A NEW DIAGNOSTIC TEST OF MODEL INADEQUACY WHICH USES THE MARTINGALE DIFFERENCE CRITERION

By Melvin J. Hinich and Douglas M. Patterson

University of Texas at Austin and Virginia Polytechnic Institute and State University

First version received October 1987

Abstract. Let $\{x(t)\}$ denote a discrete-time random process. Given a sample of increments $e(t) = x(t) - x(t-1)$ from the time series, we wish to test formally whether the sample is consistent with the assumption that $\{e(t)\}$ is a martingale difference. It is shown that the martingale criterion is more general than the white noise criterion in analyzing fitted residuals for signs of model inadequacy. In this paper we present such a test which approximately achieves a given type 1 error probability for samples. We assume that (1) the process is strictly stationary, (2) all its kth-order cumulant functions exist and (3) the kth-order cumulants are absolutely summable and satisfy a mixing condition. The martingale assumption implies that most third-order cumulants of the increment process are zero, and thus the third-order cumulant sequence is sparse. This result is used to derive test statistics based on a modified sample bispectrum. The test can be regarded as a two-dimensional portmanteau test of serial dependence. The large-sample results are demonstrated through the use of artificial data. Finally, the test is applied to a daily financial series.

Keywords. Discrete-time random process; martingale criterion; cumulant functions; model inadequacy; serial dependence; residuals; financial series; diagnostic checks.

1. BACKGROUND

A tenet of time series analysis is the idea that evidence of model inadequacy can often be found through diagnostic checks applied to the residuals of a fitted model. For example, the residuals are typically checked for autocorrelation. In the case of linear models, lack of autocorrelation is an indication of model adequacy. However, in a non-Gaussian environment lack of residual autocorrelation is a necessary, but not a sufficient, condition for model adequacy. Consequently, the white noise stopping rule needs to be revised. In this paper we present a new diagnostic test which is appropriate for non-Gaussian settings. The test can easily be applied to data analysis using a computer program called MARTIN, which is available from the authors.

It is well known that a wide variety of non-Gaussian and nonlinear models are also white noise processes. In the context of these more general models, the analog of white residuals is nonforecastable residuals in the mean square metric. This requirement is stronger than no serial autocorrelation and weaker than independence. Martingale difference processes are a class of stochastic processes which meet this new criterion. There are many nonlinear

0143-9782/92/03 233–20 $02.50/0

models whose variates are dependent but nevertheless are nonforecastable. Some examples of such models will be presented in Section 5.

In general, the conditional expectation provides an optimum forecast of the mean in the sense that it minimizes the mean square forecast error (see Papoulis, 1965, Section 11.1). In this paper we call a sequence $\{e(t)\}$ 'nonforecastable' if $E\{|e(t)|\} < \infty$ for all t and if

$$E\{e(t+1)|e(t), e(t-1), \ldots\} = 0 \qquad (t \geqslant 1). \tag{1.1}$$

Let $\{x(t)\}$ be the partial sums of $e(t)$, i.e. $x(1) = e(1)$, $x(2) = x(1) + e(2), \ldots$. Then

$$E\{x(t+1)|x(t), x(t-1), \ldots\} = x(t) \qquad (t \geqslant 1). \tag{1.2}$$

Equation (1.1) defines a martingale difference process and Equation (1.2) defines a martingale process.

The value of the martingale difference criterion as a diagnostic check can be put into better focus through consideration of the following AR(1) model:

$$y(t) = ay(t-1) + u(t) \qquad (|a| < 1) \tag{1.3}$$

where $\{u(t)\}$ is a stationary white noise series. The conditional expectation of $y(t)$ is

$$E\{y(t)|y(t-1), y(t-2), \ldots\} = ay(t-1) \tag{1.4}$$

if and only if

$$E\{u(t)|u(t-1), u(t-2), \ldots\} = 0. \tag{1.5}$$

In other words (1.4) holds if and only if $u(t)$ is a martingale difference. Next let $u(t)$ be a quadratic nonlinear error sequence:

$$u(t) = \varepsilon(t) + \sum_{m=1}^{L} a(m)\varepsilon(t-1)\varepsilon(t-m-1) \tag{1.6}$$

where the $\varepsilon(t)$ are independent and identically distributed random variates and $A(z) = \sum_{m=1}^{L} a(m)z^m$ has no zeros inside the unit circle $\mathrm{mod}(z = 1)$ in the complex plane. This error sequence is not a martingale difference since

$$E\{u(t)|u(t-1), u(t-2), \ldots\} = \sum_{m=1}^{L} a(m)\varepsilon(t-1)\varepsilon(t-m-1), \tag{1.7}$$

and the conditional expectation of $y(t)$ is not $ay(t-1)$ but rather

$$E\{y(t)|y(t-1), y(t-2), \ldots\} = ay(t-1) + \sum_{m=1}^{L} a(m)\varepsilon(t-1)\varepsilon(t-m-1). \tag{1.8}$$

Note that the error sequence in (1.8) is white noise, and that its serial dependence will not be detected by the usual diagnostic tests.

In this paper we derive a statistical test to determine whether a sample is consistent with the asssumption that the data are generated by a martingale difference process or, equivalently, that the partial sums follow a martingale. The martingale difference assumption implies that the difference series is serially uncorrelated; it also implies that most third-order cumulants of the difference process are zero. The later fact is used to develop test statistics calculated from a modified bispectrum. The test can be regarded as a two-dimensional form of the 'portmanteau' test of autocorrelation in fitted residuals because it considers the third-order cumulants (bicovariances) of these residuals taken as a whole. In developing the test, however, we work in the frequency domain in order to take advantage of certain well-known asymptotic results concerning the sample bispectrum. The test is presented in Sections 2, 3 and 4. In Section 5, the size and power of the test are examined using artificially generated data. We demonstrate the test in Section 6 using financial data which economic theory suggests should follow a martingale difference.

2. INTRODUCTION TO THE TEST

Let $\{x(t)\}$ denote a discrete-time random process for integer t. Define the increment $e(t) = x(t) - x(t - 1)$. Given a sample of increments $e(1)$, $e(2), \ldots, e(N)$ from the time series, we wish to test formally whether the sample is consistent with the assumption that $\{x(t)\}$ is a martingale. In this paper we present such a test which approximately achieves a given type 1 error probability for large values of N. The test holds for several assumptions which are required to apply large-sample theory. In particular, we assume the following.

(i) $\{e(t)\}$ is mean zero and strictly stationary.

(ii) All kth-order cumulant functions exist for $\{e(t)\}$.

(iii) The kth-order cumulants of $\{e(t)\}$ are absolutely summable and satisfy the mixing condition stated by Brillinger (1975, Assumption 2.6.2(2)).

We note that some sort of stationarity and mixing (short-memory) conditions are required to perform inferential analysis on time series data. The mixing condition holds if the process is m-dependent (Billingsley, 1979, p. 315).

The nth second-order cumulant of $\{e(t)\}$ is the covariance $c_{ee}(n) = E\{e(t)e(t + n)\}$ for integer n. Note that the second-order cumulants for $n \neq 0$ are all zero for a martingale difference process. The (m, n)th third-order cumulant is

$$\begin{aligned} c_{eee}(m, n) &= E\{e(t)e(t + m)e(t + n)\} \\ &= c_{eee}(n, m) \qquad ((2.1) \\ &= c_{eee}(m - n, - n) = c_{eee}(n - m, - m). \end{aligned}$$

For example, $E\{e^2(t)e(t - n)\} = c_{eee}(0, - n) = c_{eee}(n, n)$. We shall use this example later. We use the term (m, n)th bicovariant for Equation (2.1).

