Variance-Ratio tests and High-Frequency Data: A study of liquidity and mean reversion in the Indian equity markets.

Susan Thomas and Tirthankar Patnaik

October 23, 2002

Abstract

This paper tests for market efficiency at high–frequencies of the Indian equity markets by studying the behaviour of serial correlation in firm stock prices. We do this using the Variance Ratio test using returns data at a frequency of 5 minutes. We find that at this frequency interval, stocks show a pattern of mean–reversion. We also find that different stocks revert at different rates. Microstructure literature suggests that there is a correlation between the efficiency of a stock price and the liquidity of the stock on the market. We examine this hypothesis with liquidity measured in terms of both trading intensity as well as impact cost. We find strong evidence that there is a link between the liquidity of a stock and the patterns of serial correlations in its market price at high frequency.

Keywords: Variance-Ratios, High Frequency Data, Market Liquidity

Contents

1 Introduction

Tests of the "Efficient Market Hypothesis" (EMH) have focussed on testing the presence of serial correlation in financial market data. The presence of significant serial correlation indicates that prices are forecastable. This, in turn, implied that there may be opportunities for rational agents to earn abnormal profits if the forecasts are predictable after paying for transactions costs. Under the null of the EMH, serial correlations ought to be neglible. Indeed, most of this literature documents that small serial correlations exist in returns, which supports the hypothesis of market efficiency.

In the recent past, analysis of financial markets has become tranformed by the availability of intraday or *high frequency data*. This data was first available from foreign exchange markets, where data at frequencies of minutes could be observed all 24–hours of the day. Today, price and volumes data can be observed intervals from other asset markets that are electronic at frequencies as small as a second.

The analysis of serial correlations with high frequency data is particularly interesting for several reasons. One reason is that economic agents are likely to require time in order to react to opportunities for abnormal profits that appear in the market during the trading day. While the time required for agents to react may not manifest themselves in returns observed at a horizon of a day, we may observe agents taking time to react as patterns of forecastability in intra–day high–frequency data. For example, the market microstructure literature develops one class of models as based on asymmetric information between informed traders and uninformed traders. When a large order hits the market, there will be temporary uncertainty about whether this is a speculative order placed by an informed trader, or a liquidity-motivated order placed by an uninformed trader. This phenomenon could generate short-horizon mean reversion in stock prices.

Another reason why analysing high–frequency (HF) data would prove beneficial for market efficiency studies is that the abundance of data could yield highly powerful statistical tests of efficiency. These tests would be sensitive enough to reject subtle deviations from the null. However, the use of HF data also introduces problems such as those of asynchronous data, especially in cross-sectional studies. Since different stocks trade with different intensity, using HF data is constrained to not suffer too much of a missing data problem across all the stocks in the sample.

In this paper, we study the behaviour of serial correlation of HF stock price returns from the National Stock Exchange of India (*NSE*), using variance ratio (*VR*) tests. Since it's first application to financial data in Nelson and Plosser (1982), the VR has been constantly modified and improved as a robust test of serial correlation in financial time series. We apply a heteroskedasticity-consistent form of the VR to both the returns of the the NSE-50 index (*Nifty*) of the NSE, and to a set of 100 stocks that trade on the NSE. We choose those stocks which are the most liquid stocks in India, and therefore, have the least probability of missing data over our period of study, even at a five minute frequency.

The literature leads us to expect *positive deviations from the mean* in the index returns (**?**). This positive correlation in index returns could arise due to the asynchronous trading of the constituent index stocks: information shocks would first impact on the prices of stocks that are more liquid (and therefore, more actively traded) and impact on price of less active stocks with a lag. When the index returns are studied at very short time intervals, this effect ought to be even more severe as compared to the correlations in daily data with the positive serial correlations perhaps continuously growing for a period of time before returning to the mean (Low and Muthuswamy, 1996).

In our work however, we find that there is no evidence of any significant serial correlations in the index returns, even at an interval of five minutes. In fact, the VR at a lag of one is non-zero and *negative*. This implies that there is no asynchronous trading at a five–minute interval within the constituent stocks of the index. It would appear that to find evidence of asynchronous trading in the index, we would have to go to higher frequencies than five minutes.

We would expect to find *negative serial correlations* in stocks. This effect is attributed to the "bid–ask bounce" of traded prices (Roll, 1984), where the probability of a trade executing at the bid price being followed by a trade executing at the ask price is higher than a trade at the bid followed by another trade at the bid.

Our results on the VR behaviour of individual stocks is more consistent with the literature. All the 100 stocks show significant negative deviations from the mean at a lag of five minutes. This means that stock price changes at five minute intervals have temporal dependance. We infer that the market makes corrections to pricing upto that interval at which the serial correlation becomes insignificant (which in the case of our fastest mean–reverting stock are RELIANCE, ZEETELE, HIMACHLFUT). However, there is a strong heterogeneity in the behaviour of the serial correlation across the 100 stocks. While the shortest time to mean reversion is 2, the longest is 1955 for HDFCBANK (1922 for IDBI).

We attempt to analyse the cross–sectional differences in behaviour of the VRs as correlated with their cross–sectional differences in liquidity. Kyle (1985) characterises three aspects of liquidity: the *spread*, the *depth* of the limit order book (LOB) and the *resiliency* with which it reverts to its original level of liquidity. Roll (1984) shows how the liquidity measure of the bid–ask bounce can lead to negative serial correlations in price. Hasbrouck (1991) shows that the smaller the bid–ask spread of this stock, the smaller is the impact of the trade, which would lead to a small correlation. Whereas an illiquid stock with the same depth but larger spread would suffer a larger serial correlation at the same lag.

Similarly, we postulate that the resiliency of the LOB is higher in the case of a liquid stock (ie, the LOB reverts back to its original state faster) compared with the resiliency of a LOB of an illiquid stock. The bid–ask spread after a liquidity shock persists in the case of an illiquid stock. This would imply that the serial correlation in an illiquid stock would die out at a much slower rate than a liquid stock. This could mean that liquidity has an impact not only on the *sign* of the serial correlation but also its magnitude and persistence, and therefore, the rate at which the VR reverts back to the mean.

The NSE disseminates the full limit order book information available for all listed stocks, observed at four times during the day. We construct the spread and the *impact cost* of a trade size of Rs.10,000 at each of these times for each of the stocks. We use the average impact cost as a measure of intra–day liquidity of the stocks. We then construct deciles of the stocks by their liquidity, and examine the average VRs observed for each decile.

We find that the top decile by liquidity (ie, the most liquid stocks) have the smallest deviation from mean at the first lag as well as the shortest time to mean reversion in VRs The bottom decile by liqudity have the largest deviation from mean at the first lag as well as the longest time to mean reversion. The pattern of increasing time to mean reversion is consistently observed as correlated with decreasing liquidity as measured by the impact cost.

The paper is organised as follows: Section 2.2 lays out the issues in analysing intra–data. Subsection 2.2.1 deals with issues about HF prices and returns while subsection 5.1 deals with measures of intra–day liquidity. In Section 2.1, we look at the developments in the VR methodology since Nelson and Plosser (1982), and describe the specific forms that we choose to calculate the VR as well as the method of inference. We describe the dataset in Section 3, and go on to describe our results in Section 5.2. We conclude our findings in Section 6.

2 The Variance Ratio Approach

Beginning Cochrane (1988), a large number of studies have used the variance-ratio test to examine the random walk behaviour of prices of commodities and financial instruments. These have ranged from testing for serial correlation to martingales differences in financial returns. Some of the important ones are summarised below.

Poterba and Summers (1988) tested stock prices and market returns in the US and 17 other countries over many decades and found that stock prices are positively correlated over short horizons and and negatively correlated over longer horizons. They suggested two reasons for mean reversion in stock prices, *viz.*, time-varying required returns and "price–fads"—mistaken, long-memory notions about prices that make the stock price differ from it's fundamental values for long periods of time. Lo and MacKinlay (1988) tested their Variance Ratio (VR) test on weekly data over 1962–1985, and rejected the Random Walk (RW) hypothesis for many stock indexes. Brock and LeBaron (1989) used the VR test in a theoretical production-based asset-pricing model to examine the temporal effects of credit constraints on firms. They showed that credit constraints lead to mean-reversion in stock prices. Fama and French (1987) challenged Poterba and Summers (1988) conclusions on long-horizon returns. They aver that univariate tests on long-horizon returns are imprecise. They contend that Poterba and Summers (1988) are unable to reject the RW even though they use VRs estimated on returns from 1871 to 1985.

