Tests for weak form market efficiency in stock prices: Monte Carlo evidence

Mohammed S Khaled and Stephen P Keef

The Working Paper series is published by the School of Economics and Finance to provide staff and research students the opportunity to expose their research to a wider audience. The opinions and views expressed in these papers are not necessarily reflective of views held by the school. Comments and feedback from readers would be welcomed by the author(s).

Tests for weak form market efficiency in stock prices: Monte Carlo evidence

M. S. KHALED[*] AND S. P. KEEF[**]

School of Economics and Finance, Victoria University of Wellington,
Wellington, New Zealand

[*] Contact author, School of Economics and Finance, Victoria University of Wellington,
PO Box 600, Wellington, New Zealand (Mohammed.Khaled@vuw.ac.nz)

[**] School of Economics and Finance, Victoria University of Wellington,
PO Box 600, Wellington, New Zealand (stephen.keef@vuw.ac.nz)

Abstract

Efficiency in financial markets is tested by applying variance ratio (VR) tests, but unit root tests are also used by many, sometimes in addition to the VR tests. There is a lack of clarity in the literature about the implication of these test results when they seem to disagree. We distinguish between two different types of predictability, called 'structural predictability' and 'error predictability'. Standard unit root tests pick up structural predictability. VR tests pick up both structural and error predictability.

Key Words: Weak Form Efficiency, Unit Root, Random Walk, Autocorrelation, Variance Ratio.

INTRODUCTION

Efficiency in financial markets is typically formulated as unpredictability of market returns in excess of normal returns. If share market returns are to be predictable, profit-seeking investors will exploit the opportunity until the predictability disappears. This is known as the Efficient Market Hypothesis (EMH). Given a time series of stock prices p_t, the rates of return can be measured as $r_t = \ln p_t - \ln p_{t-1}$. According to weak form market efficiency, the current rate of return on the stock cannot be predicted by past experience, e.g., the past rates of return. Consider the regression of returns

$$r_t = \beta_0 + \beta_1 r_{t-1} + \beta_2 r_{t-2} + ... + \beta_k r_{t-k} + \varepsilon_t \tag{1}$$

where ε_t is an unpredictable error component. The EMH in its weak form is represented in this regression by the null hypothesis,

$$H_0 : \beta_1 = 0, \beta_2 = 0, ..., \beta_k = 0 \tag{2}$$

Under this hypothesis, returns $r_t = \beta_0 + \varepsilon_t$, and stock prices are related as

$$\ln p_t = \beta_0 + \ln p_{t-1} + \varepsilon_t \tag{3}$$

The series $\ln p_t$ in equation 3 has a unit root. A random walk process is a special case of a unit root process. It is where error terms are independently and identically distributed *iid* . In the context of the latter, the presence of serial correlation is the important dimension – it is a denial of weak form market efficiency. Serial correlation

reflects predictability. Other denials of *iid* in the errors, such as heteroscedasticity, do not reflect predictability per se.

Weak form market efficiency requires the absence of predictability using past information, that is, prices follow a unit root with uncorrelated errors.   Henceforth, the unit root process is called 'structural predictability' and the correlation in the errors is called 'error predictability'.

Unit roots are typically tested by an Augmented Dickey-Fuller (ADF) test or a Phillips-Perron (PP) test. These tests do not test for correlation in errors. Autocorrelation or its test consequences are indeed eliminated before they test for a unit root. The ADF test controls for autocorrelation by including an adequate number of the lagged dependent variable in the test regression. The PP test controls for both autocorrelation and heteroscedasticity by a non-parametric adjustment of the coefficient covariance matrix. Thus, the ADF/PP tests cannot strictly be viewed as a test of the efficient market hypothesis. Campbell, Lo and MacKinlay (1997: 65) note: '… *since there are also nonrandom walk alternatives in the unit root null hypothesis, tests of unit roots are clearly not designed to detect predictability* …'.

Lo and MacKinlay (1988) propose the variance ratio (VR) test as a test of random walk. However, the ADF/PP tests continue to be used on their own or alongside VR tests for the same purpose (e.g., Alimov *et al*. 2004; Asiri 2008; Hassan, Abdullah and Shah 2007; Murthy, Washer and Wingender 2011; Narayan and Smyth 2004; Ozdemir 2008).

4

In many cases though, the VR test results contradict the ADF/PP test results. Typically, VR tests reject a random walk in stock prices, but a unit root in stock prices is accepted by ADF and/or PP tests.

Lo and MacKinlay (1988: 44) characterize the VR test as being '… *sensitive to correlated price changes*'. In this paper, we offer Monte Carlo evidence to show that the extreme sensitivity of VR tests to correlated price changes is the main reason for the seemingly divergent results obtained by researchers using both the ADF/PP and the VR tests. If predictability of returns (non-random walk) arises entirely from an auto-correlated error component, the VR test of a random walk is just a test of autocorrelation. In this sense, a VR test will pick up both 'structural predictability' and 'error predictability'.