The bispectrum of the process is its Fourier transform in two indices. For $m, n = -\infty, \ldots, \infty$

$$s_{eee}(f, g) = \sum_m \sum_n c_{eee}(m, n) \exp\{-i2\pi(fm + gn)\}. \tag{2.2}$$

The bispectrum is periodic in (f, g), and its principal domain D is the triangle defined by the lines $g = f$, $g = 0$ and $2f + g = 1$. The inverse Fourier transform of s_{eee} yields

$$c_{eee}(m, n) = \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} s_{eee}(f, g) \exp\{i2\pi(fm + gn)\}\, df\, dg. \tag{2.3}$$

Assumption 2.6.2(2) of Brillinger (1975) requires that $\sum_n n^2 |c_{eee}(n, n)| < \infty$, which implies that $\partial^2 s_{eee}(f, g)/\partial f^2$ exists for all f.

We shall now show that the bicovariants of the martingale difference process $\{e(t)\}$ are zero unless $m = n \geq 0$, or $m = 0$ and $n \leq 0$, or $n = 0$ and $m \leq 0$. If $\{x(t)\}$ is a martingale, the conditional expectation of $x(t)$, given a realization of $x(t-1), x(t-2), \ldots$, is equal to the realization of $x(t-1)$ for all values of t. Then, from Doob (1953, expression (7.4)),

$$E\{e(t)e(t-r)e(t-s)\} = 0 \qquad \text{for all } r, s > 0. \tag{2.4}$$

The expected value can be nonzero for $r = 0$ and $s \geq 0$. Thus $c_{eee}(-r, -s) = 0$ unless $r = 0$ and $s \geq 0$ (or $s = 0$ and $r \geq 0$, or $r = s \leq 0$ by the symmetries of the cumulant function). Its bispectrum is

$$s_{eee}(f, g) = Ee^3(t) + \sum_{n=0}^{\infty} c_{eee}(n, n)[\exp(i2\pi fn) + \exp(i2\pi gn) + \exp\{-i2\pi(f + g)n\}] \tag{2.5}$$

where the potentially nonzero bicovariants are nuisance parameters. If the process $\{e(t)\}$ is a time-reversible martingale difference, then $c_{eee}(r, s) = 0$ for all r, s except $r = s = 0$. We do not restrict ourselves to time-reversible martingales.

In general, $E\{e(t)E(t - m_1)e(t - m_2) \ldots e(t - m_k)\} = 0$ for all positive $m_1, \ldots, m_k$. This implies that increments of a martingale process are white noise since $E\{e(t)e(t - m)\} = 0$ for $m > 0$.

Out test procedure exploits the sparseness property of the bicovariance function of a differenced time-irreversible martingale process. Failure to reject the null hypothesis does not imply that the process is a martingale difference since we only deal with the third-order cumulants. More to the point, a differenced nonmartingale process can be white noise and have zero bicovariants, and our test will fail to reject such a process with probability $1 - \alpha$ where α is the type 1 error probability of the test. Our test is a first step of a more general iterative procedure that uses the trispectrum and higher-order cumulant spectra.

3. LOGIC OF THE TEST

An exposition of our method requires an introduction to bispectrum estimation. The bispectrum of a time series, as with its spectrum, is estimable from a finite data record using two asymptotically equivalent approaches. In the first approach, the estimate is computed by Fourier transforming a 'windowed' sample bicovariance array. The second or 'direct' approach computes estimates of the bispectrum directly from the Fourier transform of the data record. A review of the windowed approach is given by Subba Rao (1983). We point out that the direct approach is computationally more efficient because it avoids the time-consuming calculation of lagged products.

The first step in the direct approach is to apply the discrete Fourier transform to the data record. The Fourier transform ordinates are then used to compute a set of complex values which is a two-dimensional analog of the periodogram. This 'bispectrogram' is smoothed to produce estimates of the bispectrum in a grid of frequency pairs in its principal domain. If the data record is long, the sample should be divided into segments and the bispectrum computed for each segment and then averaged to give an estimate for the sampled time series. This direct method is sketched by Brillinger and Rosenblatt (1967b), Hinich and Clay (1968) and Huber *et al.* (1971). Hinich (1982) presents computational details for smoothing over squares in the principal domain. Patterson (1983) developed a computer algorithm which calculates the Hinich bispectral estimator. Rosenblatt (1983) reviews the general approach to estimating kth-order polyspectra, which is discussed in more detail by Brillinger and Rosenblatt (1967b). The spectrum is the second-order polyspectrum and the bispectrum is the third-order case.

The idea behind the test is easier to explain using the windowed approach to estimation. Let $\{C_{eee}(m, n): m, n = -N+1, \ldots, N-1\}$ denote the sample bicovariance computed from a sample of size N with the symmetries of expression (2.1). Let $\{w_N(m, n)\}$ denote a (double) lag window whose associated bispectrum smoothing kernel is

$$W_N(f, g) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} w_N(m, n) \exp\{-i2\pi(fm + gn)\} \tag{3.1}$$

In practice W_N is of the form $B_N^{-2} W(B_N^{-1} f, B_N^{-1} g)$ where W is some kernel and B_n is a bandwidth parameter to be specified later. The bispectrum estimator, on an equally spaced grid of bifrequencies in D (i.e. the principal domain), is then given by

$$S_{eee}\{f(j), g(k)\} = \sum_m \sum_n C_{eee}(m, n) w_N(m, n) \exp[-i2\pi\{mf(j) + ng(k)\}]. \tag{3.2}$$

The asymptotic results presented by Brillinger and Rosenblatt (1967a) and Rosenblatt (1983) imply that, for large N and under adequate conditions, the estimates shown in (3.2) are approximately independent Gaussian variates with mean $S_{eee}\{f(j), g(k)\}$ and variance of the real or imaginary part of $A\sigma_e^6/2NB_N^2$, where σ_e^2 denotes the variance of $e(t)$ and

$$A = \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} |W(f, g)|^2 \, df \, dg. \tag{3.3}$$

Thus, for each frequency pair in D, the distribution of

$$\mathrm{CH}(j, k) = \frac{2NB_N^2 |S_{eee}\{f(j), g(k)\} - s_{eee}\{f(j), g(k)\}|^2}{A\sigma_e^6} \tag{3.4}$$

is approximately χ^2 with two degrees of freedom, and that of the sum of the $\mathrm{CH}(j, k)$ is approximately χ^2 with $2P$ degrees of freedom where P is the number of frequency pairs in D. This result yields the test for the null hypothesis that $S_{eee}(f, g) = 0$ for all f, g (i.e. a test for Gaussianity).

In the present problem, the null hypothesis does not imply $S_{eee}(f, g) = 0$ but only that $C_{eee}(m, n) = 0$ outside the set ($m = n = 0$, or $m = 0$ and $n \leq 0$, or $n = 0$ and $m \leq 0$). Thus, let $d(m, n)$ be the indicator function of this set, and then the Fourier transform of $\{1 - d(m, n)\}C_{eee}(m, n)$ is zero under the null hypothesis. By analogy with the bispectrum estimator, the last can be estimated by (3.2) with $v_N(m, n) = w_N\{1 - d(m, n)\}$ replacing $w_N(m, n)$, and denoted by $S_{\mathrm{mod}}\{f(j), g(k)\}$. It will be shown in the next section that this adjusted bispectrum estimator has the same asymptotic covariance structure as before. Thus the sum of

$$\sum_m \sum_n v(m, n) C_{eee}(m, n) \exp\{-i2\pi(fm + gn)\} \tag{3.5}$$

can be used as the test statistic, as it is approximately a χ^2 variate with $2P$ degrees of freedom under the null hypothesis. In practice, σ_e is replaced by its sample estimate computed from the N observations. The error in the estimate is negligible in terms of the asymptotic variance of the bispectrum estimator and does not affect the above result (Hinich, 1982).

