of $\ln(C(l))$ vs. $\ln(l)$ for the Hénon map with $a = 1.4$ and $b = 0.3$ (thick lines) and its scrambled counterpart (dotted), for $m = 2, 3$, and 4. Higher lying lines correspond to lower m. Noise uniformly distributed on $(-0.03, 0.03)$ is dynamically added to the x -component. The changes in slopes for $l < 0.06$ ($\ln(l) < -2.8$) are clearly visible. The time series is normalized to unit sample variance, and consists of 2000 observations.

for $m = 2, 3$, and 4 with uniformly distributed noise dynamically added. In this case the attractor can be completely reconstructed when $m = 3$.[3] From Figure 1, it is clear that the slopes are different for length scales less than the magnitude of the imposed noise. Since the noise is uniformly distributed in $(-0.03, 0.03)$, the slopes for $l < 0.06$, $\ln(l) < -2.8$) are equal to the corresponding embedding dimensions. As expected, the correlation dimensions of the scrambled time series show no saturation as m grows.

This graphical procedure should, in principle, also be valid for residuals from ARCH-type

[3] The slope for $m = 2$ is only slightly less than the slopes in higher embedding dimensions, thereby suggesting that the attractor has a non-trivial structure already when $m = 2$.

models[4], where $\varepsilon_t = g(x_{t-1}, ...)u_t$, u_t IID, but it would be hard to tell if any deviation from the scrambled counterpart is caused by the conditional mean process f or the conditional variance process g or both (see Scheinkman and LeBaron (1989)). Brock, Dechert, and Scheinkman (1987) modified the Grassberger-Procaccia method in order to circumvent some of its drawbacks by developing the BDS test. A revised version of this test is found in Brock et al. (1996).

2.4 The BDS Test

The test is based on the observation that for an IID sample

$$C_m^T(l) = [C_1^T(l)]^m.$$

The identity should be understood in a statistical sense. Brock, Dechert, and Scheinkman (BDS) derived a normalization factor[5] $V_m^T(l)$ in order to make a correct statistical quantification of the departure from IID. More specifically, they showed that the BDS statistic

$$W_m^T(l) = \sqrt{T}\frac{\left[C_m^T(l) - [C_1^T(l)]^m\right]}{V_m^T(l)}$$

converges in the distribution to $N(0,1)$ as $T \to \infty$, for $l > 0$ and $m > 1$, under the null hypothesis of IID.

The BDS test has proved to be quite powerful in finding departures from IID in a number of Monte Carlo simulations; see, for example, Brock, Hsieh, and LeBaron (1991) and Hsieh (1991). It has also been verified that the finite sample distribution well approximates the asymptotic distribution for sample sizes above 1000 observations. As for the Grassberger-Procaccia method, the results depend on the magnitude of the additive noise component. It is therefore recommended (see Brock, Hsieh, and LeBaron (1991)) to apply the test for $l = 0.5 - 1.5$ times the sample standard deviation of the time series.

The BDS test suffers from the fact that a rejection of the null does not provide any hints about the cause of the rejection. It has been shown in Brock and Potter (1993) and de Lima (1996), however, that the null distribution of the test is not affected by applying the BDS test to the estimated residuals from a general class of parametric models with additive IID errors:

$$x_t = f(x_{t-1}, ...; \boldsymbol{\alpha}) + u_t,$$

[4] By ARCH-type models, we mean models that exhibit autoregressive conditional heteroscedasticity.

[5] The normalization factor is a complicated function of correlation integrals in different dimensions. It is not very illuminating and therefore not presented, but can be found in Brock et al. (1996).

where $\boldsymbol{\alpha}$ is a parameter vector and u_t is IID, *provided* that a $\sqrt{T}$-consistency estimation of the parameters is possible. The last requirement means that $\sqrt{T}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) \rightarrow N(\mathbf{0}, \boldsymbol{\Sigma})$ for some covariance matrix $\boldsymbol{\Sigma}$, and is fulfilled in maximum likelihood estimations. Moreover, de Lima (1996) shows that the nuisance parameter free properties of the BDS test remain valid for residuals from some multiplicative models

$$x_t = f(x_{t-1}, ...; \boldsymbol{\alpha}) + g(x_{t-1}, ...; \boldsymbol{\beta})u_t, \tag{4}$$

where u_t is IID, which covers many, if not all, of the known ARCH-type models, *if* the test is applied to the transformed residuals $\hat{v}_t = \ln(\hat{u}_t^2)$ where $\hat{u}_t$ are the estimated residuals in (4). The key idea is that the transformation gives rise to a model with additive IID errors:

$$\tilde{x}_t = \ln\left((x_t - f(x_{t-1}, ...; \boldsymbol{\alpha}))^2\right) = \ln\left(g^2(x_{t-1}, ...; \boldsymbol{\beta})\right) + \ln(u_t^2) = \tilde{f}(x_{t-1}, ...; \boldsymbol{\beta}) + v_t,$$

where v_t is IID because u_t is IID. It must be kept in mind that the asymptotic properties of $\hat{v}_t$ and v_t are equal only if $\boldsymbol{\alpha}$ is known, i.e. when a GARCH model is fitted to the data without the conditional mean process. However, the bias introduced is expected to be small when $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are jointly estimated "...because of the lack of predictability of the first moment compared to the second central moment..." (see Brock and Potter (1993)), as is often the case in financial time series[6].

Furthermore, the BDS test does not require any moments for the time series studied to exist, contrary to other nonlinearity tests, but when the test is applied to model residuals $\sqrt{T}$-consistency of the parameter estimates sometimes demand the existence of higher moments in the disturbance process, typically a finite variance. The assumption of the existence of higher order moments might otherwise be severe in the nonlinearity testing of financial time series, whose distributions are often heavily tailed. The disadvantage is that the BDS test may require longer samples to be as effective as some other tests.

