A New Test for Randomness with Application to Stock Market Index Data

Alicia G. Strandberg[1], Boris Iglewicz[2]
[1,2] Department of Statistics, The Fox School of Business, Temple University, Philadelphia, PA 19122

Abstract
Strandberg and Iglewicz (2012) propose a test that detects deviations from randomness, without an a priori distributional assumption. This nonparametric test is designed to detect deviations of neighboring observations from randomness, especially when the dataset consists of time series observations. The proposed test is especially effective for larger datasets. In our simulation study, this test is compared to a number of variance ratio and traditional statistical tests. The proposed test is shown to be a competitive alternative for a diverse choice of distributions and data models. In addition, this test is able to successfully detect changes in variance, which can be informative in short term investing and option trading. In our empirical application, we review and compare several transformations while evaluating common US stock market indices. We consider two commonly used transformations and a proposed new transformation from Strandberg and Iglewicz (2012). This new transformation performs surprising well for stock market index data and is the only transformation to show consistence results among the considered tests.

Key Words: Nonparametric, Randomness, Stock Market Indices, Time Series, Variance Ratio

1. Introduction

A new test, proposed by Strandberg and Iglewicz (2012), is considered for studying departures from randomness of a series of independently and identically distributed (*i.i.d.*) random variables. There are many existing tests used to determine if a time series consists of a random sample. Often these tests have restrictive distributional assumptions, size distortions, or low power. The interest here lies in developing an alternative test with minimal assumptions and a simple test statistic to determine whether a series consists of a random sample. A test with minimal assumptions is an important consideration since data distributions are often unknown. Therefore, developing a test that does not require a normality assumption or a priori knowledge of the distribution generating these data is highly desirable.

The proposed test is shown to work well for varied symmetric and skewed distributions with reliance on an upper and lower percentile and straightforward conditional binomial probability, without an a priori distributional assumption. The test is then empirically compared with two popular existing tests and shown to be nicely competitive. In our comparison varied sample sizes and distributional models are considered. The proposed test is also evaluated in an applied application using stock market index data. These data are not normally distributed; rather they are nonstationary with typically an increasing long term trend. These data are atypical because such data often consists of structural breakdowns where a relatively large number of adjacent time series observations follow a regular pattern followed by highly irregular periods with changes in mean and variance (Chu, Stinchcombe and White 1996, Bandyopadhyay, Biswas, and Mukherjee 2008). Even after making a practical transformation the distributions of these stock market index datasets are often noticeably skewed. In addition, a new data transformation, also proposed by Strandberg and Iglewicz (2012), is considered. This transformation is a modified measure of percent change (MMPC), similar to a stock return. In our stock index data analysis study, the proposed transformation was the only considered transformation that led to consistently rejecting the null hypothesis with high power for all considered tests and stock market indices.

2. Description of Tests

Our motivating application is daily closing values from stock market index data, which consist of large number of observations. Large sample datasets are becoming more common in other statistical areas, such as regression and linear and nonlinear models. There are many tests that can be used to determine whether a series, such as daily changes in index data, consists of a random sample. These tests include both classical statistics methods and methods popular in the economic and finance literature, such as variance ratio tests. In this paper we compare the proposed test with the traditional Durbin and Watson (1950) test and the most referenced of the variance ratio tests found in the literature, Lo and MacKinlay (1988).

