Nonlinearities and Nonstationarities in Stock Returns

Pedro J. F. de Lima


Stable URL:
http://links.jstor.org/sici?sici=0735-0015%28199804%2916%3A2%3C227%3ANANISR%3E2.0.CO%3B2-E

Nonlinearities and Nonstationarities in Stock Returns

Pedro J. F. DE LIMA
Department of Economics, The Johns Hopkins University, Baltimore, MD 21218 (de.lima@jhu.edu)

This article addresses the question of whether recent findings of nonlinearities in high-frequency financial time series have been contaminated by possible shifts in the distribution of the data. It applies a recursive version of the Brock–Dechert–Scheinkman statistic to daily data on two stock-market indexes between January 1980 and December 1990. It is shown that October 1987 is highly influential in the characterization of the stock-market dynamics and appears to correspond to a shift in the distribution of stock returns. Sampling experiments show that simple linear processes with shifts in variance can replicate the behavior of the tests, but autoregressive conditional heteroscedastic filters are unable to do so.

KEY WORDS: BDS test; Nonlinearity; Nonstationarity.

Constancy of the unconditional distribution of asset returns is a typical assumption of many time series models, including the vast class of autoregressive conditionally heteroscedastic (ARCH) processes. This class of processes models stock prices as nonlinear stationary processes. Modeling nonlinearities is not a simple task: Not only is the number of alternative nonlinear models vast, but their flexibility also creates the possibility of spuriously good fits, as noted by Granger and Teräsvirta (1993). Moreover, given the rate at which new financial and technological tools have been introduced in financial markets (e.g., Miller 1991) 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 Lorent and Phillips (1994) rejected the hypothesis that stock returns are covariance-stationary.

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) challenged this view, supporting the alternative hypothesis that stock prices are realizations of nonlinear stationary stochastic processes. They argued that nonstationarities would bias their frequency-domain-based methods toward acceptance of linearity. Given that their test statistics clearly indicated a rejection of linearity, they discarded the existence of nonstationarities in daily stock returns during the period July 1962 through December 1977. Using a different set of tools, Hsieh (1991) found that rejections of linearity in stock returns are mainly due to neglected conditional heteroscedasticity and cannot be attributed to structural changes (or chaotic dynamics). Hsieh also showed that a stochastic volatility model appears to capture most of the nonlinearities in stock returns.

Conditional heteroscedastic models have become the dominant time series model for stock returns. This class of models usually characterizes the volatility of high-frequency returns as an extremely persistent process. Diebold (1986) and Lamoureux and Lastrapes (1990) suggested that shifts in the unconditional variance could explain these common findings of persistence in the conditional variance. Simonato (1992) estimated a generalized autoregressive conditionally heteroscedastic (GARCH) process with changes in regime—using the Goldfeld and Quandt (1973) switching-regression method—to a group of European exchange rates and found that consideration of structural breaks greatly reduces evidence for ARCH effects. In fact, Simonato provided an example of an exchange rate—the Swiss franc—in which the ARCH effects become statistically insignificant when structural breaks in the unconditional variance are allowed. Another model that captures the idea of structural breaks in volatilities is the Cai (1994) and Hamilton and Susmel (1994) Markov switching ARCH (SWARCH) model. The idea that the pattern of conditional volatility is not constant over time was also developed by Diebold and Lopez (1995), who studied the sample autocorrelation function of the squared change in the log daily closing value of the S&P 500 stock index during different periods of time and were led to conclude that “there seems to be no GARCH effects in the 1980’s” (p. 459).

This article develops a testing methodology that formally attempts to discriminate between rejections of the null of linearity due to intrinsic nonlinearity and rejections that are due to nonstationarity in the data. The tests are based on a functional central-limit-type argument that shows that the partial sums of a test statistic introduced by Brock, Dechert, and Scheinkman (1987)—hereafter BDS—converge to Brownian motion. These tests are therefore a generalization of the popular BDS test for nonlinearity and are sensitive to shifts in variance as well as other changes in the distribution of the data. Moreover, this class of tests is robust to data generated by heavy-tailed distributions, an attractive property given the nature of the data analyzed in this article.

This testing methodology is applied to daily stock-market data covering the period January 2, 1980, to December 31, 1990—namely, the returns on the Standard and Poor’s 500 index (S&P 500) and the value-weighted index of the Cen-
ter for Research on Security Prices (VCRSP). The recursive BDS tests identify the October 1987 “crash” as a highly influential event in the study of the dynamics of stock-market returns. During the period January 1980 through October 1987, the recursive BDS tests do not reject the null hypothesis that both S&P 500 and VCRSP returns series can be viewed as a linear combination of iid random variables. The test statistics and some sampling experiments also indicate the existence of a shift in the distribution of stock returns around the stock-market crash of October 1987.

Although focusing on the impact of nonstationarities on the testing of nonlinearities, the article should not be understood as implying that nonlinear models are of little relevance for the study of stock-market dynamics. First, it is obvious that the October 1987 crash is not responsible for all rejections of iid linearity in stock returns. As pointed out by an anonymous referee, this can be clearly seen in the following experiment: Consider the daily returns on S&P 500 from 1920 to 1986; draw 1,000 sample periods of 1,500 consecutive observations. For \( m = 2 \), in 868 out of the 1,000 samples the BDS statistic is greater than 3.00 (\( p \) value of .001). From this perspective, the market dynamics of the period between January 1980 and October 1987 are unusually stable. Second, although the analysis presented in this article shows that consideration of nonstationarities reduces the evidence of nonlinearity in the sampling period under study, the tests also indicate that after October 1987 nonlinearities play a more active role in the dynamics of stock returns. Third, the methodology used in this article has some limitations. In particular, the characterization of the stock-market dynamics after the October 1987 crash is performed after a sample-splitting exercise suggested by the outcome of the testing procedure. This leads to potential biases in some of the analysis developed in the article.