Altogether, the BDS test may not only be useful in detecting chaos, but also in diagnostic testing of estimated model residuals.

3 The Data

The data we use is log-returns from two different Swedish stock indices, the OMX-Index and the Affärsvärldens Generalindex. The former was created as an underlying security for trading in standardized stock index options and forward contracts, and consists of a value-weighted

[6] With the possibility of bootstrapping small sample distributions, the asymptotic properties of the test are nowadays of less importance.

Table 1: Sample statistics for return data.

	Monthly	Daily	15 min. 1	15 min. 2	15 min. 3	15 min. all
No. of observations	936	2949	2579	2571	2767	7917
Mean ($\cdot 10^2$)	0.49	0.023	-0.0015	0.0072	0.0088	0.0049
Variance ($\cdot 10^4$)	21.61	0.30	0.03	0.13	0.036	0.06
Skewness	-0.52	-0.18	1.05	4.53	0.27	4.36
Kurtosis	8.04	9.30	32.72	96.43	9.58	130.11

combination of the 30 most traded stocks at the Stockholm Stock Exchange. The latter is the most used stock index in Sweden and it is a value-weighted index of the majority of all stocks quoted on the Stockholm Stock Exchange.

The time interval between observations is sometimes of importance in financial time series. Therefore, we look at series with different sampling frequency; monthly, daily, and 15 minutes.[7] In this way, we can detect chaotic and nonlinear behavior on different time scales. In addition, the problem of too short or too long delay times τ in the reconstruction of the chaotic attractor mentioned in section 2.2 can be mitigated.

The monthly data (Affärsvärldens Generalindex) is extended over the period 1919-1996, while the daily data (OMX-Index) is for the period 1984-1996. The intradaily series (OMX-Index), with data collected every 15 minutes, covers the period January 1992-August 1993 and is divided into three series with approximately equal length to catch any instability in the estimated parameters.[8] At the same time, the length of the series becomes comparable to the length of the daily series. The first observation every day has been removed due to the auction-like start up procedure at the exchange. All data is without dividends, but we have also looked at monthly (1919-1996) and daily (1977-1991) returns with dividends included (Affärsvärldens Generalindex).[9] The results for these series are not reported but are similar to those without dividends and any differences will be commented upon. Table 1 displays some sample statistics of the return data. The high degree of kurtosis and skewness, in particular for the 15 minute data is observed.

[7] The 15-minute time span between data points should be large enough that possible micromarket structure dependences do not shine through. Such dependences can, for instance, arise from the sequential execution of orders in the traders' limit order book, or by thin trading in some of the underlying stocks.

[8] The second subperiod covers the volatile months when Sweden abandoned the fixed exchange rate regime.

[9] For monthly returns, the series with dividends has 936 observations and for the daily returns, 3489 observations.

Table 2: BDS statistics for raw data. l is the length scale, σ is the sample standard deviation, and m is the embedding dimension. The test statistic is asymptotically distributed as $N(0,1)$.

	$\frac{l}{\sigma}$	m = 2	3	4	5	6	7	8	9	10
monthly	0.5	8.33	10.83	13.57	16.61	20.13	26.49	33.56	48.21	67.21
	1	8.43	10.74	13.39	15.72	17.81	20.54	23.64	27.34	32.20
	1.5	7.70	9.83	12.04	13.79	15.20	16.73	18.24	19.76	21.49
daily	0.5	12.36	14.66	17.55	21.05	24.82	30.05	36.82	46.94	59.21
	1	12.53	14.40	16.35	18.53	20.49	22.82	25.29	28.17	31.87
	1.5	13.25	15.12	16.73	18.13	19.21	20.34	21.47	22.56	23.86
15 min. 1	0.5	11.43	13.13	12.82	13.58	13.81	13.81	13.67	13.67	13.00
	1	12.76	13.53	12.88	12.61	12.25	11.73	11.13	10.42	9.51
	1.5	11.80	11.79	10.89	10.19	9.48	8.82	8.21	7.48	6.72
15 min. 2	0.5	16.71	19.11	20.87	22.47	24.03	25.38	26.34	27.23	27.37
	1	14.34	15.67	15.78	15.74	15.32	14.78	14.01	13.25	12.36
	1.5	9.51	10.01	9.67	9.19	8.53	7.85	7.16	6.48	5.81
15 min. 3	0.5	8.89	11.16	11.93	13.39	15.15	16.79	18.13	19.29	20.77
	1	8.36	10.59	11.23	12.40	13.48	14.27	14.79	14.95	15.14
	1.5	7.49	8.98	9.35	10.11	10.72	11.00	11.05	10.85	10.54
15 min. all	0.5	24.98	29.87	32.63	36.31	40.49	44.78	49.12	53.58	58.02
	1	24.09	6.89	27.75	28.68	29.28	29.57	29.65	29.52	29.21
	1.5	20.12	21.57	21.39	21.29	20.92	20.40	19.79	19.13	18.43

4 Empirical Results

In the following section, we apply the BDS test to the raw data to test if the different time series are IID. If the null hypothesis of IID is rejected (and it is), we try to discover the cause of the rejection by applying the BDS test to the estimated residuals from different models. More precisely, we examine whether the rejection origins from nonlinear (possibly chaotic) or linear dependences in the mean processes, nonstationarities, or if the rejection can be explained by nonlinear stochastic models exhibiting conditional heteroscedasticity.

4.1 Test of the Raw Data

Table 2 shows the BDS statistics when the embedding dimension is between 2 and 10, and l is chosen as 0.5, 1 and 1.5 times the sample standard deviation of the different time series. As can be observed, the test strongly rejects the hypothesis of IID stock returns at any conventional significance level and for all m and l, even though little can be said about the cause of the rejection. If the time series came from a noisy chaotic system with a higher dimensional attractor, then the test statistic would typically be small for low embedding dimensions and large in dimensions with enough "room" for the attractor. Obviously, this is not the case since the rejections are significant already at $m = 2$.