2.1 Strandberg and Iglewicz (2012)

The test by Strandberg and Iglewicz (2012) is based on a straightforward conditional binomial probability and test statistic that converges in distribution to the standard normal distribution. It tests for randomness of a time series based on a model that assumes *i.i.d.* random observations. Consider a time series, $Y_1, Y_2, ..., Y_N$ consisting of N observations, under the null hypothesis

$$Y_t = \mu_t + \varepsilon_t \tag{1}$$

where ε_t is $i.i.d.(0, \sigma^2)$, $Var[\varepsilon_t] = \sigma^2$ and $\mu_t = \mu$ for all *t*. Note that no additional distributional assumptions are made on ε_t in (1). Let $\mathbf{Y} = \{Y_1, Y_2, Y_3, ..., Y_N\}$. Notice that while we are dealing with a series, $\mathbf{Y}$ is a set. Consider the subsets $\mathbf{Y^*} = \{Y_t : Y_t \in R\}$ and $\mathbf{Y^{**}} = \{Y_t : Y_t \notin R\}$, where R is the simple interval $[Y_{\{0.025\}}, Y_{\{0.975\}}]$ and $Y_{\{p\}}$ is the p^{th} sample percentile. Note that $\mathbf{Y^{**}}$ is the complement of $\mathbf{Y^*}$. Other regions for R may be useful, but here we only consider the interval $[Y_{\{0.025\}}, Y_{\{0.975\}}]$. Concentrating on observations outside upper and lower percentiles allows this test to focus on tail observations, which are of special interest in stock market investigations, rather than on the entire dataset. While this choice of percentile works well in terms of providing the appropriate significance level, note that it is not directly related to the common significance level of $\alpha = 0.05$. Assume that the null hypothesis, that these observations constitute a random sample, is true. Let $N(\mathbf{Y^{**}})$ be the number of elements of $\mathbf{Y^{**}}$ and $P(Y_t \in \mathbf{Y^{**}}) = \pi$. We can estimate π by $\hat{\pi} = N(\mathbf{Y^{**}})/N$ (Strandberg and Iglewicz 2012).

Next subdivide the N observations into M consecutive intervals each having an equal number of K observations, such that $M = \lfloor N/K \rfloor$ intervals, where $\lfloor \cdot \rfloor$ is the floor (largest integer) function and K is an integer that is small relative to N. If $N/K - M \neq 0$, then using M intervals leaves out K(N/K – M) = b observations. We then recommend ignoring the first b observations in the time series. These b observations are considered an incomplete group of size less than K. Since N is large relative to K, having b $\neq$ 0 should not be a problem. For our motivating example, we consider complete weeks consisting of K = 5 days; then N/K is an integer. For simplicity assume b = 0. For i = 1, 2, …, K, j = 1, ..., M, denote $Y_i^{(j)} = Y_{K(j-1)+i}$ as the i^{th} observation in the j^{th} interval. For j = 1, ..., M, let $W_j = \sum_{i=1}^{K} I(Y_i^{(j)} \in \mathbf{Y^{**}})$, where $I(\cdot)$ is the indicator function. Under the null hypothesis that $Y_1, ..., Y_N$ are *i.i.d.*, we have $W_1, ..., W_M$ *i.i.d.* Binomial (K, π). It follows that

$$P(W_j = 1 | W_j > 0) = \frac{K\pi(1-\pi)^{K-1}}{1-(1-\pi)^K}, \; j = 1, ..., M. \qquad (2)$$

We denote the conditional probability in (2) by D. This suggests we consider all intervals such that $W_j > 0$. Let $L = \sum_{j=1}^{M} I(W_j > 0)$ and $L_1 = \sum_{j=1}^{M} I(W_j = 1)$. Thus in L out of the M intervals, we have W_j greater than 0, or at least one observation within the j^{th} interval belongs to the set **Y****. Similarly there are L_1 out of the M intervals, where exactly one observation belongs to the set **Y****.