Despite these problems, the article puts forward a significant amount of evidence providing additional support to the idea that the patterns of conditional heteroscedasticity change over time. Some of this evidence comes from some sampling experiments with ARCH filters fitted to the stock-returns data. These simulation exercises show that typical ARCH filters are not consistent with the behavior of the recursive BDS tests. Of all the ARCH models considered, however, the SWARCH model of Hamilton and Susmel (1994) is the data-generating process that comes closer to replicating the behavior of the recursive BDS tests in the stock-returns data.

Although somewhat controversial, these results are not entirely surprising. In fact, despite the vast literature on ARCH modeling, the amount of work on the statistical evaluation of ARCH models is small. Hamilton and Susmel (1994) presented predictive power comparisons between several ARCH-type models and a simple model that assumes constant variance. No matter what the loss function is, the prediction improvements derived from using ARCH estimates of conditional variances to estimate the squared returns series are quite small. In particular, if the mean squared error criterion is considered, the only model to improve on the constant-variance model is Hamilton and Susmel’s SWARCH model. The estimated SWARCH model—a 15-parameter model—implies, however, conditional heteroscedasticity and a prediction improvement of only 6% over a model with constant variance—a 2-parameter model. West and Cho (1995) performed similar predictive power comparisons for exchange-rates volatilities. Their results are in line with those of Hamilton and Susmel (1994).

The article is organized as follows. Section 1 presents the testing methodology used in the article and discusses its properties. Section 2 investigates the impact of nonstationarities on the findings of nonlinearities previously documented in the literature. October 1987 is identified as a highly influential observation in the study of nonlinearities during the sampling period considered. Section 3 assesses the robustness of those results in three different ways. First, a study of the behavior of the tests under data generated by three different ARCH filters fitted to the data is presented. It is then shown that the probability that those models provide a good approximation to the true data-generating process is quite small. Second, a very simple linear model with two unconditional variance periods is considered. Data generated from this model appears consistent with the behavior of the test statistics on the returns data. Third, this simple nonstationary model is used as the data-generating process in a simulation study of the properties of the estimators of ARCH models. The range of results obtained is consistent with the typical behavior of ARCH models in high-frequency data. In particular, the strong, persistent features of volatility processes are clearly reproduced. Section 4 presents some brief conclusions.

### 1. NONLINEARITY TESTS

Hsieh (1991) investigated the sources of nonlinearity in financial markets, concluding that conditional heteroscedasticity seems to be the main cause for the rejection of the hypothesis that (linearly filtered) stock returns are independent and identically distributed (iid). The concept of linearity used by Hsieh is iid linearity; that is, \( \{y_t\} \) is a linear process if the innovation sequence in the Wold decomposition for \( \{y_t\} \) forms an iid sequence. Hsieh’s testing strategy implicitly assumes that \( \{y_t\} \) has an ARMA\((p,q)\) representation, for \( p \) and \( q \) finite. He first fitted a linear autoregressive moving average (ARMA) model to the data and then applied the BDS test to the estimated residuals of this model.

The BDS test is based on the fact that, if \( \{y_t\} \) is an iid process, then \( C_{\varepsilon,m} - (C_{\varepsilon,1})^m = 0 \), almost surely for all \( \varepsilon > 0 \), and \( m = 1, 2, \ldots, \) where \( C_{\varepsilon,m} \trianglerighteq \lim_{T \to \infty} C_{\varepsilon,m} \) and \( C_{\varepsilon,m} = \sum \sum_{1 \leq s, t \leq T} E(Y_{t}^{(m)}Y_{s}^{(m)})/(T) \). \( Y_{t}^{(m)} \trianglerighteq \{y_{t}, y_{t+1}, \ldots, y_{t+m-1}\} \) defines an \( m \)-history process, \( T \) is the sample size, \( \|\cdot\| \) is the max-norm, and \( I_{\{\cdots\}} \) is the symmetric indicator kernel with \( I_{\{z = w\}} = 1 \) if \( \|z - w\| < \varepsilon \) and 0 otherwise. Brock et al. (1987) showed that \( V_{\varepsilon,m} = \sqrt{T}(\varepsilon_{m} - C_{\varepsilon,1})/\sigma_{\varepsilon,m} \) has a limiting standard normal distribution for all \( \varepsilon > 0 \) and \( m = 2, 3, \ldots \) under the null hypothesis of iid. Note that \( \sigma_{\varepsilon,m} \) is the asymptotic standard deviation of \( \sqrt{T}(\varepsilon_{m} - C_{\varepsilon,1}) \) under the null of iid.

The BDS test has been widely applied to financial time series (c.f., Brock, Hsieh, and LeBaron 1991; Hsieh 1991).
Simulation studies have shown that the test has power against a large class of alternatives, including ARCH models (see also Bollerslev, Engle, and Nelson 1994). As noted by Granger and Teräsvirta (1993), however, there is no Lagrangean-multiplier-type interpretation available for this test. Therefore, a rejection of the null hypothesis does not suggest the type of alternative models one should consider.