Considering the results, we proceed to study how dependences in the conditional mean might cause the rejection of IID.

4.2 Test of Dependences in the Conditional Mean

Suppose stock returns are generated as:

$$x_t = f(x_{t-1}, ...) + \varepsilon_t \tag{5}$$

where ε_t is IID and f is the conditional mean process. The expression in (5) can be used to describe noisy chaotic models as well as different stochastic processes. Now, we will investigate whether f is nonlinear (and possibly chaotic), and whether any dependences in the conditional mean can explain the rejection of IID.

If f is nonlinear, it can be modeled by nonparametric regressions, e.g. artificial neural networks. This powerful technique is exhaustively described elsewhere (see Hertz et al. (1991) and Kuan and White (1994)), and thus, we will here merely say that a multi layer perceptron (MLP), with one hidden layer of sigmoid transfer functions, can consistently model a time series generated by an arbitrary function, provided that the function is bounded and uniformly continuous. On the other hand, if the MLP only contains linear transfer functions, it can only pick up linear dependences, i.e. the nonparametric modeling reduces to a linear regression. We therefore proceed in the following way:

We split each time series into three parts of unequal size. The first 50% are used for the regression, or training, while the consecutive 25% are used as validation sets. The last 25% ($t = 1, ..., T_{test}$) are used as test sets for out-of-sample evaluation. For each returns series, we train 10 MLP's with 3, 5 and 7 hidden neurons, each with sigmoid (tanh) transfer functions, in a single hidden layer and a linear output neuron. As inputs we use $x_{t-1}, ..., x_{t-9}$ corresponding to an embedding dimension $m = 10$. For each time series, we also train linear perceptrons to capture the linear dependences. In all cases, the sum squared errors of the validation sets are minimized, and for each time series, we pick the network with the best validation performance. The different MLPs are then evaluated using the out-of-sample prediction error normalized with

Table 3: Out-of-sample evaluation for nonlinear and linear neural networks. MLP3 refers to a MLP with 3 hidden neurons etc. MLPL refers to a linear perceptron.

	E_p			
	MLP3	MLP5	MLP7	MLPL
monthly	1.0512	1.0418	1.1635	1.0357
daily	0.9913	1.0132	1.0046	1.0108
15 min. 1	0.9903	0.9871	1.0127	0.9852
15 min. 2	0.9942	1.0452	1.0015	0.9943
15 min. 3	0.9926	0.9869	1.0029	0.9820
15 min. all	0.9901	0.9925	0.9909	0.9895

the out-of-sample variance of x_t:

$$E_p = \frac{\frac{1}{T_{test}} \sum_{t=1}^{T_{test}} \left(x_t - \hat{f}\left(x_{t-1}, ..., x_{t-9}\right)\right)^2}{\hat{\sigma}_{test}^2}.$$

If f is nonlinear, E_p should be smaller for nonlinear than for linear networks. Furthermore, if f is chaotic, then the performance of nonlinear networks should improve as the forecast horizon shortens, that is, when we go from a sample interval of one month down to 15 minutes, and it is a consequence of the positive Liapunov exponent(s) in a chaotic system.

The results are reported in Table 3. Three things can be noted. First, the out-of-sample prediction performance is very weak, with E_p slightly below one (and occasionally above). Second, there is no support in the data for nonlinear regression being superior to linear regression. Sometimes, it is superior but then the improvement is negligible. Third, the performance of the nonlinear networks does not improve as the sample interval is reduced.

We also analyze the dependence of the output on the input variables by inspecting derivatives after completed regression. This is done by calculating

$$S_k(t) = \frac{\partial \hat{f}(x_{t-1}, ..., x_{t-9})}{\partial x_{t-k}}, \qquad k = 1, ..., 9$$

for $t = 1, ..., T_{test}$ and then computing

$$S_k = \frac{1}{T_{test}} \sum_t \mid S_k(t) \mid . \tag{6}$$

The measure in (6) shows which variables are the most important. For a linear network, this is just equal to (the absolute values of) the coefficients from a linear regression.

The results are found in Table 4. We see that both types of networks identify approximately the same variables as being important, often the first. The magnitudes of S_k differ somewhat

Table 4: Sensitivity dependences in the nonparametric regressions. MLP refers to a nonlinear multi layer perceptron. MLPL refers to a linear perceptron.

		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9
monthly	MLP	0.1153	0.1208	0.0360	0.0929	0.0409	0.1091	0.0263	0.0515	0.1228
	MLPL	0.1634	0.1131	0.0123	0.0230	0.0001	0.1681	0.0129	0.1049	0.1217
daily	MLP	0.1217	0.0171	0.0462	0.0249	0.0107	0.0574	0.0475	0.0325	0.0242
	MLPL	0.1843	0.0124	0.0569	0.0106	0.0210	0.0659	0.0898	0.0780	0.0281
15 min. 1	MLP	0.0537	0.0507	0.0511	0.0381	0.0191	0.0046	0.0206	0.0183	0.0105
	MLPL	0.0812	0.0365	0.0816	0.0189	0.0292	0.0069	0.0053	0.0111	0.0157
15 min. 2	MLP	0.0939	0.0499	0.0223	0.0023	0.0665	0.0572	0.0040	0.0147	0.0219
	MLPL	0.0714	0.0486	0.0298	0.0014	0.0105	0.0195	0.0381	0.0251	0.0224
15 min. 3	MLP	0.0341	0.0262	0.0102	0.0055	0.0134	0.0231	0.0068	0.0156	0.0056
	MLPL	0.0585	0.0095	0.0153	0.0270	0.0039	0.0003	0.0109	0.0062	0.0216
15 min. all	MLP	0.1330	0.0385	0.0091	0.0325	0.0070	0.0227	0.0132	0.0062	0.0303
	MLPL	0.1371	0.0393	0.0053	0.0348	0.0054	0.0162	0.0025	0.0181	0.0314

between linear and nonlinear networks, but the point is that there seem to be no substantial dependences in the data not captured by the linear networks.