Liu and He (1991b) show that the random walk hypothesis is rejected for five pairs of weekly nominal FX rate series, using the homoscedastic- and the heteroscedasticity-consistent forms of the Lo and MacKinlay (1988) test. They point towards FX overshooting or undershooting being the case of autocorrelations in weekly nominal FX rates. Liu and He (1991a) further reject the Random Walk Hypothesis (RWH) using ten pairs of monthly real FX series. They reject the assumption of an uncorrelated element in the real FX rate, which had come about as a result of previous less powerful tests. They thus show that deviations from PPP have a tendency to be reversed in a short period of time. VR tests have been used to test for the randomness properties of other times series as well. Malliaris and Urrutia (1990) apply the test to the macroeconomic series used by Nelson and Plosser (1982) to reexamine the random walk component in their series. They augment Nelson and Plosser's findings in this regard. Cogley (1990) extended the Cochrane (1988)'s work on the US economy using the VR test by extending the analysis to nine Organization of Economic Cooperation and Development (OECD) countries. He found that the limited stability of long-term growth is limited to the US. Krol (1992) uses the VR test to examine the trend properties and the degree of persistence in industrial production, and similar aggregates in the United States, post World-War-II.

2.1 Computation of Variance-Ratios

Parallel to the application of the idea of variance-ratio tests, there have been many developments in the variance-ratio methodology as well. We chart a brief course this area. Wilson and Tonascia (1971) first gave the simulated tables for the shortest confidence intervals on the variance ratios from the normal distribution. Szroeter (1978) discussed generalised VR tests for multivariate regression models. Poterba and Summers (1988) while analysing the transitory component in stock returns showed that VR tests are among the most powerful tests in detecting mean-reversion but do not much power against alternatives to the RW.

Lo and MacKinlay (1988) first proposed the VR for the random walk, and then extended it to the more general uncorrelated but possibly heteroscedastic random walk. Their VR test essentially data sampled at various frequencies. Monte Carlo studies showed that the VR test is more powerful than either the Augmented Dickey-Fuller (ADF) test or the Box-Pierce "Q" test, in testing the Poterba and Summers (1988) hypothesis of the "price-fads" in stock returns. They tested the three tests above for their power on a stationary AR(1) process, the sum of an AR(1) process, and an I(1) series. Lo and MacKinlay (1989) also talked about the sample properties of the VR test in terms of it's size and power.

Chandrakanta et al. (1989) develop the exact finite sample distribution theory of the VR test, analytically characterising the distributions of the test statistics under the null and the alternatives, and studying the corresponding power function. Cogley (1990) provide an improved approximation to the distribution of Cochrane (1988)'s VR statistic.

Faust (1988) showed that the VR is a Uniformly Most Powerful Test (UMPT) is large samples in distinguishing between an White Noise (WN) and a certain class of mean-reverting alternate models. He also gives the exact small sample distribution of the test, and a test based on simple asymptotic approximations, and also a joint test of many VR statistics. Faust (1992) generalises the concept of the VR statistic to a class of statistics called Filter Variance Ratio (FVR)s, that have the variance of a filtered time series, and the variance of the original time series. This class of statistics is optimal for testing the WN null against a class of serially dependent alternatives, which are characterised. The VR test is then shown to be an optimal test for mean-reversion. Faust also shows that a Gaussian AR series in first differences belongs to the class of these serial dependent alternatives.

Lippi and Reichlin (1991) examine the effects of temporal aggregation in time series on trend-cycle variances and persistence in time-series. They show that temporal aggregation increases the trend-cycle variance-ratio. Mikkelsen (1992) showed that the VR test has much higher power then than the Dickey-Fuller (DF) testing the null hypothesis of unit roots against fractional alternatives., using Monte Carlo simulations. He also showed that the power of the VR test rises much much faster with sample size. Chow and Denning (1993) extend the Lo and MacKinlay (1988) VR test to testing multiple VR at a time. They contend that failing to correct for the joint test size may result in very large Type I errors in VR tests. Monte Carlo results state that the Multiple Variance Ratio (MVR) test is as good as the DF and the Phillips-Perron (PP) tests for a stationary AR(1) alternative, and more powerful for two unit-root alternatives, *viz.*, an Autoregressive Integrated Moving Average (ARIMA)(1,1,1) and an ARIMA(1,1,0).

Cecchetti and sang Lam (1994) study the accuracy of asymptotic approximations of the VR in small samples, and the size distortions arising from searching over many horizons in model selection. They propose a joint test combining VR statistics at many horizons and also provide a Monte Carlo procedure for exact inference. Fong and Ouliaris (1995) compare the VR test with the spectral shape test of Durlauf (1991) to test the martingale hypothesis. The spectral shape tests are consistent against all stationary nonWN alternatives from martingale null. Proietti (1996) shows that Cochrane (1988)'s normalised VR can be extended as a non-parametric counterpart to a parametric measure of persistence based on the cumulated impulse response function. This measure would also be useful in measuring seasonally integrated processes. Ronen (1997) examines the VR test across trading and non-trading periods, using the Generalized Method of Moments (GMM) and finds that the conventional VR test has bias against the null in small samples.

Cochrane (1988) showed that the VR statistic is equal to the spectral density of the series at frequency 0. Using this, Choi (1999) calculates the VR statistic using Andrews (1991) optimal data-dependent methods. He then uses Priestly (1982) normality approximation. He also reports simulation results for these tests. Smith and Naik (1997) derive analytic finite-sample approximations of the bias and the standard error of a class of statistics used to test for serial correlations, where the underlying distribution can be Gaussian or any continuous distribution.

Wright (2000) proposes VR tests based on the ranks and signs of a time series. In contrast to standard VR procedures, these tests are exact., and found to have more power, by Monte Carlo simulations. They are also found to provide unambiguous results when tested on five FX rates, as opposed to the conventional VR.

The Lo and MacKinlay (1988) VR test is based on the assumption that the underlying distribution is normal. Lo and MacKinlay consider the $VR(q)$ as a Hausman-type test, and derive the asymptotic distribution. (See Kim et al. (1991) for more details.) It is widely reported and known that financial returns are conditionally heteroscedastic. Lo and MacKinlay (1988) have derived a heteroscedasticity-consistent $VR(q)$ to account for this effect in stock returns.

There are other problems with the $VR(q)$. It can be shown that as $q \to \infty$, the VR statistic is necessarily less than one. This means that it is biased towards mean-reversion for long aggregation periods. See Campbell et al. (1997) for more details. Since the VR test is mainly used with daily, weekly, monthly or even quarterly and yearly data, this problem is not very severe. However, with high-frequency data, the periods of aggregation are typically large. With 5-minute data, aggregation up to three months would mean calculating the $VR(q)$ to $q = 5025$ lags. At such lags, the $VR(q)$ is biased towards mean-reversion.

Next, the Lo and MacKinlay (1988) $VR(q)$ statistic assumes that the aggregation q, as an function of the sample size N, is small. In other words, $q/N \to 0$. It can be shown that as q rises with respect to N, all inference procedures break down, *e.g.,* if $q = N/2$, then

$$\min \frac{(VR(q) - 1}{\sqrt{v}} = -1.74.$$

This implies that the $VR(q)$ statistic would *never* go below -2. More severely, if $q = 2N/3$, then $Z_1 > -1.5$, always. See Campbell et al. (1997) for more details.

A number of solutions have been proposed for this problem. Poterba and Summers (1988) use Monte Carlo simulations to estimate the asymptotic distribution of the of the $VR(q)$, and employed Cochrane (1988)'s bias correction, where they divide the $VR(q)$ by it's expected value

$$E[VR(q)] = \frac{2-q}{q} + \frac{2}{q}\sum_{j=1}^{q-1}\frac{N-q}{N-j}.$$

This bias correction works well for small samples. Kim et al. (1991) use a randomisation technique (See Noreen (1989)), similar to the bootstrap, where the dataset is removed to remove temporal dependence, and then calculate the VR. The significance level is calculated by finding out the number of times these VRs are lower than that of the actual dataset.