To avoid rejections of the null hypothesis due to linear dependence, the BDS test is commonly applied to the estimated residuals of ARMA models. The asymptotic distribution of the test is the same whether one uses estimated residuals or the true (unobserved) innovations under the null hypothesis of iid (see Brock et al. 1991; de Lima 1996). If this residual-based testing procedure is followed, a rejection of the null hypothesis of iid can thus be attributed to two main factors—the process can exhibit nonlinear dependence [either deterministic (e.g., chaos) or stochastic] and/or the process can be nonstationary. This is easily understood by noting that $C_{\varepsilon,m}$ is an estimator of $Pr\{\|Y_t^m - Y_s^m\| < \varepsilon\}$, whereas $C_{\varepsilon,1}$ estimates $Pr\{\|y_t - y_s\| < \varepsilon\}$. Under the iid hypothesis,

$$Pr\{\|Y_t^m - Y_s^m\| < \varepsilon\}
= Pr\{\|y_t - y_s\| < \varepsilon\},$$

that is, the BDS test estimates the difference between the joint distribution and the product of the marginal distributions in the appropriate intervals.

Hsieh’s (1991) investigation of the causes that lead to the rejection of the iid linearity in stock-market returns points in the direction of conditional heteroscedasticity, and shows that, although ARCH-type models do not fully capture nonlinearities, a more flexible stochastic volatility model is capable of doing so. Hsieh rejected the hypothesis that structural breaks are responsible for the rejection of the null 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 concluded that “it is unlikely that infrequent structural changes are causing the rejection of iid” (p. 1859).

This testing procedure does not discriminate the effects of the October 1987 crash because the observations corresponding to October 1987 were included in almost all the subsamples considered—with the exception of the four subsamples of 15-minute returns for 1988. Given the magnitude of the price variations, October 1987 is likely to play an important role in the outcome of the tests. Evidently, even from a purely statistical point of view, there are several competing explanations for the nature of the October 1987 stock market “crash” (i.e., a large variation in stock prices). It can correspond either to some large shock in a stationary linear process, it can be the manifestation of some nonlinear process, or it can be attributed to a shift in the distribution of the innovations. Furthermore, some combinations of these factors can also be considered.

In this article, the robustness of the findings of nonlinearity is evaluated through the partial sums of the BDS statistic,

$$C_{\varepsilon,m,[Tr]} - (C_{\varepsilon,1,[Tr]})^m;$$

where $C_{\varepsilon,m,[Tr]}$ denotes the correlation integral $C_{\varepsilon,m}$ computed with the first $[Tr]$ observations. $[Tr]$ denotes the integer part of $Tr$. This testing methodology is based on the following result by de Lima (1992), which shows that the normalized partial sums

$$V_{\varepsilon,m}(r) = \frac{\sqrt{Tr}}{\sigma_{\varepsilon,m}} (C_{\varepsilon,m,[Tr]} - (C_{\varepsilon,1,[Tr]})^m),$$

converge to a standard Wiener process.

**Theorem 1 (de Lima 1992).** If $\{u_t\}$ is an iid process with distribution function $F$, then

$$\sigma_{\varepsilon,m}^2 = 4 \begin{cases} m(m - 2)C^{2m - 2}(K - C^2) + K^m - C^{2m} \\
+ 2 \sum_{j=1}^{m-1} [C^{2j}(K^{m-j} - C^{2m-2j}) - mC^{2m-2j}(K - C^2)] \end{cases},$$

where

$$C = E[I_{\varepsilon}(u_t, u_s)] = \int [F(w + \varepsilon) - F(w - \varepsilon)] dF(w)$$

and

$$K = E[I_{\varepsilon}(u_t, u_s)I_{\varepsilon}(u_s, u_r)] = \int [F(w + \varepsilon) - F(w - \varepsilon)^2] dF(w).$$

Theorem 1 remains valid when $\sigma_{\varepsilon,m}^2$ is replaced by a consistent estimator $\hat{\sigma}_{\varepsilon,m}^2$. From the properties of the standard Wiener process, it follows immediately that Theorem 1 contains the BDS test as a particular case ($r = 1$). This theorem also implies that the finite-dimensional distributions (fdd’s) of $V_{\varepsilon,m}(r)$ are Gaussian.

Figure 1 describes the path of the (normalized) partial sums of the BDS statistic for a realization of an iid Normal(0, 1) (iidN) process, with the 95% critical values for the finite-dimensional distribution. As predicted by the result in Theorem 1, this path looks like a typical realization of a Brownian motion. Note that, although the functional $V_{\varepsilon,m}(r)$ is discontinuous—$V_{\varepsilon,m}(r)$ is defined in $D[0, 1]$, the space of functions on $[0, 1]$ that are right-continuous and have left-hand limits (cadlag functions)—it converges to a stochastic process whose sample paths are continuous. Therefore, under the null hypothesis, the probability of
abrupt jumps in the path of $V_{e,m}(r)$ diminishes as the sample size increases.

Another important property of the functional $V_{e,m}(r)$ is that no moment conditions are imposed on $\{u_t\}$ to derive Theorem 1. This is a simple consequence of the fact that the correlation integral is built on the indicator kernel $I(\mathbb{U}^m, \mathbb{U}^m)$, which is a bounded random variable. It follows that this family of tests can be used to test the null of iid, regardless of the existence of moments of any order for the series under scrutiny. Therefore, a realization of an iid process $u_t$ with a marginal Paretoian distribution implies that the partial-sums path $V_{e,m}(r)$ should be similar to the one obtained when the process $u_t$ is normally distributed. Figure 1 also describes the path of the (normalized) partial sums of the BDS statistic for a realization of an iid process $u_t$ with a marginal Paretoian distribution implies that the partial-sums path $V_{e,m}(r)$ should be similar to the one obtained when the process $u_t$ is normally distributed.