Although this is not a statistical test of nonlinearities in the conditional mean process, we draw the conclusion that there are none. Either the dependences in f are truly linear, or they are so blurred with noise that the conditional mean can best be approximated by a linear process. Since chaos can only arise from a nonlinear mapping, this suggests that there is no chaotic behavior in the mean process.

Can the linear dependences in f cause the BDS test to reject IID? We try to answer this question by applying the BDS test to the residuals from the linear networks above. As appears from Table 5, the test strongly rejects the residuals as being IID at any reasonable significance level. The test statistics of the linearly filtered data do not differ a great deal from those of the raw data, indicating that the rejection is not only due to linear dependences[10].

In the following, we try to sort out what kind of nonlinearities the BDS test in fact detects in our data.

[10] We also applied the BDS test to the residuals from the nonlinear networks. As expected, the test statistics did not differ a great deal from those in Table 5, but to our knowledge, it is not clear whether the nuisance free properties of the test is valid for residuals from nonparametric regressions.

Table 5: BDS statistics for linearly filtered data. l is the length scale, σ is the sample standard deviation, and m is the embedding dimension. The test statistic is asymptotically distributed as $N(0,1)$.

		m								
	$\frac{l}{\sigma}$	2	3	4	5	6	7	8	9	10
monthly	0.5	7.28	10.09	12.59	15.04	17.70	22.11	26.18	31.35	36.66
	1	7.04	9.77	12.55	14.86	16.87	19.46	22.36	25.85	30.39
	1.5	6.50	8.97	11.30	13.19	14.65	16.22	17.73	19.26	20.97
daily	0.5	11.83	14.41	17.04	20.54	24.45	30.48	37.75	47.93	58.89
	1	12.18	14.27	16.16	18.52	20.69	23.26	26.03	29.21	33.28
	1.5	12.94	15.05	16.68	18.25	19.44	20.65	21.87	23.09	24.53
15 min. 1	0.5	10.58	12.21	11.98	12.60	12.90	12.77	12.48	12.10	11.14
	1	11.27	11.99	11.46	11.18	10.86	10.34	9.74	8.99	8.08
	1.5	10.46	10.34	9.64	8.97	8.32	7.69	7.12	6.41	5.69
15 min. 2	0.5	14.34	16.87	18.52	19.83	21.12	22.25	22.78	23.33	23.34
	1	14.27	15.45	15.53	15.51	15.16	14.67	13.95	13.19	12.27
	1.5	9.69	9.99	9.61	9.08	8.42	7.78	7.12	6.44	5.74
15 min. 3	0.5	7.99	10.21	10.80	12.02	13.49	15.04	16.65	17.92	19.78
	1	7.53	9.83	10.52	11.69	12.72	13.44	13.93	14.05	14.18
	1.5	6.67	8.37	8.87	9.76	10.42	10.73	10.77	10.56	10.23
15 min. all	0.5	22.42	27.25	30.09	33.43	37.12	40.88	44.46	48.19	52.13
	1	22.52	25.29	26.25	27.13	27.70	27.95	27.96	27.78	27.42
	1.5	19.56	20.82	20.62	20.49	20.12	19.62	19.01	18.34	17.59

4.3 Test of Nonstationarities

Financial time series are not necessarily stationary. There might be a number of structural changes creating nonstationary series: policy changes, changes in financial structures as well as important technological innovations. Therefore, it is possible that a rejection of IID by the BDS test is caused by nonstationarities and a closer study of whether this is the case is motivated. The type of nonstationarities we discuss in this section arise from switching between different linear models; that is, we allow the parameters of a linear model to change from regime to regime.

To investigate the influence of nonstationarity, we can either apply the BDS test to the residuals from a linear switching model, or we can try to judge it more qualitatively (following Hsieh (1991)). Since we have time series with different extensions in time, we do the latter. Our longest return series is 77 years and our shortest is about 7 months. In between, we have a time series of about 13 years. When we proceed to increasingly shorter time intervals, we expect the effect of structural changes to disappear (we simply assume that less and less changes occur) and consequently, any rejection of the IID hypothesis due to nonstationarities alone should disappear.

There is no support for this idea. The linearly filtered statistics in Table 5 are as high for the 15 minute data as for the monthly and daily data. As long as the linear regime shifts occur with a frequency low enough to leave our 15 minute data unaffected, they cannot be the only cause of the BDS test rejecting the IID hypothesis.

4.4 Test of Dependences in the Conditional Variance

As mentioned, there are many ways of generating non-IID data; we have treated linear and nonlinear autoregression, chaos, and nonstationarity. Now, we turn to models exhibiting nonlinearities in the conditional variance. Examples of models with this behavior are autoregressive conditionally heteroscedastic models like ARCH, GARCH, and their extensions.

Many time series exhibit periods of unusually large volatility followed by periods of tranquility. Under such circumstances, the assumption of a constant variance is obviously inappropriate. In order to detect the time-varying variance, the conditional variance can be modeled as a linear function of past squared errors. This is the ARCH model. Further extensions of the model, allowing the conditional variance to also be a function of its own lags, gives us Bollerslev's (1986) generalized ARCH model, GARCH(p, q), where p and q are the number of lagged conditional variance and squared error components, respectively. In the case of an AR(2) process with the conditional variance modeled as GARCH(1,1) we have:

$$\begin{aligned} x_t &= \alpha_0 + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \varepsilon_t \\ \sigma_t^2 &= \phi_0 + \phi_1 \varepsilon_{t-1}^2 + \phi_2 \sigma_{t-1}^2, \end{aligned}$$

where σ_t^2 is the conditional variance of ε_t, $\varepsilon_t = \sigma_t u_t$, and $u_t \sim N(0,1)$. Most empirical studies suggest that p and q larger than one are rarely needed. One problem with GARCH modeling is that the parameters are restricted; ϕ_0, ϕ_1, and ϕ_2 must all be larger than zero and the sum of $\phi_1$and ϕ_2 must be less than one. The sum $\phi_1 + \phi_2$ measures the persistence of conditional variance to shocks, which approaches infinity as the sum approaches one from below. The case where the sum of ϕ_1 and ϕ_2 equals one is referred to as Integrated in GARCH or IGARCH.