Notice that the j^{th} interval with $W_j = 1$ must also satisfy $W_j > 0$. Thus we can rewrite L_1 as $L_1 = \sum_{j=1}^{L} I(W_j = 1 | W_j > 0)$. Under the null hypothesis, it follows that $I(W_j = 1 | W_j > 0)$ is a Bernoulli random variable with success probability D = $P(W_j = 1 | W_j > 0)$. Consequently, L_1 follows Binominal (L, D). For large L, by central limit theorem, we know $\frac{(L_1/L)-D}{\sqrt{D(1-D)/L}} \xrightarrow{d} N(0,1)$ where "$\xrightarrow{d}$" means converge in distribution. For moderate L, the binomial approximation can be improved by using a correction factor with modified test statistic

$$Z = \frac{(L_1/L) - D + (cH/L)}{\sqrt{D(1-D)/L}} \qquad (3)$$

where $H = \begin{cases} 1 & if \; \frac{L_1}{L} \geq D \\ -1 & if \; \frac{L_1}{L} < D \end{cases}$ and $\frac{c}{L}$ is the useful correction factor needed to prevent the test from being too conservative (Strandberg and Iglewicz 2012). In this study we use c = 0.50 and K = 5 to mimic stock market data with full 5 day trading weeks.

2.2 Lo and MacKinlay (1988)

As an alternative, we considered the popular variance ratio test of Lo and MacKinlay (1988). Under the null hypothesis, the relation between observations is a random walk,

$$Y_t = \mu + Y_{t-1} + \varepsilon_t \qquad (4)$$

where Y_t is the value of a return at time t, μ is an unknown arbitrary drift parameter and ε_t is a disturbance term with $Var(\varepsilon_t) = \sigma_\varepsilon^2$. A random walk assumes the conditional mean and variance are linear in time, a condition that may not always be reasonable when considering stock market data.

Although, Lo and MacKinlay developed two variance ratio tests, M_1 and M_2, our simulation investigation, using the vrtest package of R (Kim 2010), showed that the simpler test, M_1, performs slightly better than M_2. As a result, our simulation results will only be reporting for M_1. Under the null hypothesis M_1 assumes ε_t are *i.i.d.* $N(0, \sigma_\varepsilon^2)$. The test statistic for M_1 is

$$M_1(k) = \frac{VR(k)-1}{\phi(k)^{1/2}} \tag{5}$$

where $\phi(k) = \dfrac{2(2k-1)(k-1)}{3kN}$, $VR(k) = \dfrac{\hat{\sigma}^2(k)}{\hat{\sigma}^2(1)}$, $\hat{\sigma}^2(k) = m^{-1}\sum_{t=k}^{N}(Y_t - Y_{t-k} - k\hat{\mu})^2$

with $m = k(N-k+1)(1-kN^{-1})$ for $k \geq 1$, and $\hat{\mu} = N^{-1}\sum_{t=1}^{N}\left(Y_t - Y_{t-1}\right)$, such that when $VR(k) = 1$ observations are serially uncorrelated; likewise when $VR(k) \neq 1$some autocorrelations between observations exist (Lo and MacKinlay 1988, Charles and Darné 2009). Note that $VR(k)$ is a variance ratio that compares the variance of the k-period return with k times the variance of the one-period return. In this study the common value of $k = 2$ is used. Under the null hypothesis, M_1 follows asymptotically the standard normal distribution. However the sampling distribution of this variance ratio test statistic is known to be skewed (Chen and Deo 2006, Charles and Darné 2009).

2.3 Durbin and Watson (1950)

A well known traditional test of randomness was introduced by Durbin and Watson (1950). They test for non-randomness in the residuals of an ordinary least squares regression equation with the test statistic,

$$d = \frac{\sum_{t=2}^{N}(e_t - e_{t-1})^2}{\sum_{t=1}^{N} e_t^2} , \tag{6}$$

where e_t is the t^{th} residual and N is the number of observations. This test assumes *i.i.d.* $N(0, \sigma^2)$ errors. The test statistic, d, is asymptotically normal with mean of 2 and variance of 4/N (Harvey 1990). When data are positively (negatively) serially correlated, d will have a value that tends to zero (four). Although this test was designed for residuals of least squares regression, it can be used to test time series data such as $Y_1, Y_2, Y_3, ..., Y_N,$ by substituting $Y_t = e_t$ in (6), assuming the assumption of *i.i.d.* $N(0, \sigma^2)$ is valid. In some applications, the restricted conditions of normality and serial correlation are limiting. When data are heavily skewed, this test does not meet the size requirement (Ali and Sharma 1993).