A bootstrapping scheme is used by Pan et al. (1997) to derive the distribution of the the $VR(q)$ statistic. They draw do a Simple Random Sampling with Replacement (SRSWR) on their sample of daily returns of four currency futures, a thousand times. This way they get the complete empirical distribution of the $VR(q)$.

The Lo and MacKinlay (1988) $VR(q)$ assumes q to be fixed, and $q/N \to 0$. Priestly (1982) derived that the asymptotic distribution of the $VR(q)$ when

$$q \to \infty,\ N \to \infty, \quad and \quad q/N \to 0$$

is

$$V\hat{R}(q) \sim N\left(VR(q), \frac{4q}{3N}VR(q)\right)$$

Cecchetti and sang Lam (1994) have used Monte Carlo simulations to derive a better approximation of the empirical distribution of the $VR(q)$ statistic under the same assumptions.

Richardson and Stock (1989) then derived the asymptotic distribution of the $VR(q)$, with

$$q \to \infty,\ N \to \infty, \quad and \quad q/N \to \delta, \quad a\ constant.$$

They have shown that under this assumption, the Non-overlapping Variance Ratio (NVR) statistic $VR_N(q)$ has a limiting Chi-Square distn given by

$$VR_N(q) \xrightarrow{d} \delta\chi^2_{1/\delta}$$

This non-normal statistic is not consistent. It is robust to heteroscedasticity and non-normality. They next show that the Overlapping Variance Ratio (OVR) statistic $VR(q)$ has the following limiting distn

$$VR(q) \to \frac{1}{\delta}\int_{\delta}^{1} Y_{\delta}(\lambda)^2 d\lambda,\ where$$

$$Y_\delta(\lambda) = W(\lambda) - W(\lambda-\delta) - \delta W(1)$$

Here $W(\lambda)$ is the standard Brownian Motion (BM) on $[0,1]$. This gives

$$E\left[\frac{1}{\delta}\int_\delta^1 Y_\delta(\lambda)^2 d\lambda\right] = \frac{1}{\delta}\int_\delta^1 E[Y_\delta(\lambda)^2]d\lambda = (1-\delta)^2$$

This shows that the limiting distn depends on the functional of a BM, rather than being normal. This distn is fully non-parametric, and Richardson and Stock use Monte Carlo to estimate it.

Tse and Zhang (2002) derive the asymptotic distn of the OVR $VR(q)$ statistic with the assumption that the underlying distn belongs to the stable Paretian family of distributions. They find that the OVR statistic converges very slowly to it's asymptotic limit.

In this paper,we follow Lo and MacKinlay (1988)'s notation throughout. Say the sample of prices has $nq+1$ observation. Define

$$\hat{\mu} = \frac{1}{nq}\sum_{k=1}^{nq}(p_k - p_{k-1}) \tag{1}$$

$$\sigma_a^2 = \frac{1}{nq-1}\sum_{k=1}^{nq}(p_k - p_{k-1} - \hat{\mu})^2 \tag{2}$$

$$\sigma_b^2(q) = \frac{1}{m}\sum_{k=q}^{nq}(p_{qk} - p_{qk-q} - q\hat{\mu})^2 \tag{3}$$

$$\sigma_c^2(q) = \frac{1}{m}\sum_{k=q}^{nq}(p_k - p_{k-q} - q\hat{\mu})^2 \tag{4}$$

$$m = q(nq-q+1)(1-\frac{q}{nq}) \tag{5}$$

Now define the statistics

$$\tilde{VR}(q) = \frac{\sigma_b^2(q)}{\sigma_a^2} \tag{6}$$

$$\overline{VR}(q) = \frac{\sigma_c^2(q)}{\sigma_a^2} \tag{7}$$

Here, $\tilde{VR}(q)$ is the non-overlapping and $\overline{VR}(q)$ is the overlapping variance-ratio statistic. Under the normality assumption for returns, and homoscedastic increments, the asymptotic distribution of the two statistics are

$$\sqrt{(nq)}(\tilde{VR}(q)-1) \xrightarrow{d} N(0, 2(q-1)) \tag{8}$$

$$\sqrt{(nq)}(\overline{VR}(q)-1) \xrightarrow{d} N\left(0, \frac{2(2q-1)(q-1)}{3q}\right) \tag{9}$$

$$\tag{10}$$

However, financial time series typically have heteroscedasticity in them. Lo and MacKinlay (1988) also gave the asymptotic distribution of the $\overline{VR}$ statistic under heteroscedastic increments. This is a more general test where the null is an uncorrelated but heteroscedastic random walk. Now, the $\overline{VR}$ statistic is not unbiased.[1] Define

$$\hat{\delta}_k = \frac{nq\sum_{j=k+1}^{nq}(p_j - p_{j-1} - \hat{\mu})^2(p_{j-k} - p_{k-1} - \hat{\mu})^2}{\left[\sum_{j=1}^{nq}(p_j - p_{j-1} - \hat{\mu})^2\right]^2} \tag{11}$$

[1] But Lo and MacKinlay (1988) have shown that the statistic is consistent, nevertheless.

$$\hat{\theta}(q) = 4\sum_{k=1}^{q-1}\left(1-\frac{k}{q}\right)^2\hat{\delta}_k \tag{12}$$

$$\sqrt{nq}(\overline{VR}(q)-1) \sim N(0,\theta). \tag{13}$$

Typically, overlapping variance-ratio statistics are calculated, since more observations are then available. However, Richardson and Smith (1991) found that overlapping VR statistic have about 22% lower standard errors than non-overlapping VR stats. Recently, Andersen et al. (2001) have used the non-overlapping VR statistic in their analysis.

2.2 Issues with VR using High-Frequency data

High-frequency finance (See Goodhart and O'Hara (1997), Dacorogna et al. (2001)) is a relatively new field, and there are new developments coming up all the time. A number of studies have been conducted on the imformational efficiency of intraday markets (See Goodhart and Figliuoli (1991), Goodhart and O'Hara (1997), Baillie and Bollerslev (1990), Gavridis (1998), Dunis (1996) and ap Gwilym et al. (1999) for more details.). Both the equity and the FX markets have been extensively studied.

The first HF data that became available were time series of prices of every single trade on the New York Stock Exchange. (Wood et al. (1985); McNish (1993),Harris (1986); Lockwood and Linn (1990)). These papers were mainly involved in the empirical characterization of the markets. The Swedish firm Olsen and Associates have been insturmental in making the high=frequency data of the largest markets in the world, the foreign exchange markets to researchers. Goodhart and Figliuoli (1991), Guillaume et al. (1994) are some of the papers in this area. An excellent survey of the literature can be found in Goodhart and O'Hara (1997). Others are Gavridis (1998), Dunis (1996) ap Gwilym et al. (1999).

There are a number of properties peculiar to intraday data. These range from the high amount of noise in the data, to strong intraday periodocity (See Andersen and Bollerslev (1997)), significant first-order negative autocorrelation Goodhart and Figliuoli (1991), U-shaped patterns in the volatility, number of trades, and the volume.

2.2.1 Choice of frequency for the price data

An important issue that has to be dealt with when working with HF data is inherently irregular nature of data observed in real–time (Granger and Ding, 1994). Most trading systems at exchanges can execute trades at finer intervals than the frequency at which exchanges record and publish data. For example, the trading system at the NSE can do trades at intervals of $1/64^{th}$ of a second. However, the data is published with timestamps at the smallest interval of a second. There could be multiple trades within the same timestamp. Some stocks do trade with the intensity of several trades within a second, whereas other may do just one. Thus, the notion of the "last traded price" (LTP) for two different stocks might actually mean prices at two different real times, which leads to comparing asynchronous or *irregular* data.[2].