We proceed one more step, this time using an asymmetric specification for the conditional variance process, by extending the standard ARCH-model to Nelson's (1991) EGARCH-model:

$$\ln \sigma_t^2 = \phi_0 + \phi_1 \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \phi_2 \left(\left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| - E \left[\left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| \right] \right) + \theta \ln(\sigma_{t-1}^2)$$

where as above σ_t^2 is the conditional variance of ε_t and $\varepsilon_t \mid \varepsilon_{t-1} \sim N(0, \sigma_t)$. Unlike simple ARCH and GARCH, EGARCH can capture the asymmetric response of the variance to the direction of ε_t, that is, a higher variance when ε_t is negative, and a lower variance when ε_t is positive. In this model, $\phi_1 < 0$ represents the asymmetric effect while $\phi_2 > 0$ produces the ARCH effect. The θ-parameter determines the persistence in variance. Unlike the GARCH model, there are no non-negativity restrictions on the parameters in the EGARCH model.

A crucial assumption in the specifications above is the normality assumption on the standardized error term u_t; the conditional heteroscedasticity of ε_t alone is expected to explain the observed kurtosis in the return distribution. Since empirical evidence strongly rejects the idea that financial returns are normally distributed, we compare estimates with normally distributed errors with t-distributed ones. In this way, a larger part of the excess kurtosis in the stock returns might be captured.[11]

Baillie and DeGennaro (1990) as well as Poon and Taylor (1992) clearly demonstrate that the t-distribution gives a better fit to the error term of financial return series. For the t-distribution one additional parameter must be estimated, the degree of freedom, υ. The t-distribution is

[11] For the EGARCH model, one must be aware of the fact that the conditional variance expression changes with the distribution; $E\left[\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right|\right]$ depends on the chosen distribution. For the normal distribution $E\left[\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right|\right] = \sqrt{\frac{2}{\pi}}$, and for the t-distribution, our calculations give $E\left[\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right|\right] = \frac{\Gamma\left(\frac{\upsilon-1}{2}\right)}{\Gamma\left(\frac{\upsilon}{2}\right)} \frac{\sqrt{\upsilon}}{\sqrt{\pi}}$. For large values of υ (over 30), this expression is very close to $\sqrt{\frac{2}{\pi}}$.

Table 6: GARCH with normally distributed errors. The asterisk indicates that the null hypothesis of an IGARCH cannot be rejected. Small numbers indicate standard errors.

	$\alpha_0 \cdot 10^5$	α_1	α_2	$\phi_0 \cdot 10^5$	ϕ_1	ϕ_2	$\phi_D \cdot 10^5$
monthly*	0.01 0.002	0.16 0.035	-0.088 0.034	5.87 1.830	0.17 0.024	0.81 0.025	
daily	33.20 8.710	0.20 0.019		0.14 0.022	0.11 0.008	0.83 0.012	
15 min. 1	-1.36 2.23	0.21 0.021		0.07 0.0024	0.10 0.013	0.27 0.017	1.89 0.119
15 min. 2	4.55 3.640	0.26 0.019		0.19 0.0068	0.20 0.021	0.17 0.013	13.75 0.452
15 min. 3	4.43 2.993	0.14 0.022		0.12 0.0063	0.15 0.017	0.34 0.025	1.70 0.163
15 min. all	1.98 1.650	0.21 0.012		0.13 0.002	0.24 0.013	0.17 0.0073	6.35 0.099

symmetric, but has fatter tails and a higher kurtosis than the normal distribution. When υ goes to infinity, the t-distribution approaches a normal distribution.

In this section, we estimate the parameters in different GARCH models. EGARCH models have also been estimated but the results are not presented since these are very similar to those of the GARCH models. For each time series, a number of different GARCH and EGARCH models have been estimated, and the models with the best fit have been chosen, on basis of parsimony, likelihood value, and behavior of the standardized residuals.

For the t-distribution, one problem that has earlier been found for the 15-minute data is the low degrees of freedom (below 4), which means that fourth-order moments do not exist (Hansson and Hördahl (1994)). One possible explanation for these results is that the density for the 15 minute returns has very fat tails, driving the degrees of freedom of the t-distribution to low levels. In fact, these problems are solved by adding a daily overnight dummy to the 15-minute series in order to account for the increased variance in overnight returns.[12] Using a process like the following for the conditional variance in the GARCH model,

$$\sigma_t^2 = \phi_D D + \phi_0 + \phi_1 \varepsilon_{t-1}^2 + \phi_2 \sigma_{t-1}^2$$

where D is the dummy variable with parameter ϕ_D, we capture a considerable part of the kurtosis and get well-behaved residuals.[13]

The maximum likelihood estimates (the BHHH algorithm) of GARCH models with normally distributed, as well as t-distributed, errors are presented in Tables 6 and 7. For all our models and return series, we get significant parameter estimates, except for α_0 for the 15-minute returns.

[12] It would also be straightforward to include a weekend dummy but its impact should be smaller than the overnight dummy; the number of nontraded days each weekend is much smaller than the number of nontraded 15-minute periods overnight.

[13] In the EGARCH-case, a dummy is introduced in a similar manner; $\ln \sigma_t^2 = \ln(1 + \phi_D D) + \phi_0 + ...$

Table 7: GARCH with t-distributed errors. The asterisk indicates that the null hypothesis of an IGARCH cannot be rejected. Small numbers indicate standard errors.