METHODOLOGY

The model $\ln p_t = \beta_0 + \ln p_{t-1} + \varepsilon_t$ is a random walk process (with drift) if the error ε_t is *iid* with mean zero and constant variance σ^2. In this case, the variance of the qth price change or difference ($\Delta_q \ln p_t = \ln p_t - \ln p_{t-q}$) increases linearly with q, i.e., $\text{var}(\Delta_q \ln p_t) = q \, \text{var}(\Delta \ln p_t)$. Hence, the variance ratio

$$VR(q) = \frac{\text{var}(\Delta_q \ln p_t)}{q \, \text{var}(\Delta \ln p_t)} \qquad (4)$$

will equal 1 for any q. If the price changes tend to revert to the mean in the long run, then $\ln p_t$ and $\ln p_{t-q}$ would be negatively correlated for large q. If so, the variance of

5

the qth difference will increase less than linearly with q, i.e., $VR(q)$ is likely to be smaller than 1 for any $q > 1$. A value of the ratio significantly below 1 denies a random walk process, i.e., the price is not weak form efficient.

Suppose we have a total of $T + 1$ asset prices p_t, i.e., a sample of T asset returns $\Delta \ln p_t$. The sample mean of the returns is $\hat{\beta}_0 = (\sum_{t=1}^{T} \Delta \ln p_t)/T$. The variance ratio (Equation 4) is then estimated as

$$VR(q) = \frac{\sum_{t=q}^{T}(\Delta \ln p_t + \Delta \ln p_{t-1} + ... + \Delta \ln p_{t-(q-1)} - q\hat{\beta}_0)^2}{q\sum_{t=1}^{T}(\Delta \ln p_t - \hat{\beta}_0)^2} \qquad (5)$$

Stock returns are frequently found to be auto-regressively heteroscedastic (ARCH). They are also typically found to be non-normal in distribution, with fat tails and a slight negative skew (i.e. more negative returns than positive ones). Lo and MacKinlay (1988) offer a test statistic of $VR(q)$ which is robust to both heteroscedasticity and non-normality. The test statistic is

$$z(q) = \frac{VR(q) - 1}{\sqrt{\theta(q)}} \qquad (6)$$

where

$$\theta(q) = \sum_{j=1}^{q-1} \left[\frac{2(q-j)}{q} \right]^2 \hat{\delta}_j$$

and

$$\hat{\delta}_j = \sum_{k=j+1}^{T}(\Delta \ln p_k - \hat{\beta}_0)^2 (\Delta \ln p_{k-j} - \hat{\beta}_0)^2 / \left[\sum_{k=1}^{T}(\Delta \ln p_k - \hat{\beta}_0)^2 \right]^2$$

In the case of mean reversion, the statistic $z(q)$ would be negative for large q. Under the null hypothesis of a unit root with uncorrelated changes, the statistic $z(q)$ is asymptotically standard normal for any q.

Typical choices of q are the even numbers increasing geometrically up to half of the sample size. Chow and Denning (1993) propose an appropriate multiple variance ratio test whereby the random walk hypothesis cannot be rejected at 5% significance only if $Max_q |z(q)|$ is smaller than $SMM(k,v,0.05)$, which is the critical value with a 5% upper tail area of the Studentized Maximum Modulus (SMM) distribution. The parameters of this distribution are the number of variance ratios being compared (k) and the degrees of freedom ($v = T - k$). In our discussions below, this version of the VR test will be referred to as the VR-SMM test. For $v = \infty$, the SMM critical values can be read off the standard normal table as $SMM(k,v,0.05)$ equals the standard normal critical value with an upper tail of $0.5[1 - (1 - 0.05)^{1/k}]$. For example, with seven choices of q as in our experiments below, $k = 7$ and $SMM(7, \infty, 0.05) = 2.68$.

MONTE CARLO EVIDENCE

Let the data generating process (DGP) for the prices be:

$$\ln p_t = \beta_0 + \ln p_{t-1} + \varepsilon_t \tag{7}$$

where

$$\varepsilon_t = \rho \, \varepsilon_{t-1} + x_t u_t \tag{8}$$

and

7

$$\ln x_t^2 = \lambda \ln x_{t-1}^2 + v_t \qquad\qquad (9)$$

with u_t and v_t being independent $N(0,1)$ random variables. The degree of serial correlation in errors ε_t is given by the coefficient ρ. Conditional on the x_t values, heteroscedasticity of the ARCH type is allowed in the error term ε_t to reflect the time varying volatility thought to be present in most financial time series.

The Monte Carlo results, presented in table 1, are based on four DGPs with heteroscedastic errors. Autocorrelation in the errors is either absent ($\rho = 0$) or set at one of the values $\rho = 0.05, 0.1$ or 0.2. The VR statistics are based on variance ratios calculated at the seven q values 2, 4, 8, 16, 32, 64 and 128. Since the maximum q is typically chosen as half of the sample size, a sample of 257 ($= 128 \times 2 + 1$) values of the heteroscedasticity-causing x variable is generated first, using equation 9 with $\lambda = 0.5$, $x_1 = 1$ and $v \sim N(0,1)$. Conditional on these x values, a sample of $\ln p$ values of size 257 can be obtained using equations 7 and 8, with $\beta_0 = 0.005$, $p_1 = 1000$ and $\varepsilon_1 = 0$. Such samples of $\ln p$ values are then replicated 10 000 times by random drawings of $u \sim N(0,1)$.

