Available online at www.sciencedirect.com
ScienceDirect
Computational Statistics & Data Analysis 52 (2008) 2504–2513
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

Testing the random walk hypothesis through robust estimation of correlation

Andrei Semenov*

Department of Economics, York University, Vari Hall 1028, 4700 Keele Street, Toronto, Ont., Canada M3J 1P3

Received 7 February 2007; received in revised form 16 August 2007; accepted 18 August 2007
Available online 31 August 2007

Abstract

This paper uses Monte Carlo simulations to examine the properties of the conventional Pearson and some of the most well-known robust to outliers estimators of correlation in the presence of general heteroskedasticity. We show that the tests of a random walk based on the Pearson autocorrelation coefficient, including the Lo and MacKinlay [1988. Stock market prices do not follow random walks: evidence from a simple specification test. Rev. Financial Studies 1, 41–66] robust form of the variance-ratio test, can be unreliable in the presence of some forms of conditional heteroskedasticity. As an alternative to the Pearson autocorrelation coefficient, we propose the median coefficient of autocorrelation. Our simulation results show that, in contrast to the Pearson autocorrelation coefficient, the median coefficient of autocorrelation is robust to conditional heteroskedasticity. When applied to exchange rate returns, the variance-ratio test based on the median autocorrelation coefficient provides stronger evidence against the random walk hypothesis compared with the Lo and MacKinlay [1988. Stock market prices do not follow random walks: evidence from a simple specification test. Rev. Financial Studies 1, 41–66] robust variance-ratio test.

Keywords: Autocorrelation; Conditional heteroskedasticity; Correlation coefficient; Monte Carlo experiment; Random walk; Variance ratio

1. Introduction

One of the central questions in empirical finance and applied econometrics is the issue about whether the prices of financial assets evolve according to a random walk and thus cannot be predicted. Under the simplest version of the random walk, the increments, or first differences of prices, are independently and identically distributed (IID). A more general random walk model makes allowance for unconditional heteroskedasticity in increments and assumes that they are independent, but not identically distributed over time. The version of the random walk, which contains the first two random walk models as special cases, consists in including processes with dependent, but serially uncorrelated increments therefore allowing for conditional heteroskedasticity in increments.

To test the hypothesis of a random walk with IID increments some nonparametric tests such as the Spearman rank correlation and footrule tests and the Kendall correlation test, e.g., can be applied (e.g., Randles and Wolfe, 1979; Serfling, 1980). More recent tests fall into the class of semiparametric tests (e.g., Shorack and Wellner, 1986). Under parametric assumptions, statistics such as the likelihood-ratio statistic and the canonical correlation, e.g., can be used

* Tel.: +1 416 736 2100x77025; fax: +1 416 736 5987.
E-mail address: asemenov@econ.yorku.ca.

0167-9473/$ - see front matter
doi:10.1016/j.csda.2007.08.016

to construct tests of the null hypothesis that a time series is a random walk with IID increments (e.g., Muirhead, 1983). The filter rule developed by Alexander (1961, 1964) and technical analysis are widely used to test a random walk model, in which increments are independent, but not identically distributed.

A random walk model with dependent, but serially uncorrelated increments is typically tested using the Lo and MacKinlay (1988) robust to general heteroskedasticity form of the variance-ratio test. Wright (2000) however finds in Monte Carlo simulations that the robust form of the variance-ratio test suffers from size distortions in the presence of conditional heteroskedasticity. As an alternative, Wright (2000) proposes to use the variance-ratio tests based on the ranks and signs. Using Monte Carlo experiments, he finds that these tests are more powerful compared with the conventional Lo and MacKinlay (1988) heteroskedasticity-robust test with little or no risk of size distortions.

The use of the ranks and signs of returns rather than the raw returns, as proposed by Wright (2000), is, in fact, equivalent to replacing the Pearson autocorrelation coefficients in the equation for the variance ratio with the rank and sign autocorrelation coefficients, respectively. Within the approach proposed by Wright (2000), the robustness of the variance-ratio test to general heteroskedasticity is expected to be achieved through using a measure of autocorrelation that is robust to heteroskedasticity in returns rather than through correcting the variance of the Pearson autocorrelation coefficient for heteroskedasticity, as suggested by Lo and MacKinlay (1988).

Shevlyakov and Vilchevski (2002) use Monte Carlo simulations to investigate the behavior of the conventional Pearson and some of the most well-known robust to outliers estimators of correlation and find that the Pearson correlation coefficient as well as the rank and sign correlation coefficients are heavily biased in the case of an ε-contaminated bivariate normal distribution (the contamination scheme is described by the model of a mixture of normal distributions). Arguing that in the case of contamination the bias of an estimator is a more informative characteristic than its variance, they show that under contamination the best behavior is possessed by the median correlation coefficient (Shevlyakov, 1988; Shevlyakov and Lee, 1997). This estimator is asymptotically minimax with respect to the bias and has the maximum value of the breakdown point ($\varepsilon^* = \frac{1}{2}$).

The result in Shevlyakov and Vilchevski (2002) makes it interesting to examine the properties of the conventional Pearson, rank, sign, and median correlation coefficients under the assumption that increments exhibit unconditional and/or conditional heteroskedasticity. If, as in the case of contamination, the median correlation coefficient possesses the best properties, then it will be of interest to researchers to adapt this measure of correlation in tests of random walk based on the use of autocorrelation coefficients.

In this paper, we use Monte Carlo simulations to investigate the behavior of the conventional Pearson and some of the most well-known robust to outliers estimators of correlation (including the rank, sign, and median correlation coefficients) as well as the behavior of the variance ratios based on these estimators in the presence of unconditional and/or conditional heteroskedasticity in increments.

The plan of the remainder of this paper is as follows. In Section 2, we present various versions of the random walk hypothesis. In Section 3, we describe the approaches to testing the null hypothesis of random walk. The sampling distribution properties of the Pearson and some of the most well-known robust to outliers correlation coefficients as well as the properties of the variance ratios based on these coefficients are examined under unconditional and/or conditional heteroskedasticity in increments in the Monte Carlo simulations of Section 4. Section 5 reports the empirical results for exchange rate returns. Conclusions follow in Section 6.

2. The random walk hypothesis

The simplest form of the random walk hypothesis is the Random Walk 1 model, or RW1. In this model, the natural logarithm of prices, $p_t \equiv \ln P_t$, follows a random walk with IID increments:

$$p_{t+1} = \mu + p_t + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim IID(0, \sigma^2). \tag{1}$$

Independence implies that increments ε_t are serially uncorrelated and that any nonlinear functions of the ε_t's are also uncorrelated. The most common distributional assumption for ε_t is normality. Under this assumption, a series of the natural logarithm of prices is a random walk with normally distributed increments:

$$p_{t+1} = \mu + p_t + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim IID\,\mathrm{N}(0, \sigma^2). \tag{2}$$