There are several methods of dealing with irregular data, in the literature. Dacorogna et al. (2001) follow a time-deformation technique called ν-time, where they model the time series as being a subordinated process of another time series. Another approach is to model the time series directly in trade-time, as opposed to calendar time. This approach is followed to check for non-linearities and heavy-tailedness in the data. Marinelli et al. (2001) We follow the approach of Andersen and Bollerslev (1997), and impose a discrete-time grid on the data. The width of the grid is non–trivial. If the width of the interval is too

[2] For example, stock A at the NSE trades four times within the second, with trades at the $1^{st}, 2^{nd}, 3^{rd}$ and 15^{th} 64^{th} interval of a second. Stock B trades once with a trade at the 63^{rd} interval. Then the LTP recorded for A is at $15/64^{th}$ of a second and B at $63/64^{th}$ of a second which is asynchronous in real time

high, then information about the temporal pattern in the returns may be lost. On the other hand, a thin interval may lead to high incidence of non—trading which is associated with spurious serial-correlation in the returns. Therefore, the grid interval must be chosen to minimize the informational loss while avoiding the problem of spurious autocorrelation. When intraday data is concatenated across days, the first return on any day is not an intraday return, but an overnight return. Returns over differing periods can lead to temporal aggregation problems in the data, and result in spurious autocorrelation. Here we follow Andersen and Bollerslev (1997), and drop the first observation of the day. For more details on the issues of data-filtering and cleaning, see Patnaik and Shah (2002).

2.2.2 Patterns in HF data

Another issue that arises when analysing the serial correlations in the data are the the U–shaped patterns, documented for markets over the world. It is observed that returns in the beginning of the trading day and the end of the trading day tend to be different as compared with data in the middle, etc. There is a definite U-shape in the intraday volatility of returns. Some of the many papers that have documented this fact, in various markets across the world are Wood et al. (1985); Stoll and Whaley (1990); Lockwood and Linn (1990); McNish and Wood (1990b,a, 1991, 1992); Andersen and Bollerslev (1997), and many others. The U-shape is also seen in the trading volume, and in the bid-ask spread. This means that there are strong intraday seasonalities in intraday data, which have to be filtered out.

The use of the variance-ratio technique to examine market efficiency has been relatively low for intraday markets. Low and Muthuswamy (1996) tested three foreign exchange quotes, *viz.,* USD/JPY, DEM/JPY, and USD/DEM for the period October 1992 — September 1993. They calculate variance ratios for these FX returns where the holding period (aggregation) ranges from 5 minutes till 3525 minutes (corresponding to 750 5-minute intervals). They find that the variance ratios do not scale up and that this negative correlation in the returns becomes stronger as the as the holding period increases. VRs are found to grow faster in the short-run, *i.e.,* less than 200 minutes. This shows that "serial dependencies are stronger in the long-run". (See Low and Muthuswamy (1996, Page 19)).

Andersen et al., in 2001 have averred that the standard variance-ratio tests could be seriously misleading when applied on intraday data, due to the inherent intraday periodicity present. They advocate the use of the Fourier Flexible Form (FFF) (See Gallant (1981)) regression technique when analyzing such data. However, their analysis is primarily devoted towards studying the shift in volatility patterns in HF data, and not serial correlation in the returns. Further, they use non-overlapping returns for their analysis.

3 Data description

The data for this study comprises of all trades in the Capital Market segment of the National Stock Exchange in the period Mar 1999–Feb 2001. The data set has about 253,717,939 records in 514 days. These are "time and sales" data, as defined by MacGregor (1999). A total of 1384 stocks traded in this period. The National Stock Exchange (NSE) follows a five-day trading week, with rolling settlement. Trading starts at 0955 in the morning, and continues without break till 1530. NSE is one of the most heavily traded stock exchanges in the world. There are more than 490,000 trades daily on an average. The trading system is a continuous open electronic limit-order book market, with order-matching done on a price-time priority basis.

3.1 Choice of frequency for the data

Intraday data is irregular, in the sense that there could be multiple trades with the same timestamp. This is due to the fact that the highest precision of recording the data is a second. There are many methods of dealing with irregular dat, in the literature. We follow the approach of Andersen and Bollerslev (1997), and impose a discrete-time grid on the data. We have chosen a 300 second interval as the base frequency for this grid. The choice of this interval is non-trivial. If the width of the interval is too high, then vital information may be lost. On the other hand, a thin interval may lead to non-trading and associated spurious serial-correlation in the returns. These would in turn affect the variance-ratio calculations. Our choice of a 300 second interval is based on various tests and diagnostic measures. For more details, see Patnaik and Shah (2002). The 300s discretization gives about 67 data points per day. There are about 10 per cent missing observations. This gives about 31,000 observations per stock. In addition, the data were filtered of outliers, and anomalous observations.

4 Patterns in Variance Ratios: Examples from data

A random walk has variance linear in time. This implies that the variance of the random walk should scale up linearly as the holding period increases. In other words, the VR, should be horizontal at unity, once the holding period is factored in. Statistically, it should be insignificant from 1. If a market is efficient in the weak form, then we expect that the VR of a stock price would be insignificant from unity. Any disturbances are expected to be temporary, and insignificant in the long-run. The assets that are studied here comprise the NSE index, called the S & P CNX Nifty (NIFTY), and individual stocks.

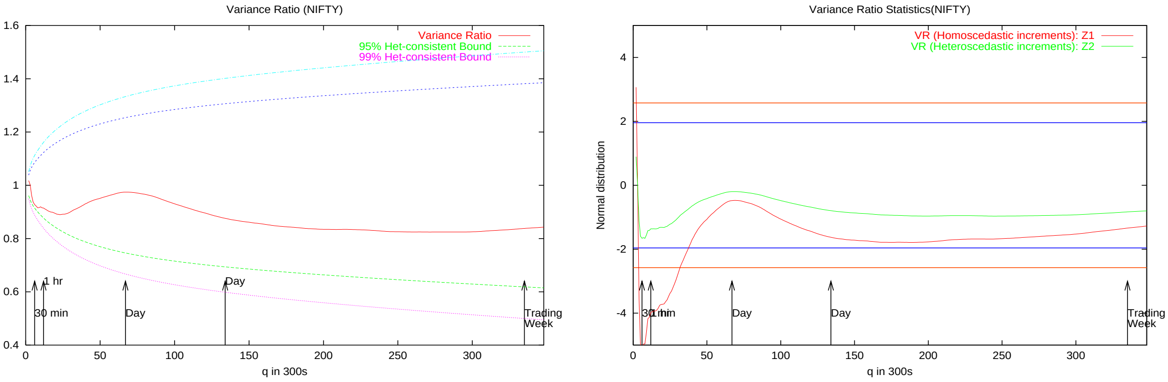
In our sample we took the hundred most liquid stocks in the NSE,(See section 5.1 for more details on the selection process.). We then calculated VRs for each of these stocks in the sample and the index S & P CNX Nifty (NIFTY) using returns discretized at 300s intervals. The holding period was from 10 minutes (aggregation q =2), to three months (q = 5025). We use the Lo and MacKinlay (1988) technique to calculate the VRs, and the heteroscedasticity-consistent standard errors, for them. Since the ratio q/N (See section 2 for more details.) in our case, is about 0.16, the large-sample inconsistency problems of the VR statistic do not arise.

There are inherent problems associated with just studying the market index. A market index consists of various stocks, which differ in their liquidity. The reaction time of stocks across the index to an information shock, is therefore different. "An index may give a completely false impression of the extent of price fluctuation in individual markets. If, in any time period, only one share in an index of fifty shows variation, the averaging would ensure that this variability on a reduced scale would be imputed to the whole market." (Kemp and Reid (1971), as quoted in Cooper (1982)). This variability should show up as spurious positive autocorrelation in index returns.

The VR and the standard statistic for the index NIFTY are shown in figure 1. It is clear that the index shows almost no serial correlation at any aggegation period. There's some small but insignificant negative serial correlation in the short-run (upto a half-day), and there is a rise in the VR after that, till the end of the day. There's a general trend towards low serial dependence and mean-reversion as the holding period increases, but it's not significant. Thus our results do not support the hypothesis of spurious correlations indiced in the index, due to varying reaction-times of stocks to external information.