Figure 1. Recursive BDS Statistics Computed on 3,000 Independent Observations Generated From a Standard Normal Distribution (left) and a Symmetric Pareto Distribution With $\alpha = 2$ (right). Dashed lines represent the 95% confidence bands.

In this section, two different stock-returns series are tested for nonlinearities and nonstationarities, the returns to the S&P 500 stock-market index and the returns to the VCRSP index of the Center for Research in Security Prices. The data correspond to 2,781 daily observations from January 2, 1980, to December 31, 1990. This sample period is interesting for at least two reasons. First, it includes the October 1987 stock-market crash, allowing for a reevaluation of Hsieh's (1991) findings that evidence of nonlinearity is robust to the consideration of different subsample periods. Second, the sample period corresponds basically to the eighties, allowing for a reevaluation of Diebold and Lopez's (1995) finding that there seems to be no evidence of ARCH effects during the eighties. Prior to the analysis to be presented, all the series were demeaned and passed by a standard linear filter. This consisted of first removing the day-of-the-week and the month-of-the-year effects by regression on dummies. The residual correlations were then

2. NONLINEARITIES AND STRUCTURAL SHIFTS IN STOCK RETURNS

In this section, two different stock-returns series are tested for nonlinearities and nonstationarities, the returns to the S&P 500 stock-market index and the returns to the VCRSP index of the Center for Research in Security Prices. The data correspond to 2,781 daily observations from January 2, 1980, to December 31, 1990. This sample period is interesting for at least two reasons. First, it includes the October 1987 stock-market crash, allowing for a reevaluation of Hsieh's (1991) findings that evidence of nonlinearity is robust to the consideration of different subsample periods. Second, the sample period corresponds basically to the eighties, allowing for a reevaluation of Diebold and Lopez's (1995) finding that there seems to be no evidence of ARCH effects during the eighties. Prior to the analysis to be presented, all the series were demeaned and passed by a standard linear filter. This consisted of first removing the day-of-the-week and the month-of-the-year effects by regression on dummies. The residual correlations were then
removed by an autoregressive filter with order chosen by the Schwarz information criterion. The application of the BDS test to these two filtered series results in an overwhelming rejection of the iid-linear null hypothesis, as found by Hsieh (1991). Table 1 suggests that this result is robust to choice of \( \varepsilon \). To test how sensitive this result is to the so-called stock-market crash, I compute the same set of statistics on the sample period 01/02/80 to 10/01/87. The findings of nonlinearity in stock-return data seem to be very sensitive to the inclusion of October 1987 data. For both the S&P 500 and the VCRSP data, the null hypothesis of iid linearity cannot be rejected at conventional significance levels.

These results give some support to the findings that characterize stock returns as noncovariance stationary processes (see Hsu et al. 1974; Pagan and Schwert 1990; Lorentan and Phillips 1994). If the data-generating process is in fact nonstationary (with a possible shift in the unconditional variance of the innovations), this would cause the BDS test to reject the null of iid. It is worth noting that it is not known how robust the cumulative sum (CUSUM) test (used by both Pagan and Schwert and Lorentan and Phillips) is to the existence of nonlinearities in the data. For instance, the CUSUM test might be misleading if the conditional variance of stock returns is not constant over time. At the same time, the results presented in Table 1 need to be confronted with the possibility that some nonlinear process could generate this type of behavior of the BDS test.

The recursive BDS-type tests presented in Section 1 can be used to shed some light on those issues. In particular, they overcome the arbitrariness of sample splitting considered in Table 1. Throughout the rest of this section, I set \( m = 2 \) and \( \varepsilon = 1 \) standard deviation of the data. Some experiments with a few other values of \( \varepsilon \) indicate that the results to be described are not very sensitive to the choice of \( \varepsilon \). Figure 2 displays the sample path of the test statistics corresponding to the S&P 500 and VCRSP data.

The evolution of both sample paths shows that for any sample split beginning at 01/02/80 and ending at any point before 10/15/87 \( (r \approx .7) \) the null of iid would not be rejected at the 5% significance level. Evidently, this analysis is constrained by the fact that the tests might have reduced power for small values of \( r \)—note that for \( r < .11 \) the number of observations is smaller than 300. Simulations by Brock et al. (1991), however, showed that, at least for sample sizes larger than the ones associated with these small \( r \) values, the sampling distribution of the BDS statistic tends to be well approximated by its asymptotic distribution. In any case, the results reproduced in Figure 2 seem robust to sample-size problems because the sample path of the recursive tests only gets out of the 95% confidence bands after approximately 1,970 observations. This seems to be a relatively large number of observations, even for test statistics that demand large sample sizes. (A more thorough investigation of power issues is presented in Sec. 3.)

Therefore, the BDS test finds very little or no evidence of nonlinear dependence in stock returns (S&P 500 and VCRSP) during the eighties prior to the October 1987 crash. The abrupt jump in the sample path of the test statistics is also worth notice. By itself, the fact that the sample path crosses the 95% confidence bands does not provide a clear interpretation of the causes leading to the rejection of the hypothesis that stock returns are iid linear. At least two nonstationarity alternatives, however, seem capable of generating an empirical path for the recursive BDS tests consistent with the tests’ behavior under the S&P 500 and VCRSP indexes. Figure 3 shows a realization of the sample path of the BDS tests for two such alternatives. These two examples are introduced for the sake of illustration, and it is by no means suggested that they are an accurate description of the data.