3. Stock Market Example

In this section we evaluate randomness of daily closing values of stock market indices using the Dow Jones Industrial Average (DJIA), the Standards & Poor's 500 (S&P 500), and the National Association of Securities Dealers Automated Quotation System (Nasdaq). We only consider data for full 5 day weeks, with data ending on December 18, 2009. These data were also analyzed by Strandberg and Iglewicz (2012, 2013). We consider three different transformations, which include two common transformations and a newly proposed transformation by Strandberg and Iglewicz (2012). Consider the following transformations: lag 1 closing daily stock market index differences, $Y_t - Y_{t-1}$; daily percentage change defined as $100\left(\frac{Y_t(i)}{Y_{t-1}(i)} - 1\right)$ where $Y_t(i)$ is the daily closing price for index i on day t (Cizeau, Potters and Bouchaud 2001, Bandyopadhyay, Biswas and Mukherjee 2008, Mukherjee and Bandyopadhyay 2011), $i = 1, 2, 3$, for DJIA, S&P 500, and Nasdaq, respectively; the modified measure of percent change, MMPC, $100\left(\frac{Y_t(i)}{MA^{(i)}_{(t-2q:t-q-1)}} - 1\right)$ where MA is a delayed moving average of q observations such that

$$MA_{(t-2q:t-q-1)} = \frac{\sum_{j=t-2q}^{t-q-1} Y_t}{q}$$

, were, as suggested by Strandberg and Iglewicz (2012) , q = 10. It is reasonable to test whether these transformed observations constitute a random sample. Results are included in Table 1. To assist in the readability of Table 1, after each test statistics * is added if significance is at the 10% level, ** if significance is at the 5% level and *** if significance is at the 1% level. Extremely high test statistic values contain only *** with blanks for associated numbers.

Table 1: Test Statistics for Stock Market Indices Application Results

DJIA	*SI*	*DW*	*M_1*
$Y_t - Y_{t-1}$	-34.47***	2.11***	-7.44***
$100((Y_t/Y_{t-1})-1)$	-16.79***	1.99	0.70
$100((Y_t/MA)-1)$	-36.64***	0.07***	***
S&P 500			
$Y_t - Y_{t-1}$	-28.01***	2.11***	-6.32***
$100((Y_t/Y_{t-1})-1)$	-11.93***	1.94**	3.00***
$100((Y_t/MA)-1)$	-30.85***	0.08***	***
Nasdaq			
$Y_t - Y_{t-1}$	-18.71***	1.98	1.03
$100((Y_t/Y_{t-1})-1)$	-14.59***	1.87***	5.68***
$100((Y_t/MA)-1)$	-26.34***	0.06***	89.22***

The test statistics show *** if significant at the 1% level, ** if significant at the 5% level and * if significant at the 10% level. Test statistics greater than 100 or less than -100 are not shown since they are highly significant at the 1% level: SI = Strandberg and Iglewicz (2012) with c = 0.5, DW = Durbin and Watson (1950), M_1 = Lo and MacKinlay (1988) with k = 2

In Table 1 notice that there are some very high absolute test statistics resulting in p-values below 0.0001. Therefore exact p-values are not included in Table 1 because we do not claim that level of accuracy. In this table test statistics differ between transformations and indices, with the exception of Strandberg and Iglewicz, SI, which is significant at the 1% level for all indices and transformations. It is possible that stock market index data are mostly moderately stable with periods of time where values are instable. These unstable periods of time may result in far more outlier values or values outside the interval $[Y_{\{0.025\}}, Y_{\{0.975\}}]$. In cases like this, SI shows high power to reject the null hypothesis of random data.