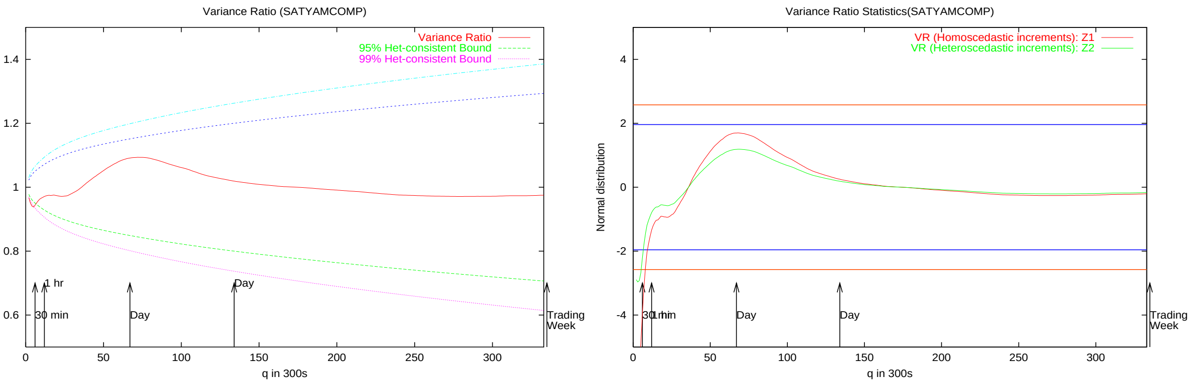
Figure 2 shows the VR for a highly liquid stock, SATYAMCOMP, in the same period. The pattern in the VR is almost similar to the index, but more pronounced, owing to the noise in the single stock. The tendency to mean-revert is lower for this stock than the index. The heteroscedasticity-consistent VR statistic $Z2(q)$ is insignificant from zero after about an hour. This implies that highly liquid stocks have almost no serial correlation in the Indian intraday markets (albeit there's some weak negative autocorrelation). There is a tendency for the stock to mean-revert, (after a fortnight) in the long-run, but it's not statistically significant.

Figure 3 shows the VR for a relatively less liquid stock, TATAPOWER in the same period. Contrary to the index and the liquid stock, here there is strong mean-reversion in the returns, in the short-term (up to two-three weeks). The heteroscedasticity-consistent $Z2(q)$ statistic is significantly negative all



(a) (b)

Figure 1: (a) Variance Ratios for NIFTY over one week, for the period Mar 1999 — Feb 2001. NIFTY index values are discretised at 300s intervals. (b) Heteroscedasticity-consistent standardized VR statistics for NIFTY for the same period.



(a) (b)

Figure 2: (a) Variance Ratios for SATYAMCOMP over one week, for the period Mar 1999 — Feb 2001. SATYAMCOMP returns are discretised at 300s intervals. (b) Heteroscedasticity-consistent standardized VR statistics for SATYAMCOMP for the same period.

the way up to a fortnight. Thereafter the serial correlation becomes insignificant, but remains weakly negative. Similar inspection of any relatively illiquid stock shows essentially the same result of significant mean-reversion in the short-run, and weak negative autocorrelation in the long-run.

5 The role of liquidity

Market liquidity for a commodity is one of it's most desirable aspects. A market is said to be liquid if "a large volume of trades can be immediately executed, with a minimum effect on prices". According to Kyle (1985), a market is liquid if it has *depth*, *tightness*, and *resilient*. For a market with depth, a large volume has to be transacted, in order to bring about even a unit change in the price. Tightness means that an agent can switch sides with minimum associated cost, implying a small spread. Prices in a resilient market would take a very small time to come back to their fundamental levels after a shock. Liquidity has an impact on the market efficiency of prices. Lack of liquidity would lead to higher trading costs, and thus lower market efficiency.

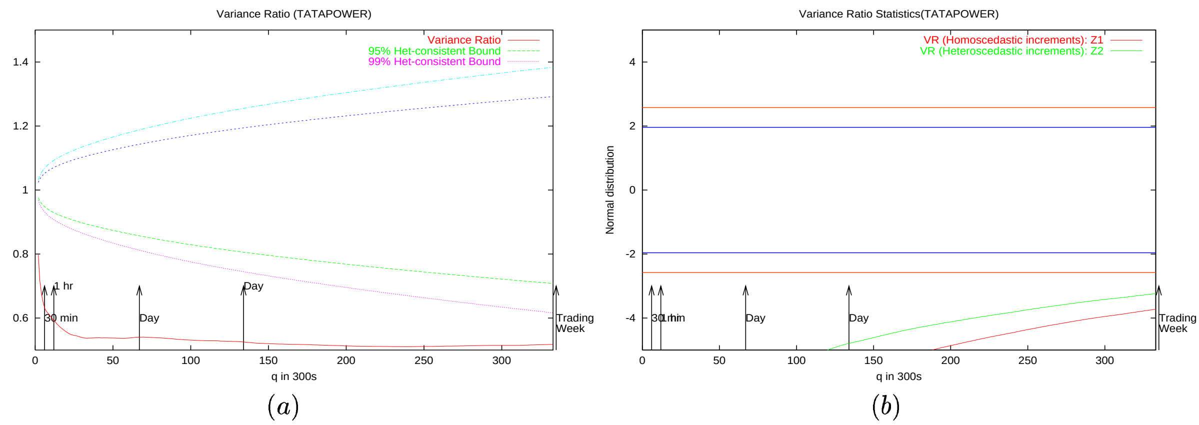


Figure 3: (a) Variance Ratios for TATAPOWER over one week, for the period Mar 1999 — Feb 2001. TATAPOWER returns are discretised at 300s intervals. (b) Heteroscedasticity-consistent standardized VR statistics for TATAPOWER for the same period.

Stocks with higher liquidity are expected to react faster to news and information as compared to those that with lower liquidity. This would mean that the standard tests of weak–form market efficiency would have different manifestations depending upon the liquidity of a stock: a stock with higher liquidity is less likely to reject standard tests compared to those with lower liquidity. One implication is that serial correlation in the patterns of returns of stocks with weaker liquidity would show stronger rejections of the random walk hypothesis compared with those having better liquidity. Such a stock will take a long time to revert to it's fundamental value (longer mean-reversion). However, there isn't much theoretical evidence to corroborate this issue. See, for example, Madhavan (1992).

5.1 Liquidity measurement

"liquidity, like . . . , is easily recognized but not so easily defined" (See O'Hara (1995)), and there are a number of proxies Some of them are the number of trades, called the "trading intensity", the trading volume, the bid-ask spread, the effective bid-ask spread and the impact cost. Each of the proxies has some advantages and disadvantages. For instance, the bid-ask spread is based on quotes which are just indicative. The trading intensity does not take in the price of the asset. There are pre- and post-trading measures of liquidity too. The liquidity ratio of a stock is the ratio between the annual trading volume over daily market capitalization.

5.1.1 Trading intensity

The trading intensity of a stock is the number of times it trades in a given time interval. NSE has one of the highest trading intensities in the world. In this period studied here, the mean trading intensity was about 24.56 trades per second.[3] For the study, we chose the hundred most traded stocks in the period. The table of the these stocks is given in the Appendix. On an average, each of these stocks traded about 4211 times a day, about one trade in every 5 seconds. These 100 stocks comprise about 83.32% of all the trades recorded in this period at the NSE.

Trading intensity is a fairly common measure of liquidity, but suffers from the fact that it does not have the price of the stock incorporated in it. Hence a high-priced but relatively lowly-traded stock would be quite illiquid in the trading intensity measure.

[3]253717939 trades over 514 days with 20100 seconds/ day, gives 24.56

Decile	Trades	Mean IC	Minimum	Maximum
1	12323.29	0.098	0.07	0.12
2	11878.61	0.131	0.12	0.14
3	6615.22	0.144	0.14	0.15
4	4587.33	0.158	0.15	0.17
5	1783.87	0.174	0.17	0.18
6	2115.46	0.191	0.18	0.21
7	1146.91	0.219	0.21	0.23
8	594.96	0.231	0.23	0.24
9	904.73	0.245	0.24	0.25
10	666.27	0.252	0.25	0.26

Table 1: Summary statistics of the portfolios. Portfolios are selected on the basis of impact costs for an order of Rs. 10,000, during the period Mar 1999-Feb 2001.