<table>
<thead>
<tr>
<th>Sample period</th>
<th>Stock index</th>
<th>( c/r )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( .5 )</td>
</tr>
<tr>
<td>01/02/80 to</td>
<td>S&amp;P 500</td>
<td>1.89</td>
</tr>
<tr>
<td>12/31/90</td>
<td>VCRSP</td>
<td>2.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( .059 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( .003 )</td>
</tr>
<tr>
<td>01/02/80 to</td>
<td>S&amp;P 500</td>
<td>-1.61</td>
</tr>
<tr>
<td>10/01/87</td>
<td>VCRSP</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( .542 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( .803 )</td>
</tr>
</tbody>
</table>

Table 1. Nonlinearity Testing—BDS Statistic

NOTE: BDS statistics for embedding dimension \( m = 2 \). \( P \) values for the iid-linear null hypothesis are in parentheses.

Figure 2. Recursive BDS Statistics Computed on the S&P 500 Index (left) and the VCRSP Index (right) for the Period 1/2/80 to 12/31/90 (2,776 observations). Dashed lines represent the 95% confidence bands.
The first alternative model is a *two-variance* model; that is, two random samples of 500 observations were drawn from an N(0, 1) and an N(0, 4) distribution, respectively. The sample path is quite similar to the ones obtained from both the S&P 500 and the VCRSP series, especially in the region concerned with the jump from one regime to the other. Note that Hsieh (1991) presented simulation results on the small-sample power properties of the BDS test on a similar data-generating process. The test has no problems detecting this type of alternatives. The second alternative model is a *transient variance change* TVC model. Here, the data-generating process is an independent Gaussian process \( (X_t, \text{ with } t = 1, 2, \ldots, 2,800) \) in which the standard deviation of the data-generating process for \( t \) between 1,960 and 2,000 is four times larger than for the rest of the sample period. This TVC model incorporates some of the elements of the stochastic volatility model previously proposed by Hsieh (1991), the main difference being that in the TVC model a large variance shock disappears after a fixed number of periods, but in the stochastic volatility model the shock dies off exponentially. The behavior of the tests under the TVC model is also quite similar to the behavior of the tests under the stock-returns data.

One relevant question is whether the periods before and
after the jump in the sample path of the recursive tests are qualitatively similar. In other words, what contribution to the rejection of linearity does the after-crash period provide?

A first approach to this problem is to run the BDS recursive tests in reverse time; that is, apply the tests to the series \( y_t = z_{T-t+1}, t = 1, \ldots, T \). If \( y_t \) is iid so is \( y_t \). The results displayed at the top of Figure 4 are qualitatively similar to the ones obtained in direct time. In particular, there is also a pronounced jump in the path of the test statistics, occurring roughly between 12/22/87 and 10/13/87. This last point corresponds to the peak in this jump and is very close to the break-point date provided by the recursive tests when run in direct time. Note that from the beginning of the sample (12/31/90) up to the jump region the recursive tests are now outside the 95% confidence bands for most of the period, although only marginally. By comparison to the results obtained in direct time, this rejection of the iid null suggests that the dynamics of stock returns after October 87 become more complex. As could have been expected, the path appears fairly stable after the jump region.

Figure 4 also includes the sample path of the recursive tests for two sample-splitting exercises in direct time, corresponding to two different starting points. These two points are both chosen to potentially represent the beginning of a second period in the sample, as in a two-variance or a TVC model. This approach presents some serious difficulties to the methodology used in this article because the determination of the jump is an outcome of the testing procedure. Therefore, the application of the recursive tests to a subsample of data starting after the (endogenously determined) break point is likely to suffer from significant biases.

is similar to breaking a regression function in two according to the outcome of a testing procedure. Moreover, even if it is conceptually possible to develop corrected statistical procedures for this testing strategy, it should be pointed out that the method pursued in this article does not formally estimate break points but only suggests their possibility. Therefore, the starting point for a second period must be chosen somewhat arbitrarily.

The two starting points considered in the graphs at the bottom of Figure 4 give rise to quite distinct results. The choice of any point between 10/15/87—approximately the first time the sample path of the test statistics for the complete sample (in direct time) crosses the 95% confidence bands—and the middle of November 1987 results in the sample path of the test statistics crossing the 95% confidence bands shortly after its initialization point (bottom left graph). If one sets the first sample point after 11/20/87, however, a different pattern emerges (bottom right graph). Although the null hypothesis is still rejected (the terminal point in this new sample path is 2.78), evidence of predictability becomes considerably weakened. Only for large values of \( r \) is the sample path outside the 95% confidence band. Together with the reverse time results, and despite of the potential biases affecting the confidence bands, this last graph suggests that evidence for nonlinearity becomes stronger after the October 1987 crash. These graphs, however, also highlight the idea that nonstationarities tend to overstate the evidence for nonlinearity detected by the conventional BDS test.

3. ROBUSTNESS OF THE EMPIRICAL FINDINGS

This section investigates the sampling behavior of the recursive BDS tests under some nonlinear and nonstationary alternatives. As for the BDS test itself, no analytical characterization of the behavior of the tests under these alternative hypotheses is available. Robustness issues are thus addressed by means of a simulation study.