Furthermore, **Z** values for SI are all negative, while **Z** values for M_1 are often positive. The negative **Z** values from SI indicate that more than expected multiple unusual transformed values tend to occur within weeks containing at least one unusual transformed value. The positive **Z** values for M_1 indicate that the variance of two time periods is higher than expected. Results from the Durbin and Watson test, DW, generally indicate that transformed indices values are positivity correlated – this occurs when d in (6) < 2.00.

Among transformations, only the MMPC results in consistently rejecting the null hypothesis at the 1% significant level for all tests, while the other two transformations show inconsistent rejection levels among tests and indices. It is interesting to observe that for DW and M_1 results can differ depending on the transformation used yet test results based on the MMPC transformation are all significant at the 1% level. Differences between the first two transformations include not only changes in significant levels, but also changes in interpretations - since for M_1 test statistics changes in value from negative to positive and for DW test statistics change from $d < 2.00$ to $d > 2.00$, where d is defined as in (6).

In summary, the only test that shows consistent results for each considered transformation is SI, while M_1 and DW show inconsistent results among transformations. In addition, the only transformation that shows consistent results among individual tests for each considered dataset is the MMPC transformation while all other transformations had inconsistent results among the considered indices. In summary, when dealing with financial data, the test used and choice of transformation can play a key role in resulting conclusions.

4. Simulation Results

In this Section we investigate each test over a diverse group of distributions and data models. We will view these tests as competitive, even though M_1 is based on a random walk model, thus not sensitive to changes in variance, while the other considered tests are based on a random sample null hypothesis. Since in practice most practitioners will not know the true distribution of their data, methods that perform well for a variety of distributions are preferred.

We will consider random samples of N = 300, and 10,000 observations and will use the g- and -h distributional family (Tukey 1977, Martinez and Iglewicz 1984, Hoaglin 1985) to obtain or approximate the standard normal, Z, distribution, the Student's t distribution with 3 degrees of freedom, t_3, and the chi-square distribution with 4 degrees of freedom, χ_4^2. The χ_4^2 observations are standardized before performing each test. These distributions are considered null cases since they are known distributions expected to preserve the test size.

We also consider some alternative data models. We considered a model where 30% of observations follow $f_1(y)$, then 40% follow $f_2(y)$, and the remaining 30% follow $f_3(y)$. A model with constant mean and changing variance is created by letting, $f_1(y)$ be $N(0, 0.75^2)$, $f_2(y)$ be $N(0, 1)$, and $f_3(y)$ be $N(0, 1.25^2)$, while a model with constant variance and changing mean is created by letting $f_1(y)$ be $N(-2, 1)$, $f_2(y)$ be $N(0, 1)$, and $f_3(y)$ be $N(2, 1)$. In addition, four correlated cases, C1, C2, C3 and C4 are considered. For these cases, motivated by stock market weekly data, observations are simulated from complete trading weeks where weeks with market closures or holidays are not included. Each correlated model is designed such that 90% of these weeks have observations that are *i.i.d.* N(0,1) and a correlation structure exists for the remaining 10% of these weeks. Consider a typical member of the 10% correlated weeks. Let Y^M be the Monday value coming from $N(0, 2^2)$. The other generated values are Y^j, j = T, W, Th, F. For C1, $Y^T = \rho Y^M + \varepsilon$ where ε is N(0,1) and the remaining three days are from *i.i.d.* N(0, 1). For C2, $Y^F = \rho Y^M + \varepsilon$ and the remaining three days are from *i.i.d.* N(0, 1). For C3, $Y^j = \rho Y^M + \varepsilon$, where j is randomly chosen from T, W, Th, F and the remaining three days are from *i.i.d.* N(0, 1). In C4 the correlation structure is present over two weeks, such that Y^M in the first week, Y_1^M, has a value from $N(0, 2^2)$. Two days are correlated observations, such that $Y^j = \rho Y^M + \varepsilon$, where j is randomly chosen twice with replacement from T_1, W_1, Th_1, F_1, M_2, T_2, W_2, Th_2, F_2, the remaining days are generated from *i.i.d.* N(0, 1). In all four correlated cases $\rho = 0.90$. More null and alternative cases are discussed in Strandberg and Iglewicz (2012, 2013).