4. LARGE-SAMPLE PROPERTIES OF THE ADJUSTED BISPECTRUM

To set up the large-sample approximation for the distribution of our new test statistic $S_{\mathrm{mod}}\{f(j), g(k)\}$, let $\{B_N\}$ denote a bandwidth sequence where $B_N = 0(N^{c-1})$ and $1/2 < c < 1$. Thus $B_N \to 0$ and $B_N^2 N \to \infty$ as $N \to \infty$. If the bispectrum at (f, g) is estimated by smoothing M^2 sample bispectral values over a square centered at (f, g), then $B_N = M/N$. In this case $M = [N^c]$. If the estimates are an average of $K = N/L$ bispectrograms computed from non-overlapping segments of L observations, then $B_N = 1/L$ and $K = [N^c]$.

Let $W_N(f, g) = B_N^{-2} W(B_N^{-1} f, B_N^{-1} g)$ where $W(f, g)$ satisfies the following conditions.

(i) W is a bounded and continuous function in each variable.
(ii) $\int_{-1/2}^{1/2}\int_{-1/2}^{1/2} |W(f, g)|\, df\, dg = 1$.
(iii) W shares the symmetries of the bispectrum, i.e. $w(f, g) = w(g, f) = w(f, 1 - f - g) = w(1 - f - g, g)$ and $W^*(f, g) = W(-f, -g)$.

THEOREM. *Under the stated conditions the asymptotic variance–covariance of $N^{1/2} B_N S_{\text{mod}}\{f(j), g(k)\}$ is the same as that of $N^{1/2} B_N S_{eee}\{f(j), g(k)\}$ as $N \to \infty$.*

PROOF. From Brillinger and Rosenblatt (1967a),

$$\sum_m \sum_n C_{eee}(m, m) w_N(m, n) \exp\{-i2\pi(fm + gn)\} = \int\int W_N(f - f', g - g') I_{eee}^{(N)}(f', g')\, df'\, dg' \tag{4.1}$$

where

$$I_{eee}^{(N)}(f, g) = \frac{d_e(f) d_e(g) d_e^*(f + g)}{N} \tag{4.2}$$

is the third-order periodogram, with the asterisk indicating the complex conjugate, and

$$d_e(f) = \sum_{t=0}^{N-1} e(t) \exp(-i2\pi f t) \tag{4.3}$$

is the discrete Fourier transform of the data record. The asymptotic properties of the bispectral estimator follow from the asmptotic properties of the third-order periodogram and the form of the smoothing kernel as $N \to \infty$. We shall now show that the kernel for the lag window $\{v(m, n)\}$ converges to the kernel of $\{w_N(m, n)\}$, and thus the asymptotic variance of S_{mod} is the same as that of the unmodified bispectral estimator.

Consider the unidimensional smoothing kernel $K(f) = \int W(f, g)\, dg$. The assumptions made for W imply that K is continuous and bounded, and that $\int |K(f)|\, df = 1$. Let $K_N(f) = B_N^{-1} K(B_N^{-1} f)$. Then, for each n,

$$w_N(n, 0) = \int\int W_N(f, g) \exp(i2\pi f n)\, df\, dg = \int K_N(f) \exp(i2\pi f n)\, df. \tag{4.4}$$

Since

$$\sum_{n=0}^{N} \exp(-i2\pi f n) = \exp(-i\pi N f) \frac{\sin\{\pi(N + 1) f\}}{\sin(\pi f)} \tag{4.4}$$

the following inverse relation holds:

$$\sum_{n=0}^{N} w_N(n, 0) \exp(-i2\pi fn) = K_N(f) + 0(N^{-1}). \tag{4.5}$$

It thus follows from (3.8) and (4.5) that the smoothing kernel for $w_N(m, n)d(m, n)$ is $K_N(-f) + K_N(-g) + K_N(f+g) + 0(N^{-1})$. The kernel for $v(m, n) = w_N(m, n)\{1 - d(m, n)\}$ and large N is

$$B_N^{-2}\{W(B_N^{-1}f, B_N^{-1}g) - B_N^{-1}WD(B_N^{-1}f, B_N^{-1}g)\} \tag{4.6}$$

where

$$WD(B_N^{-1}f, B_N^{-1}g) = K(-B_N^{-1}f) + K(-B_N^{-1}g) + K\{B_N^{-1}(f+g)\}. \tag{4.7}$$

Thus the difference between the kernel for $w_N(m, n)$ and that for $v(m, n)$ is of order $0(B_N)$ which goes to zero as $N \to \infty$. This completes the proof. ■

We next turn to the bias level of the S_{mod} estimator. The magnitude of the bias for the bispectrum estimator defined in (3.2) for frequency pair (f, g) is of order $0(GB_N^2)$ where

$$G(f, g) = \frac{|\partial^2 s_{eee}(f, g)/\partial f^2 + \partial^2 s_{eee}(f, g)/\partial g^2|}{2} \tag{4.8}$$

Now, with the approximation $B_N \approx N^{-1/2}$ the squared magnitude of the bias is of order $0(G^2/N^2)$. From (2.2), (4.8) and Parceval's theorem, the sum of squared biases for P bifrequencies is

$$\text{SSB} = 0\left\{N^{-2}\sum_m \sum_n (m^2 + n^2)^2 c_{eee}^2(m, n)\right\}. \tag{4.9}$$

If the $\{e(t)\}$ is a martingale difference, then

$$\text{SSB} = 0\left\{N^{-2}\sum_n n^4 c_{eee}^2(n, n)\right\} \tag{4.10}$$

Assume that there are l nonzero $c_{eee}(n, n)$ in (4.10) and suppose that the arithmetic average of the c_{eee} is $\bar{c}$. Then

$$\sum_{n=1}^{l} n^4(\bar{c})^2 = \frac{\bar{c}^2}{30} l(l+1)(2l+1)(3l^2 + 3l - 1) \approx \frac{l^5}{5}\bar{c}^2$$

and thus

$$\frac{1}{N^2}\sum_{n=1}^{l} n^4(\bar{c})^2 \approx \frac{l^5}{5N^2}\bar{c}^2. \tag{4.11}$$

For a finite but large N and $B_N \approx N^{-1/2}$, the complex variance of S_{mod} is approximately $A\sigma_e^6$ (see the discussion following Equation (3.2)). Hence it follows from (4.11) that the sum of the squared magnitudes of the bias divided by the variance of S_{mod} is of order

$$0\left(\frac{l^5\bar{c}^2}{5N^2\sigma_e^6}\right). \tag{4.12}$$

Clearly, $l^{5/2} \ll N$ for the bias to be small.

This bias in the large-sample variance will affect the type 1 error probability of our test if there are a number of large values of $c_{eee}(n, n)/\sigma_e^3$. In filter theory terminology, the bispectrum of the martingale 'leaks' through the sidelobes of our sharp diagonal filter $\{w(m, n)d(m, n)\}$.

A smaller bias can easily be achieved by using a tapered filter. We use the following two-dimensional simple taper in our application of the method to data:

$$d(m, n) = \begin{cases} 1 - |m - n|/V & \text{if } |m - n| < V \\ 0 & \text{otherwise} \end{cases}$$

for $m, n = 0, 1, \ldots$. This filter has a smaller bias than the 0–1 filter that we used to explain the method. The use of this V-notch shrinks the distribution of S_{mod} toward zero.

5. SIZE AND POWER OF THE TEST

In this section we examine the size of the martingale test using artificially generated data. In order to implement the test, we wrote a computer algorithm called MARTIN. The mainframe version of the program is written in FORTRAN and can handle up to 10 000 observations, whereas the personal computer (PC) version is limited to 4200 observations. The PC version is available as an executable program.