3.1 Nonlinear Alternatives: ARCH-type Models

Several nonlinear models could be considered as the potential data-generating process for the stock-returns series studied in Section 2. Models that display conditional heteroscedasticity such as ARCH appear, however, to be the dominant time series model for financial time series. For this reason, the simulation experiments presented in this section concentrate on some of the most commonly used ARCH-type filters. In particular, a GARCH(1, 1), an EGARCH(1, 0), and a SWARCH-L(3, 2) are estimated on the S&P 500 data. This choice of the lag polynomials reflects a combination of common practice [GARCH(1, 1)], automatic lag selection using Schwarz’s information criterion [EGARCH(1, 0)], and compatibility with the previous work of Hamilton and Susmel (1994) [SWARCH-L(3, 2)]. See Figure 5 for a description of the specific ARCH filters considered.

The sampling experiments consist of drawing 3,779 observations (the first 1,000 observations are discarded to avoid dependency on the initial conditions) from each pro-
cess and computing the recursive statistics $V_{\epsilon,m}(r)$, $0 \leq r \leq 1$, for the data thus obtained. As in Section 2, $m$ is set to 2 and $\epsilon$ to one standard deviation of the generated data. This procedure is repeated 1,000 times.

The results of this simulation exercise clearly suggest that none of the preceding ARCH-type processes is consistent with the behavior of the recursive BDS statistics. This statement is based on the following facts.

First, the number of cases in which the sample path of the recursive BDS tests remains inside the 95% bands after 1,960 observations is extremely small (.03% for GARCH, .02% for EGARCH, and 1.9% for SWARCH). Moreover, in none of these cases does the sample path remain inside the 95% bands throughout all the first 1,960 observations for the GARCH and EGARCH data. For the SWARCH data, only 1 out of these 19 cases remains inside the bands throughout all the first 1,960 observations. This last figure is computed without taking into consideration what happens during the first 300 observations ($r = 300/2,776 \approx .10$) to avoid small-sample problems. Remember that, as documented in Figure 2, the sample paths of the recursive tests for both the S&P 500 and the VCRSP series do not go out of the 95% bands for the first 1,960 observations (i.e., during the period 1/02/80 through 10/01/87). Finally, for the 1,000 generated series, the BDS test—$V_{\epsilon,m}(1)$—only fails to reject the null hypothesis at the 5% level in 3 of the 1,000 replications involving the SWARCH model, and it always rejects the null for the GARCH and EGARCH models.

Second, another set of interesting statistics is the estimates of the first time that the sample path $V_{\epsilon,m}(r)$ crosses the boundaries of the 95% confidence region; that is, $\{r: |V_{\epsilon,m}(r)| > 1.96\sqrt{r}, |V_{\epsilon,m}(s)| < 1.96\sqrt{s}, 0 \leq s \leq r\}$. Again, these estimates were computed for values of $r > .10$. In the simulated data, the mean of these estimates is $r = .13$ for GARCH data and $r = .14$ for both the EGARCH and the SWARCH data, with standard errors of .06, .07, and .08, respectively. Therefore, the corresponding value in the observed S&P 500 series—$r = 1.960/2,776 = .71$—is at least more than seven standard deviations away from the value obtained for any of the generated datasets.

Third, Figure 5 illustrates most of the preceding points quite clearly. This graph shows 95% confidence lower bands for the sample paths of the recursive BDS tests obtained in the simulations for each of the fitted ARCH filters, as well as the path of the recursive BDS test corresponding to the S&P 500 data. The 95% confidence lower bands represent the mean of the simulated paths minus two standard deviations. Clearly, the path of the test for the S&P 500 data comes out of the bands early on in the sampling period for any of the fitted filters. In other words, none of the data-generating processes appears consistent with the prolonged stay inside the 95% iid confidence bands displayed by the BDS recursive tests in Figure 2. After October 1987, however, the sample path of the BDS recursive tests borders on the 95% lower confidence band for the SWARCH-L(3, 2) generated data. Although this cannot be seen as evidence in favor of the SWARCH model versus the other two models of conditional heteroscedasticity, it is nonetheless interesting that the model that allows for different volatility periods is the one that comes “closer” to the behavior of the test for the S&P 500 data.

The results presented so far are very much in line with Diebold and Lopez’s (1995) general conclusion that conditional heteroscedasticity might not be as general a phenomenon as some of the modern literature seems to indicate, with patterns that may vary substantially across different time periods. The recursive BDS tests show that between January 1980 and October 1987 there is little evidence of nonlinearity and, more concretely, of conditional heteroscedasticity in the S&P 500 data.

3.2 Nonstationary Alternatives: Shifts in the Unconditional Variance

As stated in Section 2, structural shifts in the distribution of the data are a likely explanation of some of the results of the application of the recursive BDS tests to the stock returns indexes considered in this article. The remainder of this section presents some sampling experiments with models that allow for shifts in the unconditional variance. For that reason, the results already presented in the article are complemented with the analysis developed by Tsay (1988) to identify breaks in the variance of linear ARMA processes.

The Tsay (1988) approach to variance-change detection assumes that an observed time series $y_t$ is defined by

$$y_t = \begin{cases} a_t & \text{if } t < d \\ a_t(1 + \omega) & \text{if } t \geq d, \end{cases}$$

where $\omega$ is a constant and $a_t$ is an iidN sequence. Note that this setup can be easily generalized to accommodate the case in which $y_t$ is an ARMA(p, q) model. Tsay (1988) used an iterative procedure—labeled as Procedure V—to estimate $d$, the date of a variance change. With the exception of some observations at the beginning and the end of the sample period, the method assumes that every data point is a possible break point, and it computes the ratio of the sample variance before that date to the sample variance after that date. It then looks for too large or too small values in these variance ratios. If a break point is detected, the original series is scaled by the estimated variance ratio after the break point and the procedure is repeated until no new break points are detected. The distribution of the testing procedure is not known, but Tsay (1988) provided simulated critical values for a few sample sizes. All the critical values employed in this article, however, result from bootstrapping the null model of no variance changes for the relevant sample size. The first 50 and the last 50 observations are excluded in the search for potential break points. The reader is referred to Tsay (1988) for a detailed description of the method.