Table 2 summarizes the simulation results. For each case 10,000 replications are simulated and the rejection rate is the percentage of tests out of 10,000 where p-value < 0.05. For the null cases, since the significance level is set at 5.00%, we expect rejection rates to be close to 5.00%. With 10,000 replications, the standard error of a 0.05 proportion of rejections, under the null hypothesis, is $\sqrt{(0.05\times 0.95)/10000} = 0.0022$ therefore rejection rates in the range of $(0.05 \pm 1.96\times 0.0022)\times 100 = [4.57\%, 5.43\%]$ are expected. All considered null cases are within this range and therefore are reasonably close to the desired 5.00% rejection rate except for SI when N is small. When N is small SI is slightly conservative.

For the alternative cases, as expected, greater power is generally seen as N increases; however power can differ considerably across alternatives. When N is small all considered tests, excluding SI, have high power to detect changes in mean, while all tests have low power for other cases. As N becomes large, the SI is the only test with high power for all alternative cases. Only one test, SI, is able to show considerable power against changing variances especially when N is large, while rejection rates for M_1 show low power. This is not surprising, as the null hypothesis for M_1 consists of a random walk, thus tolerating changes in variance. If detecting changes in variance is a concern, the SI test should be used.

The SI test is able to show high power for all the correlated cases with large N, while M_1 and DW have modest to low power for C2, C3 and C4. When N is large all tests show high power for C1 and modest power as N decreases. In summary, when N is large, SI is the only considered test with high power for all correlated cases. This test works well irrespective of the correlation structure following a highly volatile day, without a priori knowledge of future dependencies, while tests with set autocorrelation structures may not always be able to retain power in such situations.

Table 2: Simulation Rejection Rates Comparisons

	SI	*DW*	M_1
N = 300			
Z	4.21%	5.12%	5.18%
t_3	4.06%	4.94%	4.85%
X^2_4 Standardized	4.22%	4.78%	4.90%
Changing Variance	15.83%	6.87%	6.53%
Changing Mean	13.81%	100.00%	100.00%
C1	17.43%	22.81%	18.40%
C2	16.84%	6.28%	6.34%
C3	17.21%	7.50%	6.79%
C4	8.18%	7.01%	6.11%
N = 10,000			
Z	5.29%	4.86%	5.21%
t_3	4.96%	4.93%	5.07%
X^2_4 Standardized	4.91%	4.92%	4.96%
Changing Variance	95.32%	6.89%	7.43%
Changing Mean	99.24%	100.00%	100.00%
C1	99.83%	100.00%	100.00%
C2	99.85%	6.73%	7.10%
C3	99.74%	37.63%	36.15%
C4	65.96%	26.80%	25.40%

SI = Strandberg and Iglewicz (2012) test with c = 0.5, DW = Durbin and Watson (1950) test, M_1 = the Lo and MacKinlay (1988) with k = 2

5. Conclusion

In this paper we summarize and compare three tests for detecting randomness in time series data, including comparisons using stock market index data. One of our considered tests, SI is unique because it is based on a simple test statistic that surprisingly works well for varied null and alternative data choices with minimal distributional assumptions. Since this proposed new method is not based on autocorrelations or related measures, it is shown to have high power, for larger sample sizes, to detect correlated structures not only between consecutive observations but also over longer lags. Of the studied tests, only the SI test has power to detect changes in variance and all four correlated cases when N is large. Through simulations, it is also shown that SI meets the size requirements for a variety of null cases.