Insert table 1 about here

Given a nominal test size of 5%, empirical test sizes (i.e., percentage of rejection of a true null hypothesis in the given replications) are compared for the ADF, PP and VR tests. As our sample size is very close to 250, the 5% critical value for that size (-2.88) is used for the ADF/PP t-tests with a drift but no trend.

8

When errors are not correlated ($\rho = 0$), all three tests in table 1 are for a unit root only. Results for this case show that the ADF and PP tests have relative frequencies of type I error close to the nominal 5%, i.e., the true null hypothesis is rejected about as frequently as expected. However, the VR-SMM test is quite conservative with this type of error, its empirical test size being less than half the nominal size. For a given sample, reducing the probability of rejecting a true hypothesis increases the probability of accepting a false hypothesis. Hence, as a test of unit root, the VR-SMM test is less powerful than the ADF/PP tests.

More differences between the tests appear as autocorrelation increases even by small amounts (to 0.05, 0.1 or 0.2). The null hypothesis continues to be that of a unit root for the ADF and PP tests in these cases. The empirical test size remains below the nominal size at about 4% for the ADF test, but it gets smaller at higher autocorrelation for the PP test. Thus, the ADF test performs better than the PP test in terms of remaining true to the nominal test size.

For the VR-SMM test, the null hypothesis is that of a unit root with $\rho = 0$. Clearly, the hypothesis, which is false when $\rho = 0.05, 0.1$ or 0.2, is rejected with a rapidly rising frequency as autocorrelation in the errors increases. This is consistent with the interpretation that, even in the presence of a unit root, VR tests are able to reject a false hypothesis of unpredictability. As a test of unpredictability, this advantage is lost in the ADF/PP methods, which are designed to test for a unit only.

9

CONCLUSIONS

Our interpretations of the commonly used weak form market efficiency tests can be summarised as follows:

1. When a unit root in $\ln p_t$ is rejected (with or without auto-correlated errors in this series), so is EMH/random walk. The implication is that stock prices are predictable.

2. Acceptance of a unit root in $\ln p_t$ indicates absence of structural predictability. Absence of predictability in $\ln p_t$ arising from error predictability is not guaranteed. In other words, the presence of a unit root in stock prices is a necessary but not sufficient condition for weak form market efficiency.

3. It is the acceptance of a random walk (unit root with *iid* errors) that indicates absence of predictability of any kind as required by weak form market efficiency.

An ADF or a PP test of a unit root can rule on the matter of structural predictability alone. In contrast, a VR test of a random walk can assess predictability arising structurally and /or via auto-correlated errors. If stock prices are predictable, at least partly owing to auto-correlated errors, a VR test result may contradict an ADF/PP test result. For example, if there is no structural predictability but errors are correlated, an ADF/PP test will accept a unit root, while a VR test rejects a random walk. If testing weak form market efficiency is the goal, then the choice is either (i) a VR test or (ii) a unit root test combined with a further test of error autocorrelation.

REFERENCES

Alimov, A. A., D. Chakraborty, R. A. K. Cox and A. K. Jain (2004), "The random walk hypothesis on the Bombay stock exchange", *Finance India*, 18, 1251-1258.

Asiri, B. (2008), "Testing weak-form efficiency in the Bahrain stock market", *International Journal of Emerging Markets*, 3, 38-53.

Campbell, J. Y., A. W. Lo and A. C. MacKinlay (1997), *The Econometrics of Financial Markets*, Princeton University Press, Princeton.

Chow, K.V. and K. C. Denning (1993), "A simple multiple variance ratio test", *Journal of Econometrics*, 58, 385-401.

Hassan, A., M. S. Abdullah and Z. A. Shah (2007), "Testing of random walks and market efficiency in an emerging market: An empirical analysis of Karachi stock exchange", *The Business Review, Cambridge*, 9, 271-280.

Lo, A. W. and A. C. MacKinlay (1988), "Stock market prices do not follow random walks", *The Review of Financial Studies*, 1, 41-66.

Murthy, V., K. Washer and J. Wingender (2011), "Are stock prices in the US nonstationary? Evidence from contemporary unit root tests", *Applied Financial Economics*, 21, 1703-1709.

Narayan, P. K. and R. Smyth (2004), "Is South Korea's stock market efficient?", *Applied Economics Letters*, 11, 707-710.

Ozdemir, Z. A. (2008), "Efficient market hypothesis: evidence from a small open-economy", *Applied Economics*, 40, 633-641.

Table 1. Empirical test sizes with heteroscedastic price changes

Method	Null hypothesis	Serial correlation			
		$\rho = 0$	$\rho = 0.05$	$\rho = 0.1$	$\rho = 0.2$
ADF	Unit root	0.0427	0.0404	0.0408	0.0422
PP	Unit root	0.0503	0.0440	0.0378	0.0295
VR-SMM	Unit root and $\rho = 0$	0.0218	0.0479	0.1132	0.4323

Note: Nominal test size $= 0.05$