	$\alpha_0 \cdot 10^5$	α_1	α_2	$\phi_0 \cdot 10^5$	ϕ_1	ϕ_2	$\phi_D \cdot 10^5$	υ
monthly*	0.01	0.16	-0.078	6.60	0.16	0.82		9.13
	0.002	0.035	0.035	2.56	0.031	0.033		2.580
daily	34.00	0.17	-0.037	0.09	0.11	0.86		7.28
	7.490	0.019	0.019	0.020	0.015	0.018		0.630
15 min. 1	-1.28	0.17		0.06	0.10	0.34	1.54	4.90
	1.910	0.019		0.005	0.024	0.033	0.190	0.455
15 min. 2	2.24	0.24		0.16	0.21	0.26	7.77	4.66
	3.246	0.019		0.013	0.034	0.027	0.925	0.438
15 min. 3	5.10	0.11		0.11	0.16	0.35	2.07	4.62
	2.656	0.019		0.011	0.031	0.042	0.158	0.438
15 min. all	1.35	0.17		0.10	0.25	0.31	3.07	4.02
	1.436	0.011		0.006	0.023	0.018	0.235	0.196

For the monthly data, we cannot reject the null of an infinite unconditional variance (we cannot reject the null of an IGARCH in these GARCH models) even though intuition suggests a mean-reverting variance. This conflicting evidence might be reconciled by allowing for fractional orders of integration (Bollerslev and Mikkelsen (1996)). Including AR(2) terms in the conditional mean gives better fit to data and less (linearly) correlated standardized residuals for some data series. In general, the coefficients of the AR(2) terms are very small, however. The estimates of the degrees of freedom parameter υ for the t-distribution are clearly finite, indicating non-normally distributed errors. The degrees of freedom parameters differ between time series but are similar for the GARCH and EGARCH models, respectively. Overall, the parameter estimates are fairly non-sensitive to the specification of the error distribution.

For our daily OMX-Index data, Hansson and Hördahl (1997) have estimated EGARCH models where the errors are described by the normal distribution as well as the generalized error distribution (GED). A comparison of our estimates with the results from Hansson and Hördahl (1997) shows almost identical estimates for ϕ_1, ϕ_2 and θ when the errors are assumed to be normal. The same holds for our parameters in the t-distributed model, when compared to the estimated GED models by Hansson and Hördahl.[14]

In practice, all models retain some skewness in the residuals, a finding common in empirical studies of conditionally heteroscedastic models. A comparison of the values of kurtosis between the models shows that, as expected, the t-distributed model has residuals with higher kurtosis than the normally distributed one. While the normally distributed standardized residuals show excess kurtosis for all data series, the t-distributed standardized residuals show a lower than theoretically predicted kurtosis of about the same size for all but the daily series.

[14] It is important to note that, as Nelson (1991) shows, EGARCH models with a t-distributed u_t give rise to strictly stationary although not covariance stationary σ_t^2 and ε_t; i.e. σ_t^2 and ε_t have no finite unconditional moments.

Table 8: BDS statistics for transformed GARCH residuals, normal distribution. l is the length scale, σ is the sample standard deviation, and m is the embedding dimension. The test statistic is asymptotically distributed as $N(0,1)$.

		m								
	$\frac{l}{\sigma}$	2	3	4	5	6	7	8	9	10
monthly	0.5	-1.13	-1.10	-0.98	-0.79	-0.43	0.04	0.35	-0.17	-2.24
	1	-0.69	-0.46	-0.28	-0.15	-0.13	-0.21	-0.20	-0.27	-0.13
	1.5	-0.49	-0.22	-0.08	-0.07	-0.06	-0.24	-0.29	-0.46	-0.46
daily	0.5	0.82	1.18	1.37	1.74	2.01	1.84	1.41	0.60	0.13
	1	0.50	0.65	0.80	1.29	1.38	1.29	1.09	0.97	0.92
	1.5	0.67	0.37	0.41	0.87	0.86	0.71	0.57	0.50	0.45
15 min. 1	0.5	0.19	-0.26	-0.48	-0.45	-0.25	-0.11	0.02	1.04	1.38
	1	-0.52	-0.44	-0.29	-0.08	0.21	0.55	0.91	1.28	1.46
	1.5	-0.89	-0.36	0.05	0.37	0.80	1.20	1.59	1.96	2.16
15 min. 2	0.5	1.30	1.47	1.38	1.59	1.69	1.61	0.80	-0.52	-2.69
	1	1.51	1.37	1.33	1.52	1.60	1.80	1.95	2.14	2.00
	1.5	1.83	1.86	1.89	1.97	2.06	2.22	2.31	2.43	2.39
15 min. 3	0.5	-0.44	-0.93	-1.04	-0.68	-0.71	-0.43	-0.87	-1.49	-2.94
	1	-0.27	-0.71	-0.98	-0.72	-0.57	-0.55	-0.59	-0.60	-0.59
	1.5	-0.39	-0.90	-0.97	-0.67	-0.49	-0.55	-0.59	-0.62	-0.61
15 min. all	0.5	0.40	0.03	0.01	0.45	0.70	0.90	0.95	0.83	0.84
	1	0.27	0.06	0.10	0.63	1.01	1.40	1.66	1.91	2.04
	1.5	0.38	0.30	0.56	1.03	1.43	1.80	2.11	2.38	2.51

4.4.1 BDS Statistics on Transformed Standardized Residuals

A comparison of BDS statistics for the linearly filtered financial data and for the GARCH residuals, the extent to which the rejection of IID residuals are due to ARCH-effects is captured. As discussed in section 2.4, we circumvent the problem in GARCH filtering related to the issue of nuisance parameters by looking at the transformed standardized residuals. In Table 8 and Table 9, we present the BDS statistics for the transformed standardized residuals from the fitted GARCH models in Table 6 and Table 7, respectively.