At the 1% significance level, the Tsay procedure detects three break points, occurring at $t = 1,950$, $t = 2,150$, and $t = 2,675$, which correspond to 09/21/87, 07/06/88, and 08/02/90. Therefore, the first break point ($t = 1,950$) almost coincides with the first time that the recursive BDS tests get out of the 95% confidence bands ($t = 1,960$). One remark is in order. The significance level used for the Tsay
procedure is more conservative than the one used for the BDS recursive tests. This asymmetric treatment of the two tests is due to some of the potential problems associated with the application of the Tsay procedure to stock returns. First, it is not known how the Tsay procedure might be affected by nonlinearities; namely, it might be misled by potential conditional heteroscedasticity. Second, the Tsay procedure is a moments-based technique, and even under the assumption that variances are finite, variance-change detection might be affected by the nonexistence of higher-order moments—namely, fourth moments. For these reasons, the critical values employed in the application of the Tsay procedure are deliberately conservative. In any case, it should be stressed that consideration of a 5% significance level does not lead to the detection of any break point before \( t = 1,950 \); that is, there are no break points between January 1980 and September 1987.

Tsay’s model assumes that, in between break points, (linearly filtered) stock returns can be described as iid processes. The following sampling experiment tries to assess the likelihood of this hypothesis: For each of the periods identified by the Tsay procedure, estimate the sample variance of the S&P 500 data; then use a normal distribution with zero mean and variance set equal to that sample variance to draw as many independent observations as the number of data points in the period under consideration. Finally, apply the recursive BDS test to the resulting data series. As for the sampling experiments with ARCH-type models, repeat this procedure 1,000 times.

This sampling experiment has some major shortcomings. First, the extensive work on the characterization of the marginal distribution of stock returns leaves little reason to believe that the innovations are drawn from a normal distribution. Second, although the first of the three break points detected by Tsay’s procedure is associated with an extensively described period of unusual activity in the stock market, the other two break points do not seem to be directly associated with economically meaningful events. For this last reason, I also considered a similar sampling experiment in which only one break point is considered, corresponding to the first of the break points determined by Tsay’s procedure.

Figure 6 displays the path of the BDS recursive tests for S&P 500 data as well as 95% confidence bands for the test when the data is generated according to the break-point models just described. These confidence bands are constructed like the ones in the sampling experiments involving ARCH-type models. This graph shows that the behavior of the recursive BDS test for S&P 500 data is consistent with a data-generating process that draws independent data from two distributions with different unconditional variances. This is in clear contrast to the results of the sampling experiments for ARCH-type models, as well as to Hsieh’s (1991) results that indicate that the rejections of linearity associated with the BDS test cannot be explained by nonstationarities. The results displayed in both Figures 5 and 6 suggest that, during the sample period analyzed in this article, structural shifts in the unconditional variance overstate the evidence for conditional heteroscedasticity.

### 3.3 ARCH Models and Nonstationary Alternatives

One question that the sampling experiments with the recursive BDS tests cannot address is why the estimated ARCH models presented in Section 3.1 provide strong evidence of ARCH effects in the data. This part of the article tries to address this question by considering the following experiment: Generate data from the one-break model considered in the previous section and fit a \( \text{GARCH}(1, 1) \) and an \( \text{EGARCH}(1, 0) \) to those series. Repeat this experiment 1,000 times and then investigate the sampling distribution of the parameter estimators. Although of considerable interest, this experiment does not include the SWARCH-L(3, 2) model because the estimation of the parameters in this model is extremely slow.

The data generated by the one-break model is taken by the maximum likelihood estimators of the parameters as coming from highly persistent ARCH models. This is revealed by the sampling distribution of the parameter estimators in these two ARCH filters (see Table 2). Restricting attention to the autoregressive coefficient in both the \( \text{GARCH}(1, 1) \) and \( \text{EGARCH}(1, 0) \) filters, the median estimates are .968 and .871, respectively. The corresponding standard errors are .002 and .092. This means that the one-break model is capable of generating data that is translated into point estimates similar to the ones obtained for the stock-returns series. Once again, it should be stressed that the one-break model used throughout these sampling experiments is not necessarily seen as a good approximation to the true data-generating process. Its purpose is to illustrate that a model with shifts in the unconditional variance

### Table 2. ARCH Models and Nonstationary Data

<table>
<thead>
<tr>
<th>Quantile</th>
<th>( \omega )</th>
<th>( \phi )</th>
<th>( \theta )</th>
<th>( \alpha )</th>
<th>( \psi_1 )</th>
<th>( \psi_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5%</td>
<td>9.991e-07</td>
<td>.963</td>
<td>.019</td>
<td>.916</td>
<td>.671</td>
<td>.035</td>
</tr>
<tr>
<td>50%</td>
<td>1.056e-06</td>
<td>.968</td>
<td>.022</td>
<td>.910</td>
<td>.871</td>
<td>.000</td>
</tr>
<tr>
<td>97.5%</td>
<td>1.117e-06</td>
<td>.972</td>
<td>.027</td>
<td>.905</td>
<td>.916</td>
<td>.034</td>
</tr>
</tbody>
</table>

NOTE: Sampling distribution of the maximum likelihood estimator of the parameters of two ARCH filters is fitted to data generated by the one-break model fitted to the S&P 500 data.
can replicate many of the features of some of the statistical tools commonly used in the analysis of stock-returns data. In particular, note that the autoregressive parameter in the GARCH(1, 1) model with the S&P 500 index data—.879—is not within the 95% confidence interval constructed from the corresponding simulated distribution—(.963, .972). The estimate of the autoregressive parameter in the EGARCH(1, 0) model—.909—lies within the bounds of the analogous 95% confidence interval, however, and, more importantly, the range of values obtained in the simulation experiment is consistent with strong persistence in the volatility process, as it is usually detected in stock-market data.