For our stock market index data analysis we review and compare three transformations and demonstrate that the choice of transformation can have a noticeable effect on test results. Only the MMPC transformation results in consistent rejection of the null hypothesis with high power for all tests and indices. While the SI test is the only test that is able to strongly reject the null hypothesis for all three studied indices and considered transformations. We note with interest that for the other considered tests, results can differ depending on the transformation used. The only test that does not show evidence of this concern is SI.

When deciding among tests, consideration must not only be given to meeting size requirements for the null hypothesis, but also the possible alternative hypotheses choices as power can differ considerably across alternatives. In our simulation study we considered a number of alternative cases to show the advantages and disadvantages of each of the studied tests. In summary, we believe that these comparisons and results will be helpful, especially when dealing with financial data and tests of randomness.

Acknowledgements

The authors would like to thank Drs. Pallavi Chitturi, and Yuexiao Dong, both from the Fox School of Business at Temple University, Philadelphia PA, for their time and helpful suggestions. They would also like to like to thank the referees and Editor from *Communications in Statistics- Simulation and Computation* for their helpful comments.

References

Ali, M. M., and Sharma, S. C. (1993), Robustness to Nonnormality of the Durbin-Watson Test for Autocorrelations, *Journal of Econometrics*, 57, 117-136.

Bandyopadhyay, U., Biswas, A., and Mukherjee, A. (2008). Controlling Type-I Error Rate in Monitoring Structural Changes Using Partially Sequential Procedures. *Communication in Statistics – Simulation and Computation* 37, 3:466-485.

Charles, A., and Darné, O. (2009), Variance-Ratio Tests of Random Walk: An Overview, *Journal of Economic Surveys*, 23, 503-527.

Chen, W. W., and Deo, R. S. (2006), The Variance Ratio Statistic at Large Horizons, *Econometric Theory*, 22, 206-234.

Chu, C. J., Stinchcombe, M., and White, H. (1996). Monitoring Structural Change, *Econometrica* 64:1045–1065.

Cizeau, P., Potters, M., and Bouchaud, J. (2001), Correlation Structure of Extreme Stock Returns, *Quantitative Finance*, 1, 217-222.

Durbin, J., and Watson, G. S. (1950), Testing for Serial Correlation in Least Squares Regression: I, *Biometrika*, 37, 409-428.

Harvey, A. C. (1990), *The Econometric Analysis of Time Series* (2nd ed.), Cambridge, MA: MIT Press.

Hoaglin, D. C. (1985), Summarizing Shape Numerically: The g-and-h Distributions, in *Exploring Data Tables, Trends, and Shapes: Robust and Exploratory Techniques,* Hoaglin, D. C., Mosteller, F., Tukey, J. W. (eds.), Hoboken, NJ: John Wiley & Sons, Inc., pp 461-513.

Kim, J.H. (2010), vrtest: Variance Ratio tests and other tests for Martingale Difference Hypothesis. R package version 0.95. http://CRAN.R-project.org/package=vrtest (accessed July 14, 2011).

Lo, A. W., and MacKinlay, A. C. (1988), Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test, *The Review of Financial Studies*, 1, 41-66.

Martinez, J., and Iglewicz, B. (1984), Some Properties of the Tukey g and h Family of Distributions, *Communications in Statistics-Theory and Methods*, 13, 353-369.

Mukherjee, A., and Bandyopadhyay, U. (2011). Some Partially Sequential Nonparametric Tests for Detecting Linear Trend. *Journal of Statistical Planning and Inference* 8: 2645-2655.

Strandberg, A.G., and Iglewicz, B. (2012), A Nonparametric Test for Deviation from Randomness with Applications to Stock Market Index Data. *Communications in Statistics- Simulation and Computation*, to Appear.

Strandberg, A.G., and Iglewicz, B. (2013), Detecting Randomness: A Review of Existing Tests with New Comparisons. *Communications in Statistics- Simulation and Computation*, to Appear.

Tukey, J. W. (1977), *Exploratory Data Analysis*, Reading, MA: Addison-Wesley.