The most obvious result is the substantial decrease in rejecting frequency after GARCH filtering[15]. We cannot reject the null hypothesis for monthly and daily data with normal errors.

[15] The BDS test is asymptotically distributed as $N(0,1)$ and considering our long data series, we assume the

Table 9: BDS statistics for transformed GARCH residuals, t-distribution. l is the length scale, σ is the sample standard deviation, and m is the embedding dimension. The test statistic is asymptotically distributed as $N(0, 1)$.

		m								
	$\frac{l}{\sigma}$	2	3	4	5	6	7	8	9	10
monthly	0.5	-1.14	-1.27	-1.16	-0.94	-0.68	-0.38	-0.43	-0.81	-1.31
	1	-0.73	-0.49	-0.28	-0.12	-0.11	-0.20	-0.19	-0.25	-0.19
	1.5	-0.49	-0.20	-0.03	-0.04	-0.06	-0.26	-0.32	-0.51	-0.53
daily	0.5	0.72	1.12	1.22	1.47	1.76	1.57	0.93	0.28	0.27
	1	0.55	0.68	0.84	1.21	1.23	1.13	0.97	0.84	0.73
	1.5	0.72	0.55	0.65	1.04	0.94	0.79	0.65	0.60	0.59
15 min. 1	0.5	0.27	0.02	-0.07	-0.00	0.16	0.53	1.04	1.64	2.77
	1	-0.39	-0.38	-0.21	-0.04	0.30	0.54	0.77	0.99	1.13
	1.5	-0.76	-0.28	0.16	0.48	0.99	1.36	1.65	1.89	2.02
15 min. 2	0.5	1.27	1.26	1.18	1.21	1.20	0.72	-0.14	-0.75	-2.45
	1	1.38	1.19	1.27	1.48	1.64	1.86	2.04	2.22	2.18
	1.5	2.00	1.97	2.06	2.16	2.31	2.57	2.72	2.89	2.88
15 min. 3	0.5	-0.58	-0.93	-1.03	-0.82	-0.56	-0.28	-0.62	-0.80	-1.69
	1	-0.34	-0.74	-1.03	-0.83	-0.70	-0.68	-0.74	-0.80	-0.78
	1.5	-0.57	-0.97	-1.07	-0.85	-0.70	-0.75	-0.80	-0.85	-0.83
15 min. all	0.5	0.35	-0.03	-0.15	0.15	0.28	0.51	0.71	1.07	1.47
	1	0.32	0.02	-0.00	0.44	0.88	1.25	1.47	1.67	1.75
	1.5	0.64	0.40	0.55	0.93	1.37	1.72	1.93	2.13	2.19

Some test statistics are outside the often used 95% confidence intervals, but that is not unexpected with as many as 27 test statistics for each return series. We still reject IID returns for the 15-minute data in period 2 and the overall period. For the third period, one statistic is very large. We should reject IID for this period, but the asymptotic properties of the test might not be applicable for m as high as 10, with these sample sizes. A proper bootstrapping may clear these doubts, however.

For the t-distributed residuals, we have no rejection for monthly and daily data, and a slightly better fit of 15-minute returns, but we still have problems with period 2. It seems that the data

same holds in our finite samples. Any rejection mentioned is at the 5% significance level. For a two-sided test at this level, the critical values are ± 1.96.

series contain linear or nonlinear dependences that we have not successfully captured. Treating overnight effects with only one dummy might be too simple, at least for this period.

There might be another explanation, however. Table 8 and Table 9 clearly suggest that there are ARCH-effects in all our data, although the specification for period 2 is not correct. As mentioned above, this period includes a highly volatile part in the Swedish financial history when the fixed exchange rate regime was abandoned. In section 4.3, we concluded that *linear* regime shifts could not explain the rejection of IID in our returns, but we cannot exclude *nonlinear* regime shifts in the volatility. Such nonlinearities may be captured by the Markov switching model, SWARCH, where a Markov-chain governs the transition probabilities between the volatility states (Hamilton and Susmel (1994)). In the SWARCH model, the conditional variance is described by an ARCH model which is the same in the different regimes, but the scale of the conditional variance differs across regimes. Moreover, the regimes are solely identified from the data in the estimation procedure and involves no subjective classifications other than the number of volatility states.

It might be argued that if we have regime shifts in our 15-minute data, then the residuals from ordinary GARCH models for the monthly and daily data covering that period should not be IID. However, nonstationarities detected in one time scale are not necessarily detectable in another. For example, a shift to a higher volatility regime for some months would probably not be detectable in the analysis of monthly returns, but would most likely be detected when analyzing 15-minute returns. Some indications of nonstationarites are indeed observed in our parameters; in Tables 6 and 7 the variation in parameter estimates for the 15 minute return sample and its subsamples can be noted.

5 Summary and Conclusions

We have searched for nonlinearities and chaos in Swedish stock index returns and find strong evidence of nonlinear behavior, but no evidence of low-dimensional chaos. To our help, we use the BDS test to detect deviations from IID errors in time series. This test detects both deterministic and stochastic dependences, and lacks some of the disadvantages of the more commonly used Grassberger-Procaccia method. Applying this test to Swedish stock index returns of different sampling frequencies (monthly, daily, 15 minutes) strongly rejects the IID hypothesis. Since many possible dependences in the data, linear as well as nonlinear, may be the cause of this rejection, we apply the test to residuals from different models.

Linear dependences do not explain much of the deviation from IID. The same applies to nonstationarity; IID is not more strongly rejected for low frequency data. In order to detect chaos in stock indices, we compare the predictability of nonlinear and linear neural networks.

No significant improvement is found, giving low importance to nonlinearities in mean or chaotic dependences in the time series. However, this does not exclude the possibility of chaos in individual security returns, since a combination (like in a stock index) of chaotic time series might very well lose its chaotic structure, as examined by Atchison and White (1996).