5.1.2 Impact cost using order book snapshots

The impact cost of a stock is the premium that is paid over and above the price of a stock, while trading the stock. The buy-price of a quantum of stock is always higher than the price of that quantum in the market. This higher price is due to the increased demand on the stock, from an equilibrium. Simliarly, when a quantum of stock is sold in the market, increased supply of that stock reduces the price of the stock. This results in the seller realizing a lower price than the market price of the stock. This reduction/increase is proportional to the supply/demand of the stock in the market. Impact cost is thus a measure of the liquidity of a stock. It is inversely proportional to the liquidity of the stock.

The impact cost for a stock is calculated using order-book snapshots of the market which are recorded four times a day by the NSE, at 1100, 1200, 1300, and 1400 respectively. The complete Market By Price (MBP) is available at these times.[4]

For our study, we calculated the buy/sell impact cost for an order of Rs. 10000 for each of the 1382 stocks that traded in the sampling period. The median buy and the sell prices were selected. Stocks were then selected on the basis of mean buy and sell impact cost. The list of stocks is in the appendix. We took the hundred most liquid stocks on the market, using this technique. A glance at this table shows that the most liquid stock is RELIANCE, which has an inpact cost of Rs. 7 for a transaction size of Rs. 10000, and the least liquid stock is CORPBANK which has Rs. 26 as it's mean impact cost.

The table for all the stocks selected and their respective trading intensities and impact costs, is given in the appendix.

5.2 Cross–sectional variation of serial correlations

According to our hypothesis, a highly liquid stock would necessarily show almost no serial correlation. This implies that the VR of such a stock would become indistinguishable from that of a random walk in a very short time interval. However, as liquidity goes down across stocks, we expect this interval to lengthen. The least liquid stocks would take a siginificantly long time to return to the random walk.

We have seen the effects of liquidity on individual securities, in the previous section. For a highly liquid stock, the VR is weakly positive in the short-run, up to a day and then weakly negative. A nice example is the stock SATYAMCOMP. There is little mean reversion in the liquid stocks. As liquidity goes down, the tendency of a stock to revert to a mean, rises. For a low-liquidity stock, the VR is significantly negative. (See TATAPOWER). However, since individual stocks have a lot of "idiosyncratic noise", (See Campbell et al. (1997, Page 72)) we analyze the VRs for a portfolio of stocks, and of a

[4]I thank my colleague Santosh Kumar for helpful discussions on this issue.

Portfolio/ Decile	Time in 300s
1	36
2	54
3	118
4	64
5	290
6	253
7	108
8	309
9	370
10	451

Table 2: Time taken to for a Cross-sectional average VR to be insignificant from unity.

cross-sectional average of the VRs of individual stocks, in addition to liquidity-based portfolios.

5.2.1 Liquidity-ranked Portfolios and Cross-sectional Averages

Portfolio VRs are calculated in two ways. In the first, we get portfolio returns from the portfolios selected, and then calculate the VR's from them and their heteroscedasticity-consistent standard errors using Lo and MacKinlay (1988). Our findings are presented in tables 3, 5.

While differences in the stocks due to their varied liquidity would show up across the ten portfolios, the portfolios would even out the idiosyncratic characteristics of individual stocks. Further, microstructural problems associated with a single stock, like bid-ask bounce, non-synchronous trading, etc. would be mitigated for a portfolio. The constituents of each of the portfolios are in the appendix.

We also calculate the portfolio-wise cross-sectional averages of the variance ratios of the each of the stocks in the ten portfolios. See 6. Since individual stocks are found to be negatively correlated, indicating mean-reversion, we expect that cross-sectional averages to be negatively correlated as well. We find that this is indeed so. The cross-sectional average VRs are uniformly lower than unity.

The VR of the most liquid portfolio is similar to that of the liquid stocks and the index. We expect portfolio VRs to be more than unity, and thus returns to be positively correlated. This is expected due to the strong cross-correlation that exists between stocks. This cross-correlation makes the portfolio returns positively correlated, even though the components are negatively correlated. In the portfolios studied, we find that this is indeed the case. The VRs are postively correlated, and mean-revert, as the holding period increases.

However, this behaviour is seen only after durations of one day or more. Intraday, portfolio returns are found to be negatively-correlated. Portfolio returns are thus negatively correlated in the extremely short-term (less than one day), and then positively correlated. Our findings thus augment those of Poterba and Summers (1988), who find stock returns to be postively correlated in the short-term, and mean-reverting in the long-run. While we find similar behaviour in horizons longer than a day, we also find negative serial correlation intraday. Morover, these results are strongly dependent on the liquidity of the portfolio. While the first seven portfolios exhibit the behaviour dsecribed above, the eighth, ninth, and the tenth portfolios do not show the negative serial correlation intraday. These returns are positively correlated are just positively correlated and then negatively correlated in the long-run.

The most liquid portfolios thus show intraday mean-reversion. The duration to mean-revert differs across portfolios, and varies by liquidity, showing a clear linkage between liquidity and mean-reversion across stocks. The most liquid stocks have very low costs of trading attached to them, are traded in large quantities as a rule. This makes any changes to the prices of such stocks temporary. When the holding period is small, intraday, like 30 minutes to three hours, price changes are small, and thus there's a high

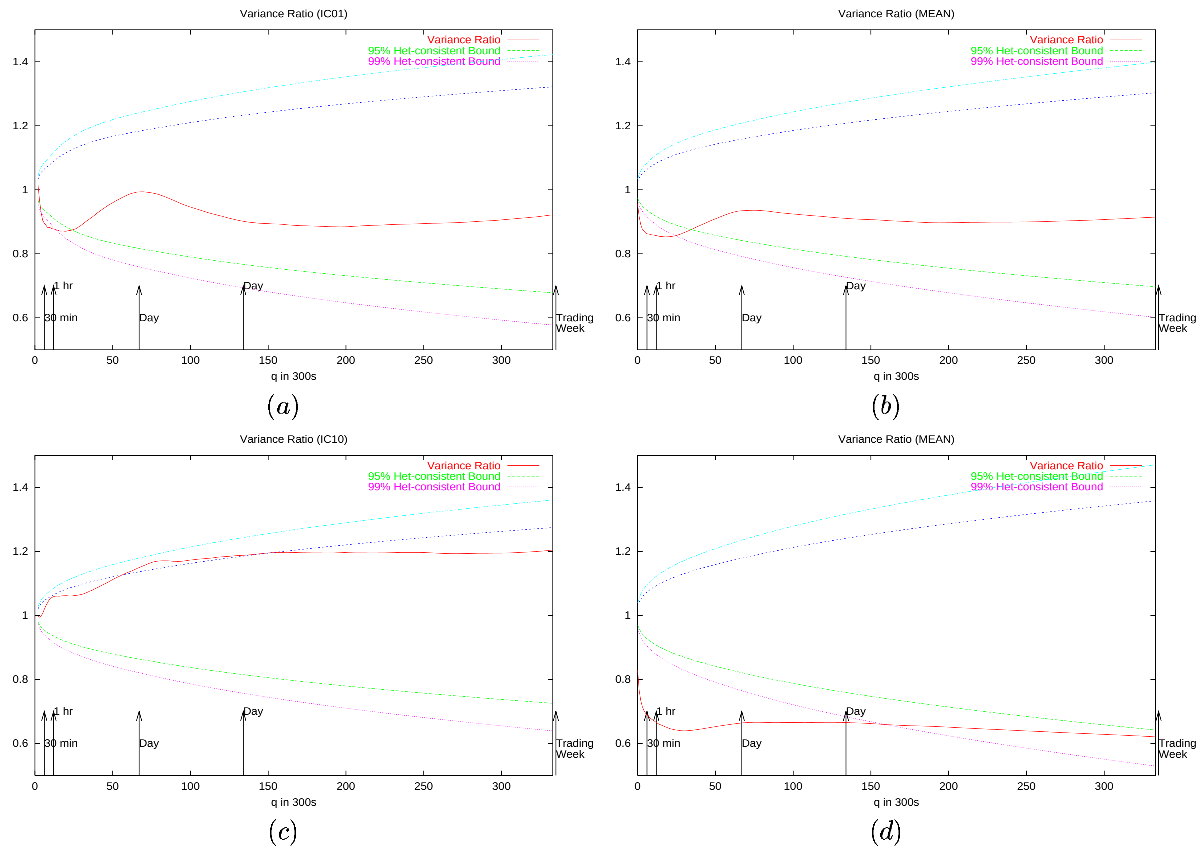


Figure 4: (a) is the portfolio VR of the top ten most liquid stocks from the NSE during the period Mar1999–Feb2001. (b) is the cross-sectional average VR of the same stocks taken separately. (c) is the portfolio VR of the bottom decile of the top hundred most liquid stocks in the NSE during the period Mar1999-Feb2001. (d) is the cross-sectioanl average VR of the same stocks taken separately.

tendency for the price of a liquid stock to mean-revert, in the extremely short-term.