The program breaks a time series of length N into smaller even-sized data frames (these can range from a length of 10 to a length of 128). The end of each frame is padded with zeros (the number of zeros is equal to the frame length) so as to double the period of the periodic extension of the data. Next, the third-order periodogram of each frame is computed and these estimates are averaged. This approach of averaging the sample bispectrum of data frames is discussed by Hinich and Clay (1968) and Rosenblatt (1985, Section 5.5). The estimated average bispectrum is transformed back into the time domain and passed through the tapered filter described at the end of Section 4. After filtering, the estimated bicovariants are transformed back into the frequency domain to produce the estimated adjusted bispectrum $S_{\text{mod}}\{f(j), g(k)\}$ for a grid of width 1/NF, where NF is the number of observations in each frame. It is necessary to transform back to the frequency domain because the estimated bicovariants are correlated, i.e. the Fourier transform turns correlations among lag indices into heteroskedastic complex normal variates over bifrequency indices as in the spectrum correlation function case. We note that the bispectral estimates obtained in the frequency domain are asymptotically independent.

Under the null hypothesis that the series is a martingale, $S_{\text{mod}}\{f(j), g(k)\} = 0$. Hence, (3.4) can be written as

$$\mathrm{CH}(j, k) = \frac{2NB_N^2|S_{\mathrm{mod}}\{f(j), g(k)\}|^2}{A\sigma_e^6}. \tag{5.1}$$

The sum of the CH(j, k) with σ_e^6 replaced by its estimate is the martingale test statistic. It is distributed as approximately a χ^2 $2P$ variate under the null hypothesis.

The behaviour of the martingale test was investigated using nonlinear models of the quadratic type with a simple linear term. This class of models can be written as

$$e(t) = \varepsilon(t) + \sigma \sum_{m=1}^{L} \sum_{n=0}^{K} a_m(n)\varepsilon(t-n)\varepsilon(t-m-n) \tag{5.2}$$

where the $\varepsilon(t)$ are independent and identically distributed random deviates with zero mean and variance 1.0. The quadratic model (5.2) is also a martingale difference if $K = 0$:

$$e(t) = \varepsilon(t) + \sigma\sum_{m=1}^{L} a_m(0)\varepsilon(t)\varepsilon(t-m). \tag{5.3}$$

The bicovariances of this model are $\{a_m(0)\sigma\}$, and thus

$$\sum_m m^4 c_{eee}^2(m, m) = \sum_{m=1}^{L} m^4 a_m^2(0)\sigma^2.$$

This implies that the bias problem increases as σ or L increases.

Although the behaviour of the test was studied using a variety of quadratic models, we shall restrict out attention here to the following two models, of which the first is a martingale difference and the second is not.

Model 1. Fifteen-term quadratic martingale difference:

$$\begin{aligned} e(t) = \mathrm{e}(t) + \sigma\varepsilon(t)\{&-0.429\varepsilon(t-2) - 0.949\varepsilon(t-4) + 0.872\varepsilon(t-7)\\ &+ 0.489\varepsilon(t-8) - 0.694\varepsilon(t-10) + 0.683\varepsilon(t-15)\\ &+ 0.905\varepsilon(t-17) - 0.954\varepsilon(t-20) - 0.922\varepsilon(t-24)\\ &- 0.349\varepsilon(t-27) - 0.657\varepsilon(t-29)\\ &+ 0.703\varepsilon(t-31) + 0.570\varepsilon(t-33) + 0.105\varepsilon(t-34)\\ &+ 0.590\varepsilon(t-37)\}. \end{aligned}$$

Model 2. Fifteen-term quadratic nonmartingale difference:

$$\begin{aligned} e(t) = \varepsilon(t) + \sigma\{&-0.429\varepsilon(t-1)\varepsilon(t-17) + 0.570\varepsilon(t-2)\varepsilon(t-25)\\ &+ 0.105\varepsilon(t-3)\varepsilon(t-27) + 0.590\varepsilon(t-4)\varepsilon(t-29)\\ &- 0.657\varepsilon(t-5)\varepsilon(t-20) + 0.683\varepsilon(t-6)\varepsilon(t-9)\\ &- 0.949\varepsilon(t-7)\varepsilon(t-23) + 0.703\varepsilon(t-8)\varepsilon(t-21)\\ &- 0.694\varepsilon(t-9)\varepsilon(t-14) + 0.872\varepsilon(t-11)\varepsilon(t-17) \end{aligned}$$

$$+ 0.489\varepsilon(t-14)\varepsilon(t-25) + 0.905\varepsilon(t-16)\varepsilon(t-27)$$
$$- 0.954\varepsilon(t-20)\varepsilon(t-28) - 0.992\varepsilon(t-24)\varepsilon(t-32)$$
$$- 0.349\varepsilon(t-26)\varepsilon(t-41)\}.$$

The $\varepsilon(t)$ variates were generated by the Gaussian pseudo-random number routine GGNML in the International Mathematics and Statistics Library (IMSL) and scaled by either $\sigma = 0.5$ or $\sigma = 0.3$. The lags on the $\varepsilon(t)$ variates in Model 1 are distributed between 2 and 37, and the coefficients of the quadratic terms are independent realizations from a rectangular density on $[-1, 1]$. Figure 1 shows a realization of Model 1 of length 500 with a scale parameter equal to 0.5. The relatively large number of outliers seen in the figure illustrates the non-Gaussian nature of this model.

Model 2 is constructed in a 'freehand' manner from Model 1, but with lagged quadratic terms which break the martingale difference property of Model 1. The same two values of σ are used to simulate the power of the test for Model 2.

Table I summarizes the results from simulating Model 1 in order to estimate the size of the test for the nominal 5% and 1% levels. The column headings show the width of the V-notch used to reduce the bias. Model 1 was replicated 200 times for each V. The same initial 'seed' was passed to GGNML at the start of each set of 200 simulations in order to facilitate comparisons between entries in Table I. For panels (a) and (b) the sample size of each simulation was 3844 and the frame length NF was 62. The first and second rows of each panel report on the 5% and 1% size of the test. The remaining rows provide information about the distribution of the martingale test statistic. For convenience, the test statistic is expressed as a standardized normal variate (z statistic) in the table. The third, fourth and fifth rows show the 90%, 95% and 99% fractiles of the sample distribution of the test statistic for each set of 200 replications. The last two rows of each panel give the minimum and maximum values of the z statistic. Panels (c) and (d) of Table I show the results of the simulation when the sample size is increased to 10 000 with NF = 100 (still 200 replications for each V). It is obvious from the table that the appropriate width for the V-notch depends on the form of the martingale model generating the data.

The estimated power of the test against a nonmartingale (Model 2) is reported in Table II. The organization of the table is similar to that of Table I. Also, 200 replications were made for each V-notch, and the combinations of sample sizes and frame lengths were unchanged. The observed reduction in power when σ goes from 0.3 to 0.5 reflects the increase in bias mentioned just below Equation (5.3).