Instead, heteroscedastic conditional variance models are found to explain much of the rejection of IID. The BDS statistics decrease substantially when GARCH or EGARCH models with normal or t-distributed errors are fitted to the time series. In particular, the t-distributed GARCH model explains the rejection of IID for monthly and daily data, while no model fully explains the nonlinearities found for the 15-minute data, but there is little doubt that the answer to a better fit lies in the specification of the conditional variance. If the rejections are caused by (not too many) regime shifts in the volatility process, these might be successfully captured by SWARCH models. We find our results quite robust and nonsensitive to the chosen type of GARCH model, and the inclusion of dividends or not. Our findings increase the strength of the already strong support for ARCH-type models of financial returns. For future research, we suggest searching for nonlinearities in other financial variables like implicit option volatilities, interest rates, and currencies. Analyzing nonlinear dependences both in time series and panel data might be of interest. The availability of transaction data also gives the possibility to study nonlinear effects in the micromarket structure.

References

Atchison, M., and White, M. (1996), "Disappearing Evidence of Chaos in Security Returns: A Simulation", *Quarterly Journal of Business and Economics*, 35(2), 21-37.

Baillie, R., and DeGennaro, R. (1990), "Stock Returns and Volatility", *Journal of Financial and Quantitative Analysis*, 25, 203-214.

Ben-Mizrachi, A., Grassberger, P., and Proccacia, I. (1984), "Characterization of Experimental (noisy) Strange Attractors", *Physical Review A* 29 (February), 975.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroscedasticity", *Journal of Econometrics*, 31, 307-327.

Bollerslev, T., and Mikkelsen, H. (1996), "Modeling and Pricing Long Memory in Stock Market Volatility", *Journal of Econometrics,* 73, 151-184.

Brock, W., Dechert, W., and Scheinkmann, J. (1987), "A Test for Independence based on the Correlation Dimension", University of Wisconsin at Madison, University of Houston, and

University of Chicago.

Brock, W., Dechert, W., Scheinkmann, J., and LeBaron, B. (1996), "A Test for Independence Based on the Correlation Dimension", *Econometric Reviews*, 15(3), 197-235.

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

Brock, W., and Potter, S. (1993), "Nonlinear Time Series and Macroeconometrics", in *Handbook of Statistics Volume 11: Econometrics*, ed. by Maddala, G., Rao, C., and Vinod, H., 195-229.

de Lima, P. (1996), "Nuisance Parameter Free Properties of Correlation Integral Based Statistics", *Econometric Reviews*, 15(3), 237-259.

Engle, R. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflations", *Econometrica,* 50, 987-1007.

Grassberger, P., and Proccacia, I. (1983), "Measuring the Strangeness of Strange Attractors", *Physica D,* 9, 189.

Hamilton, J., and Susmel, R. (1994), "Autoregressive Conditional Heteroscedasticity and Changes in Regime", *Journal of Econometrics,* 64, 307-333.

Hansson, B., and Hördahl, P. (1994), "Stock Returns and Conditional Volatility: An Empirical Investigation of the Swedish Stock Market", Working Paper, Department of Economics, Lund University.

Hansson, B., and Hördahl, P. (1997), "Forecasting Variance Using Stochastic Volatility and GARCH" in Hördahl (1997).

Hénon, M. (1976), "A Two-dimensional Mapping of a Strange Attractor", *Commun. Math. Phys.,* 50, 69.

Hertz, J., Krogh, A., and Palmer, R. (1991), *Introduction to the Theory of Neural Computation,* Addison-Wesley, Redwood City, Ca.

Hsieh, D. (1989), "Testing for Nonlinear Dependence in Daily Foreign Exchange Rates", *Journal of Business*, vol. 62, no. 3, 339-368.

Hsieh, D. (1991), "Chaos and Nonlinear Dynamics: Application to Financial Markets", *Journal of Finance*, vol. XLVI, no. 5, 1839-1877.

Hördahl, P. (1997), "Financial Volatility and Time-Varying Risk Premia", Ph.D.-Thesis, Department of Economics, Lund University.

Kuan, M. and White, H. (1994), "Artficial Neural Networks: An Econometric Perspective", *Econometric Reviews*, 13(1), 1-91.

Larrain, M. (1991), "Testing Chaos and Nonlinearities in T-Bill Rates", *Financial Analysts Journal*, September-October, 51-62.

Nelson, D. (1991), "Conditional Heteroscedasticity in Asset Returns: A New Approach, *Econometrica,* 59, 347-370.

Packard, N., Crutchfield, J., Farmer, J., and Shaw, R. (1980), "Geometry from a Time Series", *Phys. Rev. Lett.,* 45, 712.

Peters, E. (1991), "A Chaotic Attractor for the S&P 500", *Financial Analysts Journal*, March-April, 55-62.

Poon, S-H, and Taylor, S. (1992), "Stock Returns and Volatility: An Empirical Study of the UK Stock Market", *Journal of Banking and Finance*, 16, 37-59.

Ramsey, J., and Yuan, H. (1989), "Bias and Error Bias in Dimension Calculation and their Evaluation in Some Simple Models", *Physical Letters A,* 134, 287-297.

Scheinkmann, J., and LeBaron, B. (1989), "Nonlinear Dynamics and Stock Returns", *Journal of Business*, vol. 62, no. 3, 311-337.

Strogatz, S. (1996), *Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry, and Engineering,* Addison-Wesley, Reading, MA.

Takens, F. (1981), "Detecting Strange Attractors in Turbulence", *Lect. Notes in Math.,* 898, 396.

Varson, P., and Doran, P. (1995), "The Search for Evidence of Chaos in FTSE-100 Daily Returns", *European Financial Management*, vol. 1, no. 2, 201-210.