These results are also consistent with the work of Diebold (1986) and Lamoureux and Lastrapes (1990). These authors suggested that breaks in the unconditional variance could explain some of the findings of persistence in the conditional volatility.

4. CONCLUSIONS

This article uses a testing methodology developed by de Lima (1992) to address the question of whether findings of nonlinearities in stock returns have been contaminated by possible shifts in the distribution of the first differences of the logarithms of some commonly used stock-price indexes.

This article clearly identifies the October 1987 crash as a highly influential event in the study of the dynamics of stock-market returns. It also shows that some forms of nonstationarity have to be carefully considered in the modeling of financial time series—namely, that the patterns of conditional heteroscedasticity are not constant over time.

ACKNOWLEDGMENTS

Part of this research was conducted while I was a Visiting Scholar at the University of Strathclyde, Glasgow. I would like to thank all the members of the Department of Accounting and Finance for their hospitality and Craig Hsieh for extensive discussions on the topic of this article. I am also indebted to Michelle Barnes, Steve Blough, William Brock, M. Luisa Ferreira, Douglas Steigerwald, and one anonymous referee for helpful comments, as well as to Jim Hamilton for kindly sending me the GAUSS code for the estimation of SWARCH models. The usual disclaimer applies.

[Received March 1996. Revised April 1997.]

REFERENCES


You have printed the following article:

**Nonlinearities and Nonstationarities in Stock Returns**
Pedro J. F. de Lima
Stable URL: [http://links.jstor.org/sici?sici=0735-0015%28199804%2916%3A2%3C227%3ANANISR%3E2.0.CO%3B2-E](http://links.jstor.org/sici?sici=0735-0015%28199804%2916%3A2%3C227%3ANANISR%3E2.0.CO%3B2-E)

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

References

**The Stable-Law Model of Stock Returns**
Vedat Akgiray; G. Geoffrey Booth
Stable URL: [http://links.jstor.org/sici?sici=0735-0015%28198801%296%3A1%3C51%3ATSMOSR%3E2.0.CO%3B2-6](http://links.jstor.org/sici?sici=0735-0015%28198801%296%3A1%3C51%3ATSMOSR%3E2.0.CO%3B2-6)

**A Markov Model of Switching-Regime ARCH**
Jun Cai
Stable URL: [http://links.jstor.org/sici?sici=0735-0015%28199407%2912%3A3%3C309%3AAMMOSA%3E2.0.CO%3B2-Y](http://links.jstor.org/sici?sici=0735-0015%28199407%2912%3A3%3C309%3AAMMOSA%3E2.0.CO%3B2-Y)

**The Distribution of Futures Prices: A Test of the Stable Paretoian and Mixture of Normals Hypotheses**
Joyce A. Hall; Wade Brorsen; Scott H. Irwin
Stable URL: [http://links.jstor.org/sici?sici=0022-1090%28198903%2924%3A1%3C105%3ATDOFPA%3E2.0.CO%3B2-6](http://links.jstor.org/sici?sici=0022-1090%28198903%2924%3A1%3C105%3ATDOFPA%3E2.0.CO%3B2-6)

**Evidence of Nonlinearity in Daily Stock Returns**
Melvin J. Hinich; Douglas M. Patterson
Stable URL: [http://links.jstor.org/sici?sici=0735-0015%28198501%293%3A1%3C69%3AEONIDS%3E2.0.CO%3B2-W](http://links.jstor.org/sici?sici=0735-0015%28198501%293%3A1%3C69%3AEONIDS%3E2.0.CO%3B2-W)
Chaos and Nonlinear Dynamics: Application to Financial Markets
David A. Hsieh
Stable URL:
http://links.jstor.org/sici?sici=0022-1082%28199112%2946%3A5%3C1839%3ACANDAT%3E2.0.CO%3B2-P

On the Stable Paretian Behavior of Stock-Market Prices
Der-Ann Hsu; Robert B. Miller; Dean W. Wichern
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28197403%2969%3A345%3C108%3AOTSPBO%3E2.0.CO%3B2-I

Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance
Carla Inclan; George C. Tiao
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28199409%2989%3A427%3C913%3AUOCSOS%3E2.0.CO%3B2-H

Persistence in Variance, Structural Change, and the GARCH Model
Christopher G. Lamoureux; William D. Lastrapes
Stable URL:
http://links.jstor.org/sici?sici=0735-0015%28199004%298%3A2%3C225%3APIVSCA%3E2.0.CO%3B2-P

Measuring Tail Thickness to Estimate the Stable Index #: A Critique
J. Huston McCulloch
Stable URL:
http://links.jstor.org/sici?sici=0735-0015%28199701%2915%3A1%3C74%3AMTTTET%3E2.0.CO%3B2-H