Power can always be traded for a conservative position *vis-à-vis* type 1 error probability by using a sufficiently large notch width (as a fraction of N). We recommend using various notch settings, say from 3% to 8% of N. If the test statistic is not consistent with the null hypothesis when V is large, it is

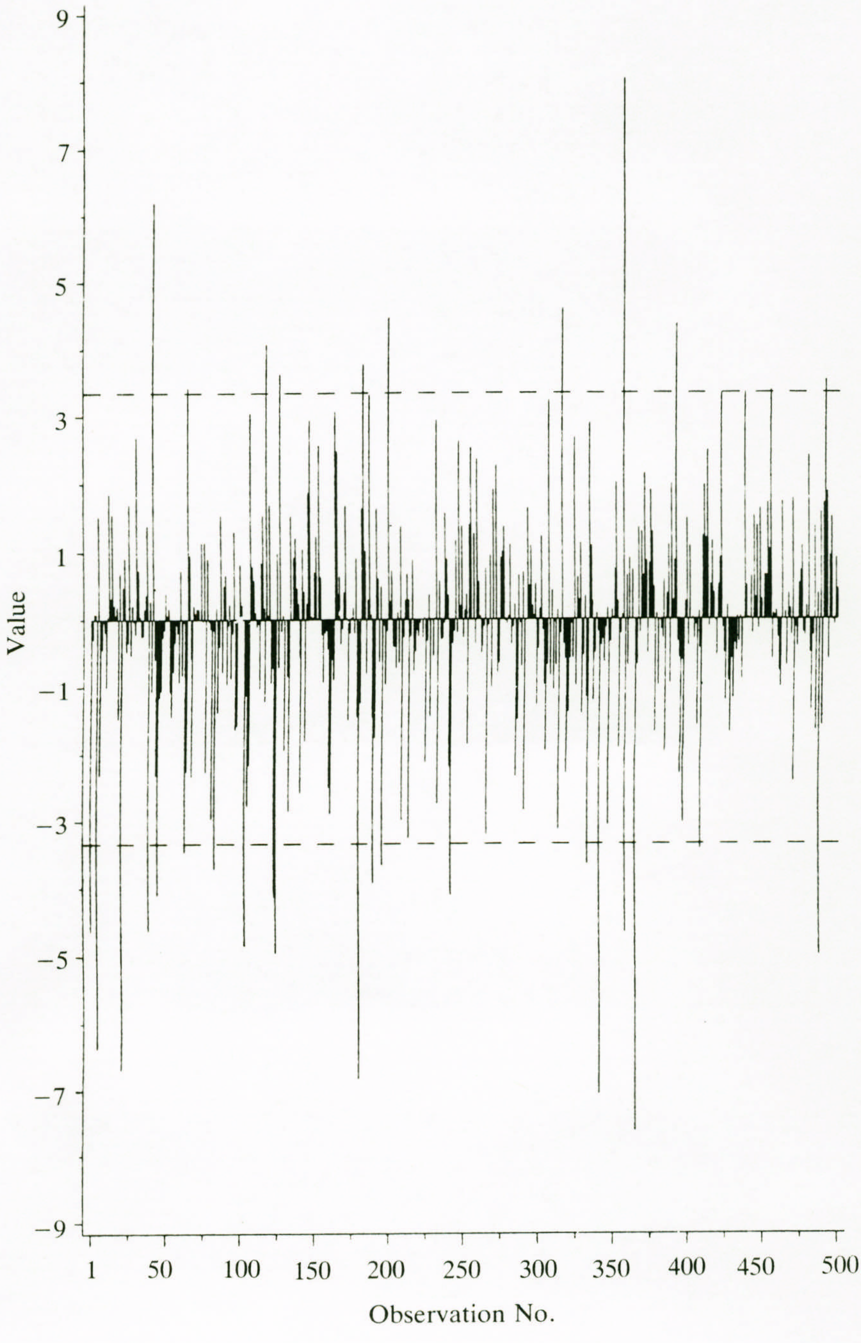

FIGURE 1. Plot of the realization of the 15-term martingale difference process. The broken reference lines are two standard deviations above and below the zero mean of the process.

TABLE I

5% AND 1% SIZE OF TEST FOR THE 15-TERM MARTINGALE MODEL AS A FUNCTION OF THE WIDTH OF THE V-NOTCH

	$V = 1$	$V = 2$	$V = 3$	$V = 4$	$V = 5$	$V = 6$	$V = 7$	$V = 8$
(a) $N = 3844$ and $\sigma = 0.3$								
5%	30.5%	17.0%	9.0%	5.0%	2.5%	1.0%	0.5%	0.0%
1%	20.0%	8.5%	4.0%	1.5%	1.0%	0.5%	0.0%	0.0%
0.90	2.95	2.03	1.49	0.96	0.26	−0.35	−1.10	−1.85
0.95	3.78	3.08	2.27	1.66	0.89	0.28	−0.31	−1.07
0.99	6.16	4.16	3.61	3.09	2.44	1.76	1.01	0.25
Min	−2.63	−3.48	−4.27	−4.84	−5.48	−6.01	−6.55	−7.22
Max	6.35	5.33	4.61	4.11	3.46	2.85	2.27	1.62
(b) $N = 3844$ and $\sigma = 0.5$								
5%	69.0%	54.0%	32.0%	23.0%	13.0%	9.0%	5.0%	1.5%
1%	57.5%	36.0%	21.5%	13.0%	9.0%	5.0%	1.5%	1.0%
0.90	4.97	4.17	3.48	2.78	2.17	1.49	0.82	0.09
0.95	6.49	5.79	4.76	4.02	3.11	2.44	1.80	1.01
0.99	9.47	6.76	6.13	5.39	4.74	4.06	3.29	2.54
Min	−2.33	−3.01	−3.60	−4.19	−5.00	−5.62	−6.21	−6.93
Max	9.59	8.50	7.64	6.95	6.21	5.57	4.94	4.22
(c) $N = 10\,000$ and $\sigma = 0.3$								
5%	32.5%	14.0%	4.0%	3.0%	2.0%	0.5%	0.5%	0.0%
1%	16.5%	6.5%	3.0%	1.5%	0.5%	0.5%	0.0%	0.0%
0.90	2.74	1.89	1.18	0.59	−0.11	−0.73	−1.41	−2.08
0.95	3.32	2.49	1.56	0.97	0.30	−0.36	−0.94	−1.55
0.99	5.01	4.19	3.30	2.59	1.87	1.17	0.51	−0.19
Min	−2.06	−2.85	−3.49	−4.10	−4.86	−5.63	−6.34	−6.96
Max	5.87	5.31	4.53	3.96	3.26	2.54	1.73	0.98
(d) $N = 10\,000$ and $\sigma = 0.5$								
5%	74.4%	56.8%	38.7%	20.6%	10.6%	6.0%	3.0%	1.5%
1%	59.8%	39.7%	18.6%	12.1%	6.5%	2.5%	1.5%	1.0%
0.90	4.96	4.09	3.32	2.59	1.89	1.15	0.44	−0.31
0.95	5.61	4.93	4.09	3.40	2.73	2.08	1.25	0.49
0.99	8.15	7.57	6.83	6.24	5.47	4.72	3.90	3.12
Min	−1.78	−2.85	−3.49	−4.07	−4.77	−5.48	−6.13	−6.75
Max	9.68	8.75	7.70	6.81	5.82	4.99	4.14	3.33

The frame lengths are 62 and 100 for $N = 3844$ and $N = 10\,000$ respectively. The 0.90, 0.95 and 0.99 entries show the indicated fractile of the distribution of the z values of the test statistic.

strong evidence against the null hypothesis since the test statistic is shrunk toward zero.