On the other hand, if a stock is not very liquid, then trading costs for that stock are high. Such stocks do not trade very often. There is scant activity at the intraday level. Therefore there is mean-reversion, but only at the daily and multi-day level. This behaviour is common across all relatively illiquid stocks, resulting in strong cross-correlation among the stocks in the portfolio. Thus the portfolio returns are all positively correlated, and the VRs of the illiquid portfolios are all greater than unity. However, the VRs are almost all insignificant. We do not find any reason to reject the hypothesis of weak-form of market efficiency, in the Indian equity markets.

To control for this cross-sectional correlation among stocks, we consider the cross-sectional averages of the VRs, as opposed to the returns from the portfolios. These averages variance-ratios, give a better estimate of the serial correlation prevalent in stocks, than the portfolio VRs. Sure enough, there is hardly any positive serial correlation apparent from these VRs. As noted above, the lack of liquidity in a stock, and consequently in a portfolio, would lead to longer durations to a random-walk behaviour. This is clear from the cross-sectional VRs. The time taken to converge to random-walk becomes more and more across portfolios arranged by liquidity. The linkage between liquidity and the weak form of market efficiency is thus apparent even after controlling for cross-correlation across stocks.[5]

[5] The standard errors of the average VRs will not be independent (See Lo and MacKinlay (1988)). Hence the bands should be taken as indicative only.

Duration	q	$VR(q)$'s of Portfolios (Deciles) returns									
		1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}	6^{th}	7^{th}	8^{th}	9^{th}	10^{th}
10 minutes	2	1.014	1.021	1.027	1.041	1.071	1.024	1.026	0.937	0.980	0.998
15 minutes	3	0.965	0.991	1.008	1.032	1.076	1.024	1.022	0.914	0.951	0.995
20 minutes	4	0.927	0.971	0.985	1.019	1.068	1.021	1.020	0.904	0.939	0.999
25 minutes	5	0.905	0.965	0.970	1.012	1.065	1.025	1.021	0.899	0.931	1.007
30 minutes	6	0.895	0.967	0.963	1.014	1.065	1.031	1.028	0.900	0.930	1.019
35 minutes	7	0.890	0.973	0.963	1.018	1.066	1.036	1.037	0.901	0.932	1.030
40 minutes	8	0.883	0.973	0.959	1.021	1.064	1.039	1.039	0.904	0.931	1.039
45 minutes	9	0.882	0.979	0.958	1.024	1.065	1.042	1.046	0.906	0.932	1.047
50 minutes	10	0.881	0.981	0.956	1.027	1.065	1.045	1.053	0.908	0.934	1.054
55 minutes	11	0.878	0.982	0.951	1.025	1.062	1.045	1.059	0.908	0.934	1.056
1 hour	12	0.878	0.982	0.948	1.024	1.060	1.045	1.063	0.908	0.933	1.059
2 hours	24	0.874	0.985	0.923	1.017	1.035	1.055	1.067	0.897	0.932	1.061
3 hours	36	0.912	1.016	0.940	1.033	1.057	1.094	1.080	0.903	0.969	1.077
1 day	67	0.993	1.079	0.993	1.079	1.113	1.148	1.139	0.955	1.050	1.150
2 days	134	0.902	0.980	0.939	0.947	1.106	1.117	1.120	0.966	1.058	1.188
1 week	335	0.923	0.892	0.972	0.866	1.200	1.123	1.193	1.108	1.138	1.205
2 weeks	670	1.012	0.857	0.976	0.755	1.261	1.052	1.240	1.181	1.144	1.235
1 month	1675	1.050	0.742	1.022	0.662	1.501	0.754	1.285	1.521	1.389	1.466
2 months	3350	1.005	0.523	0.943	0.524	1.797	0.641	1.116	1.856	1.640	1.639

Table 3: Variance Ratios of top ten portfolios of NSE stocks arranged by liquidity. The stocks are selected from all the 1382 stocks that traded in the NSE in the period Mar1999-Feb2001. The stocks are arranged by liquidity and the top hundred are selected. The portfolios are then formed from the ten deciles. Impact cost for a buy/sell order of Rs. 10,000 is the liquidity measure.

6 Conclusion

The Efficiency Market Hypothesis (EMH) is one of the most hotly-debated topics in economics. This paper takes a fresh look at market efficiency by examining a possible linkage between market liquidity and weak-form efficiency, in the Indian stock markets.

High-frequency data is now available for many markets trading different assets. It allows fresh insights into the behaviour of many key market variables. The use of intraday high-frequency data for the period March 1999- Feb 2001, enables us to look at the market at the microstructural level. Due care has been taken to discretize the irregular intraday to a discrete-time grid.

Stock selection was done on the basis of a proxy for liquidity, called the impact cost of a stock. This proxy is better than conventional proxies for liquidity, like the bid-ask spread, the effective spread, the trading intensity, the trading volume, etc. Impact costs were calculated for an order of Rs. 10,000 for all the stocks traded in the exchange in the sample period. The top hundred were chosen, and portfolios were constructed out of each of the deciles.

Variance-ratios were calculated for portfolios, and for individual stocks. It was found that liquidity has a serious impact on the serial-correlation exhibited by a stock, and consequently a portfolio. The link between liquidity and market efficiency remains even after the heteroscedasticity and stock-specific effects have been filtered.

Duration	q	Homoscedastic $Z1$ statistics of Portfolios (Deciles) returns.									
		1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}	6^{th}	7^{th}	8^{th}	9^{th}	10^{th}
10 minutes	2	2.373	3.758	4.714	7.246	12.510	4.284	4.575	-11.050	-3.564	-0.296
15 minutes	3	-4.059	-1.116	0.927	3.762	8.970	2.866	2.623	-10.148	-5.790	-0.646
20 minutes	4	-6.804	-2.734	-1.449	1.810	6.419	1.935	1.866	-9.012	-5.777	-0.108
25 minutes	5	-7.625	-2.806	-2.429	0.929	5.234	2.015	1.721	-8.166	-5.541	0.543
30 minutes	6	-7.419	-2.335	-2.594	0.966	4.622	2.225	2.024	-7.128	-4.979	1.365
35 minutes	7	-7.035	-1.735	-2.360	1.135	4.228	2.340	2.379	-6.381	-4.432	1.952
40 minutes	8	-6.942	-1.584	-2.451	1.248	3.803	2.306	2.325	-5.732	-4.120	2.310
45 minutes	9	-6.483	-1.174	-2.312	1.317	3.590	2.352	2.534	-5.241	-3.776	2.636
50 minutes	10	-6.155	-0.961	-2.292	1.382	3.392	2.347	2.783	-4.811	-3.432	2.804
55 minutes	11	-5.971	-0.889	-2.417	1.224	3.039	2.246	2.893	-4.571	-3.295	2.798
1 hour	12	-5.717	-0.835	-2.406	1.135	2.816	2.138	2.948	-4.333	-3.152	2.770
2 hours	24	-4.012	-0.487	-2.465	0.547	1.106	1.786	2.160	-3.322	-2.189	1.966
3 hours	36	-2.266	0.415	-1.551	0.853	1.467	2.442	2.083	-2.513	-0.807	2.010
1 day	67	-0.124	1.482	-0.139	1.489	2.119	2.786	2.623	-0.858	0.945	2.830
2 days	134	-1.295	-0.270	-0.808	-0.698	1.401	1.550	1.590	-0.452	0.767	2.499
1 week	335	-0.639	-0.895	-0.235	-1.113	1.663	1.033	1.617	0.906	1.157	1.722
2 weeks	670	0.071	-0.837	-0.138	-1.440	1.537	0.308	1.415	1.067	0.855	1.390
1 month	1675	0.185	-0.956	0.083	-1.253	1.860 -	0.920	1.066	1.944	1.458	1.745
2 months	3350	0.012	-1.252	-0.149	-1.250	2.093 -	0.949	0.306	2.258	1.693	1.692

Table 4: Homoscedastic Lo and MacKinlay (1988) $Z2$ statistics of the variance ratios of top ten portfolios of NSE stocks arranged by liquidity. The stocks are selected from all the 1382 stocks that traded in the NSE in the period Mar1999-Feb2001. The stocks are arranged by liquidity and the top hundred are selected. The portfolios are then formed from the ten deciles. Impact cost for a buy/sell order of Rs. 10,000 is the liquidity measure.