A convenient method for displaying the estimated adjusted (filtered) and unadjusted bispectra is through contour plots of the probability that the $\mathrm{CH}(j, k)$ are not zero (see Equation (3.4)). Figure 2 is a plot of the unadjusted bispectrum (i.e. before applying the martingale filter) for a typical realization of Model 1 with $\sigma = 0.5$, $N = 3844$ and a frame length of 62. Contours are plotted for probabilities of 50%, 80%, 90% and 95%. Note that there are a number of peaks at or above the 95% level. Figure 3 shows the probability levels for the estimated bispectrum of the same series after

TABLE II

THE POWER OF THE TEST FOR THE 15-TERM NON-MARTINGALE MODEL WIDTH OF THE V-NOTCH VARIED BETWEEN 1 AND 8

	$V = 1$	$V = 2$	$V = 3$	$V = 4$	$V = 5$	$V = 6$	$V = 7$	$V = 8$
(*a*) $N = 3844$ *and* $\sigma = 0.3$								
5%	100.0%	99.5%	96.5%	92.0%	76.5%	57.0%	30.0%	8.0%
1%	100.0%	98.0%	91.5%	78.5%	60.5%	36.0%	14.0%	2.5%
0.90	7.69	6.78	5.99	5.27	4.53	3.66	2.62	1.44
0.95	8.06	7.26	6.37	5.56	4.83	4.01	2.97	1.97
0.99	9.31	8.49	7.77	7.07	6.16	5.02	3.94	2.90
Min	2.35	1.51	0.91	0.34	−0.33	−1.14	−2.25	−3.29
Max	10.40	9.40	8.41	7.55	6.83	5.85	4.43	2.97
(b) $N = 3844$ *and* $\sigma = 0.5$								
5%	99.0%	97.5%	87.5%	72.5%	54.0%	30.5%	13.5%	2.5%
1%	97.5%	88.5%	76.0%	56.5%	35.5%	18.5%	6.5%	1.0%
0.90	7.34	6.28	5.40	4.55	3.69	2.76	1.81	0.87
0.95	7.81	6.73	5.93	5.19	4.38	3.49	2.48	1.48
0.99	9.16	7.58	6.78	6.04	5.22	4.24	3.30	2.38
Min	1.26	0.00	−0.73	−1.41	−2.11	−2.91	−3.76	−4.63
Max	9.19	8.25	7.55	6.79	5.85	4.85	3.78	2.59
(c) $N = 10\,000$ *and* $\sigma = 0.3$								
5%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
1%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	99.5%
0.90	15.20	14.19	13.32	12.58	11.58	10.45	9.06	7.46
0.95	15.66	14.62	13.83	13.00	11.96	10.72	9.22	7.85
0.99	16.91	15.85	14.99	14.21	13.19	11.92	10.46	9.04
Min	9.75	8.60	7.54	6.74	5.88	4.99	3.76	2.24
Max	17.39	16.45	15.58	14.81	11.84	12.74	11.40	9.76
(d) $N = 10\,000$ *and* $\sigma = 0.5$								
5%	100.0%	100.0%	100.0%	100.0%	100.0%	99.5%	95.5%	80.5%
1%	100.0%	100.0%	100.0%	100.0%	99.5%	98.0%	87.5%	65.0%
0.90	12.01	10.83	9.95	9.12	8.21	7.10	5.87	4.56
0.95	12.30	11.32	10.42	9.59	8.53	7.40	6.27	5.01
0.99	13.94	12.81	11.93	11.14	10.21	9.06	7.81	6.48
Min	5.80	4.88	3.86	3.08	2.22	1.24	0.22	−0.80
Max	13.98	12.99	12.10	11.26	10.34	9.41	8.30	6.86

The frame lengths are again 62 and 100 for $N = 3844$ and $N = 10\,000$ respectively. The 0.90, 0.95 and 0.99 entries show the indicated fractile of the distribution of the z values of the test statistic.

filtering with a V-notch width of 6. The tapered martingale filter has removed all the 95% peaks and most of the 90% peaks seen in Figure 2.

6. THE TEST APPLIED TO STOCK RETURN DATA

Martingale difference processes are of interest to economists because they provide a mathematical model of a fair game. In this section the test will be applied to a real financial time series which, according to economic theory,

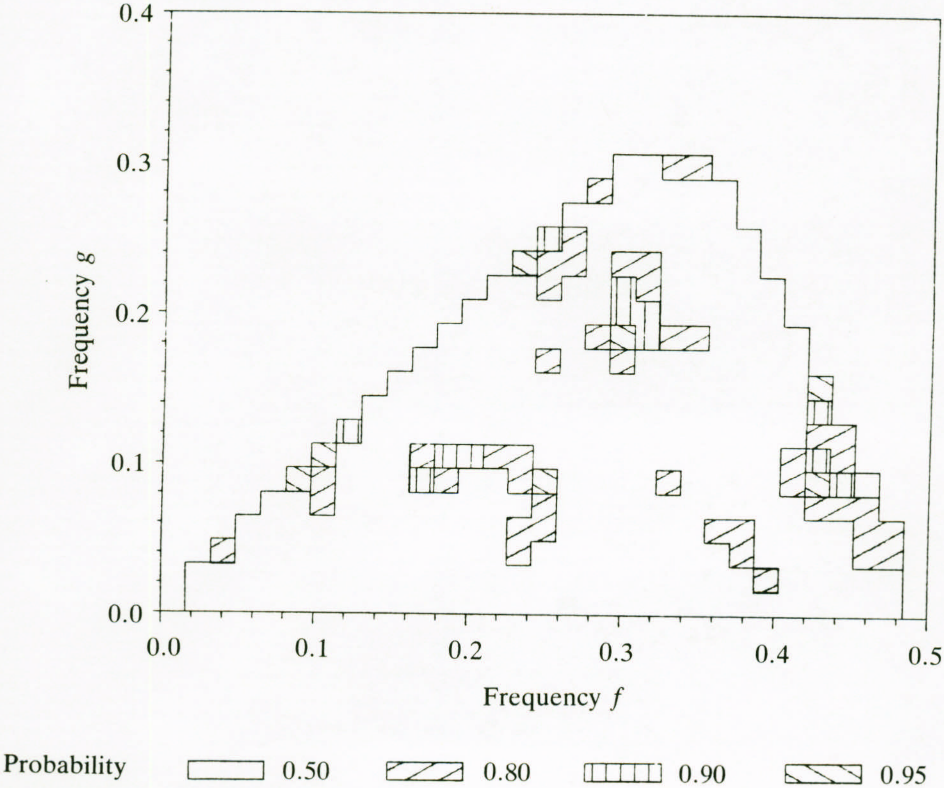

FIGURE 2. Probability that the estimated bispectrum for the 15-term martingale difference before filtering is not zero.

should behave as a fair game. The series is the 'unanticipated' component of the daily stock return to the General Electric Corporation (GE).

A brief sketch of the economic theory follows. The reader with more than a passing interest in the economic arguments is directed to the papers by Samuelson (1965), Fama (1970), LeRoy (1973) and Lucas (1978). Consider the intertemporal behaviour of the price per share of a common stock issue. Denote the price per share at the close of trading on day t as P_t. According to the theory of stock price formation, today's price P_t is proportional to a consensus forecast of tomorrow's price P_{t+1} predicted by currently available information. Next, define the continuously compounded rate of return from day t to day $t+1$ as $r_{t+1} = \ln(P_{t+1}/P_t)$. Given P_t and the market's forecast of P_{t+1}, there are techniques for estimating the market's implicit prediction of the next period's rate of return $\hat{r}_t(1)$, where $\hat{r}_t(1)$ denotes a one-period-ahead prediction of the return made at time t. Finally, define the unanticipated return e_{t+1} as the difference between the realized return and the predicted return: $e_{t+1} = r_{t+1} - \hat{r}_t(1)$. The economic theory states that e_{t+1} is non-forecastable given today's information set, i.e. e_{t+1} must follow a martingale

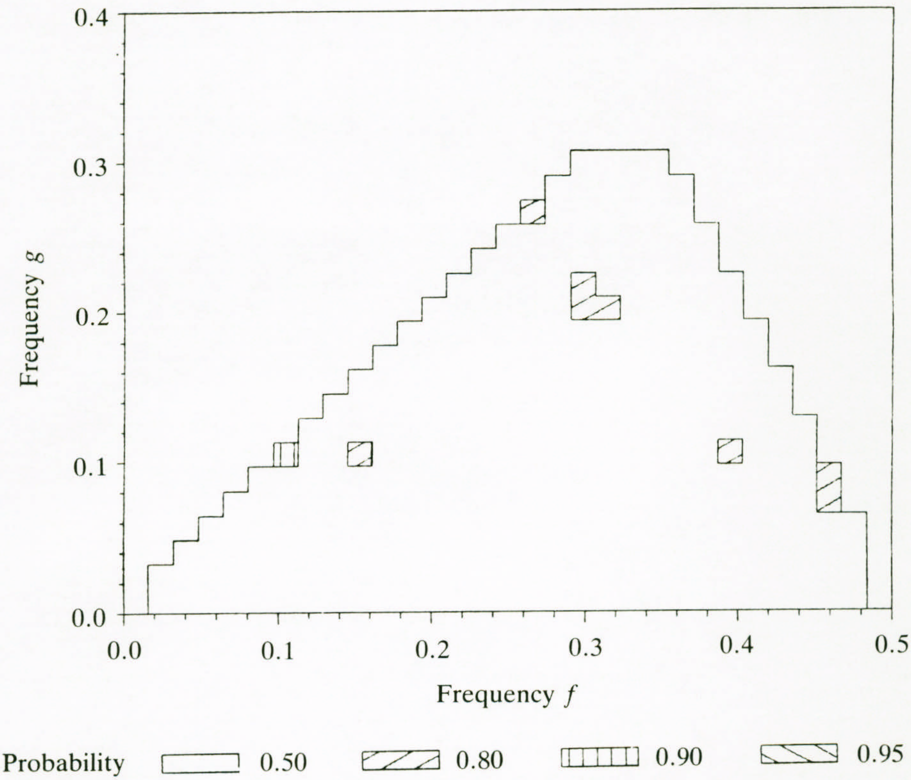

FIGURE 3. Probability that the estimated bispectrum for the 15-term martingale difference after filtering is not zero.

difference process. The theory argues that, if this were not the case, economic agents, as a consequence of their trying to exploit a forecast of e_{t+1}, would cause today's price to change in a direction which would drive the profits implied by the forecast to zero.

Estimates of the unanticipated returns e_{t+1} are contained in a database available from the Center for Research in Security Prices (CRSP) at the University of Chicago called the 'CRSP Daily Excess Returns File'. This state-of-the-art database contains the daily returns for every stock listed on the New York Stock Exchange or the American Stock Exchange in 'excess' of an estimate of its expected, or forecast, return. In other words, the database uses an instrument for the forecast return.

The period chosen for analysis of the GE excess returns was 18 September 1973 through 30 December 1983. The length of the time series is 2600 days. Figure 4 is a plot of the estimated autocorrelation function for the GE excess returns. The horizontal broken lines in the plot show the two standard error levels for the estimates. The excess returns for GE appear to correspond

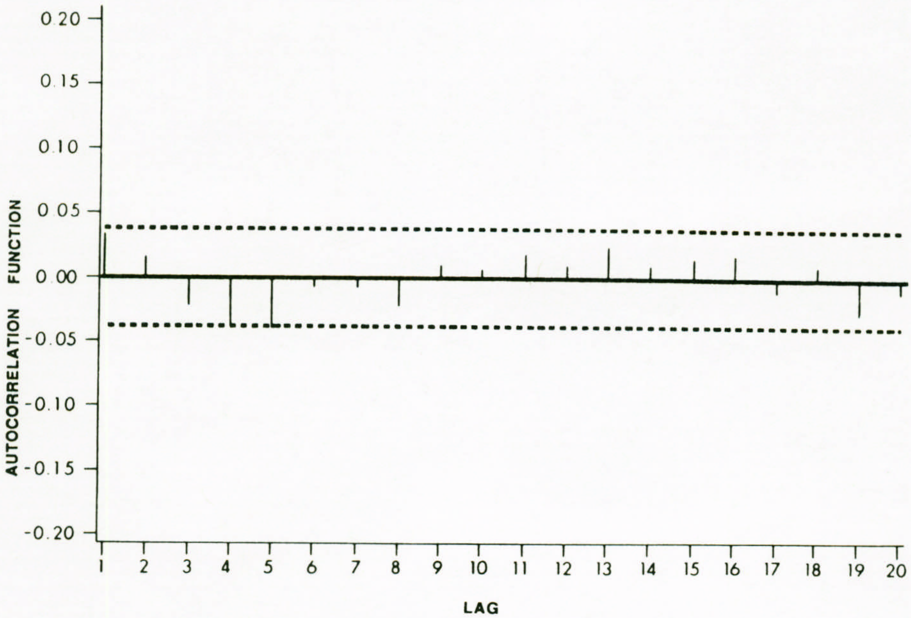

FIGURE 4. Autocorrelation function for GE excess returns.

nicely to white noise, a necessary condition under the fair game hypothesis. Although the correllogram is not shown, the realized or raw returns r_t exhibit significant autocorrelation at lags 1 and 2—this correlation is not necessarily inconsistent with theories of stock price formation (see, for example, Lucas, 1978).

The V widths, expressed as a percentage of the frame length, were 14%, 20%, 26%, 32%, 38% and 42%. Seven frame lengths were used for each V:NF = 44, 46, 48, 50, 52, 54 and 56. We calculated the average value of the martingale test statistic for each V setting, i.e. we averaged the test statistic over the NF values used for a particular V percentage. Performing the test on the 2600 observations using the PC version of the MARTIN program and running on a machine equipped with an 80386 central processor and a 20 MHz clock requires about 60 s for each V setting.

Figure 5 is a contour plot of the probability that the estimated unadjusted bispectrum is not zero for the excess returns before applying the martingale filter and using a frame length of 50. Loosely speaking, the peaks in Figure 5 show that the time series is neither Gaussian nor linear. Figure 6 is a contour plot of the probabilities for the estimated adjusted bispectrum, i.e. the estimates after applying a martingale filter with $V = 13$ (26%). Although the height of many of the peaks has been reduced by the martingale filter, they have not been reduced sufficiently for us to fail to reject the null of a martingale difference. Again, under the null hypothesis that the excess

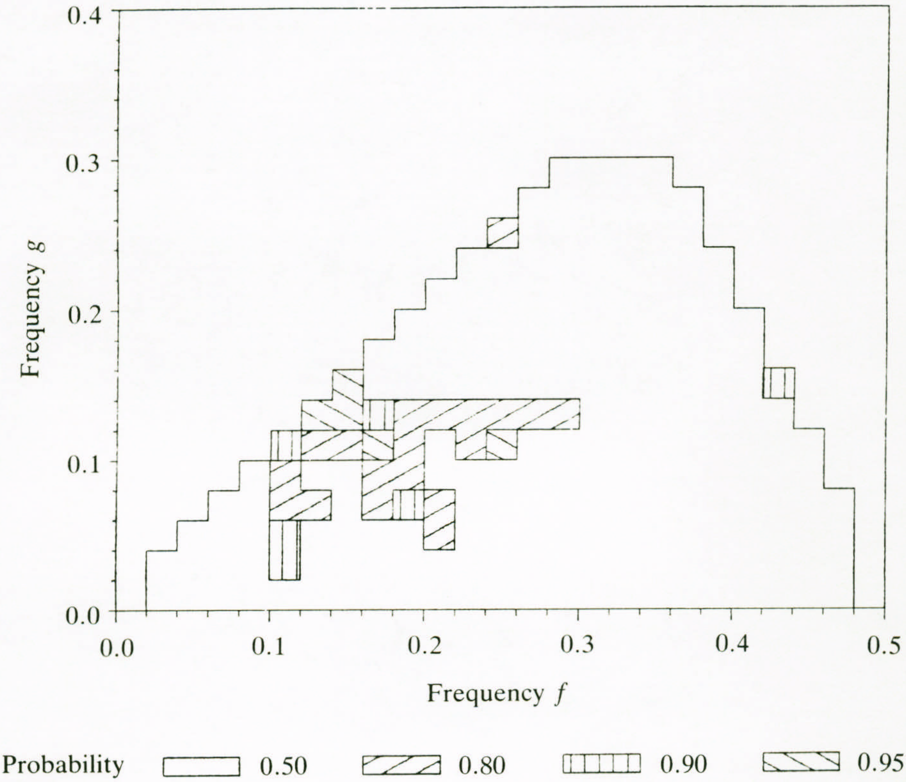

FIGURE 5. Probability that the estimated bispectrum for GE excess returns before filtering is not zero.

returns follow a martingale difference process, the adjusted bispectrum should be zero. The martingale test statistic, when expressed as a standardized normal variable, has a computed value of 11.35 for this particular combination of NF and V. Hence the test statistic is about 11 standard errors greater than its expected value under the null hypothesis.

Raising the V-notch level reduces the test statistic, but the null hypothesis is rejected for all the V-notch settings tried. The average z values over the indicated NFs were as follows: $\bar{z} = 19.6/V = 14\%$, $\bar{z} = 16.2/V = 20\%$, $\bar{z} = 12.5/V = 26\%$, $\bar{z} = 8.6/V = 32\%$, $\bar{z} = 4.92/V = 38\%$ and $\bar{z} = 2.6/V = 42\%$. The rejection of the martingale difference hypothesis is clearly not marginal.

We should point out that the strong rejection of the null hypothesis is not a peculiarity of the GE excess returns. This result is typical of other stocks we have analyzed, although in this particular application we are, in fact, testing two hypotheses: (i) the instrument properly measures the market's forecast of next period's return, and (ii) unanticipated returns follow a martingale

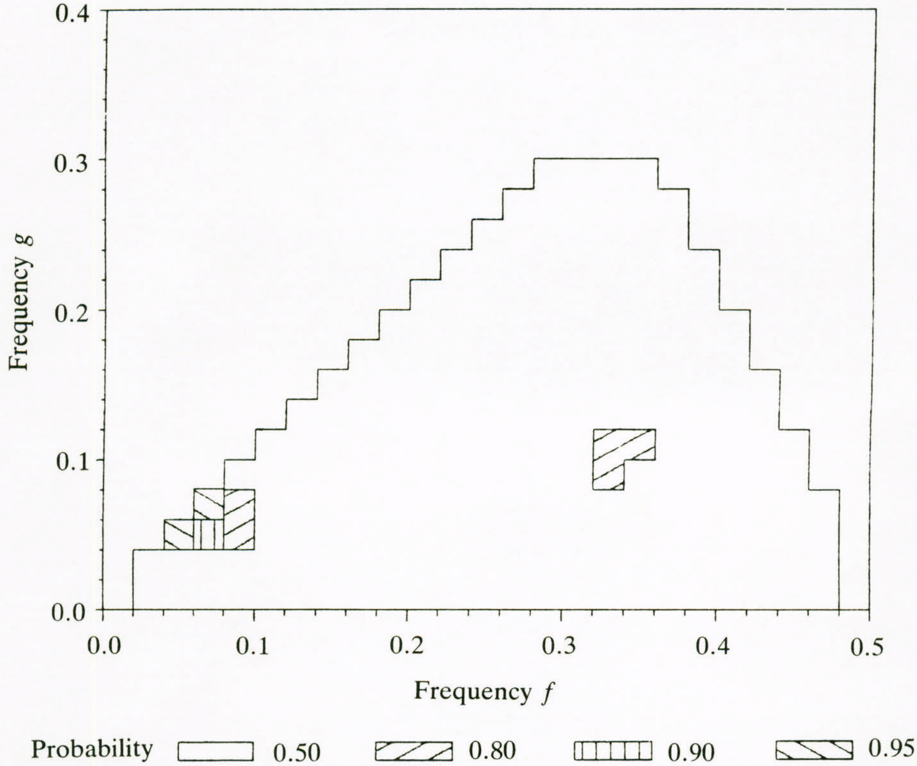

FIGURE 6. Probability that the estimated bispectrum for GE excess returns after applying the martingale filter is not zero.

difference process. Hence the rejection of the null hypothesis could be evidence that daily returns are not generated by the linear model proposed in economic theory. In fact, Hinich and Patterson (1985) have used a bispectral linearity test to reject the hypothesis that daily stock returns are generated by a linear stochastic process. As a consequence, we wonder whether any instrument based on linear methods will provide unanticipated returns which will pass the martingale difference test.

ACKNOWLEDGEMENTS

MJH is the Hogg Professor of Government, The University of Texas at Austin. He was supported by a contract from the Office of Naval Research. DMP was supported by Contract N60921-83-G-A165-B005 from the Naval Surface Weapons Center. Both authors wish to thank David Brillinger, Pat Brockett, Hugh McCulloch, Brian Roberts, Paul Shaman, Warren Weber and two anonymous referees for providing them with many insightful comments and helpful suggestions.

REFERENCES

BILLINGSLEY, P. (1979) *Probability and Measure*. New York: Wiley.

BRILLINGER, D. R. (1975) *Time Series: Data Analysis and Theory*. New York: Holt Rinehart and Winston.

—— and ROSENBLATT, M. (1967a) Asymptotic theory of kth order spectra. In *Spectral Analysis of Time Series* (ed. B. Harris). New York: Wiley, pp. 153–88.

—— and —— (1967b) Computation and interpretation of kth order spectra. In *Spectral Analysis of Time Series* (ed. B. Harris). New York: Wiley, pp. 189–232.

DOOB, J. L. (1983) *Stochastic Processes* New York: Wiley.

FAMA, E. F. (1970) Efficient capital markets: a review of theory and empirical work. *J. Finance* 25, 383–417.

HINICH, M. J. (1982) Testing for Gaussianity and linearity of a stationary time series. *J. Time Ser. Anal.* 3(3), 169–76.

—— and CLAY, C. (1968) The application of discrete Fourier transform in the estimation of power spectra, coherence and bispectra of geophysical data. *Rev. Geophys.* 6, 347–63.

—— and PATTERSON, D. M. (1985) Evidence of nonlinearity in daily stock returns. *J. Bus. Econ. Statist.* 3, 69–77.

HUBER, P., KLEINER, B., GASSER, T. and DUMMERMUTH, G. (1971) Statistical methods for investigating phase relations, *IEEE Trans. Audio Electroacoust.* 19–20, 78–86.

LEROY, S. F. (1973) Risk aversion and the martingale property of stock prices. *Int. Econ. Rev.* 14, 436–46.

LUCAS, R. E. (1978) Asset prices in an exchange economy. *Econometrica* 46(6), 1429–45.

PAPOULIS, A. (1965) *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill.

PATTERSON, D. M. (1983) BISPEC: a program to estimate the bispectrum of a stationary time series. *Amer. Statist.* 37, 323–24.

ROSENBLATT, M. (1983) Cumulants and cumulant spectra. In *Handbook of Statistics*, Vol. 3 (eds D. Brillinger and P. Krishnaiah). Amsterdam: North-Holland, pp. 369–82.

—— (1985) *Stationary Sequences and Random Fields*. Boston, MA: Birkhauser.

—— and VAN NESS, J. (1965) Estimation of the bispectrum. *Ann. Math. Statist.* 36, 1120–36.

SAMUELSON, P. A. (1965) Proof that properly anticipated prices fluctuate randomly. *Ind. Manag. Rev.* 6, 41–49.

SUBBA RAO, T. (1983) The bispectral analysis of nonlinear time series with reference to bilinear time series models. In *Handbook of Statistics*, Vol. 3 (eds D. Brillinger and P. Krishnaiah). Amsterdam: North-Holland, pp. 293–319.