Duration	q	Heteroscedasticity-consistent $Z2$ statistics of Portfolios (Deciles) returns.									
		1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}	6^{th}	7^{th}	8^{th}	9^{th}	10^{th}
10 minutes	2	0.796	1.232	1.310	3.459	5.849	2.229	1.819	-3.327	-1.618	-0.159
15 minutes	3	-1.469	-0.392	0.281	1.831	4.286	1.566	1.104	-3.317	-2.731	-0.357
20 minutes	4	-2.636	-1.028	-0.478	0.904	3.158	1.107	0.832	-3.183	-2.819	-0.062
25 minutes	5	-3.123	-1.116	-0.857	0.476	2.653	1.198	0.806	-3.075	-2.786	0.316
30 minutes	6	-3.182	-0.974	-0.968	0.505	2.406	1.364	0.989	-2.830	-2.572	0.813
35 minutes	7	-3.136	-0.752	-0.923	0.603	2.250	1.471	1.202	-2.647	-2.347	1.186
40 minutes	8	-3.188	-0.706	-0.992	0.672	2.061	1.478	1.205	-2.469	-2.229	1.428
45 minutes	9	-3.027	-0.529	-0.950	0.715	1.972	1.530	1.329	-2.330	-2.081	1.651
50 minutes	10	-2.892	-0.434	-0.943	0.754	1.883	1.544	1.466	-2.199	-1.922	1.775
55 minutes	11	-2.814	-0.400	-0.992	0.669	1.700	1.491	1.525	-2.143	-1.873	1.786
1 hour	12	-2.697	-0.375	-0.983	0.623	1.583	1.431	1.555	-2.077	-1.815	1.781
2 hours	24	-1.956	-0.227	-1.025	0.315	0.640	1.276	1.197	-1.876	-1.403	1.349
3 hours	36	-1.158	0.206	-0.684	0.515	0.882	1.821	1.229	-1.548	-0.548	1.435
1 day	67	-0.071	0.840	-0.070	0.991	1.394	2.232	1.736	-0.597	0.696	2.150
2 days	134	-0.827	-0.173	-0.473	-0.508	1.013	1.295	1.160	-0.349	0.606	1.988
1 week	335	-0.468	-0.661	-0.165	-0.896	1.346	0.914	1.315	0.774	0.994	1.462
2 weeks	670	0.056	-0.673	-0.109	-1.230	1.325	0.279	1.217	0.961	0.771	1.235
1 month	1675	0.161	-0.840	0.072	-1.138	1.694	-0.855	0.965	1.846	1.378	1.629
2 months	3350	0.011	-1.165	-0.138	-1.179	1.967	-0.901	0.287	2.205	1.652	1.626

Table 5: Heteroscedasticity-consistent Lo and MacKinlay (1988) $Z2$ statistics of the variance ratios of top ten portfolios of NSE stocks arranged by liquidity. The stocks are selected from all the 1382 stocks that traded in the NSE in the period Mar1999-Feb2001. The stocks are arranged by liquidity and the top hundred are selected. The portfolios are then formed from the ten deciles. Impact cost for a buy/sell order of Rs. 10,000 is the liquidity measure.

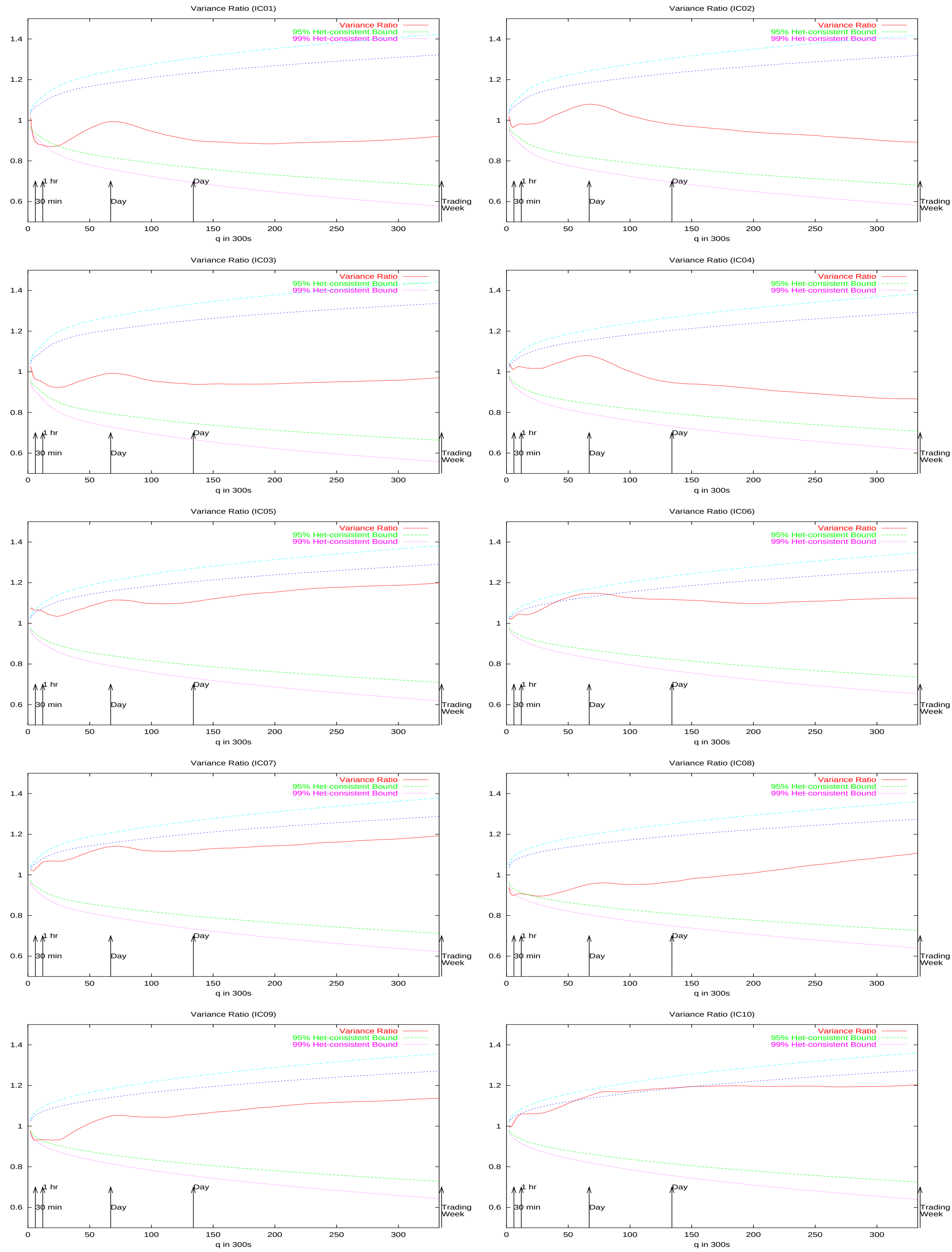


Figure 5: Variance Ratios of portfolio deciles, arranged by impact cost on an order of Rs. 10,000. The sample period is Mar 1999 — Feb 2001.

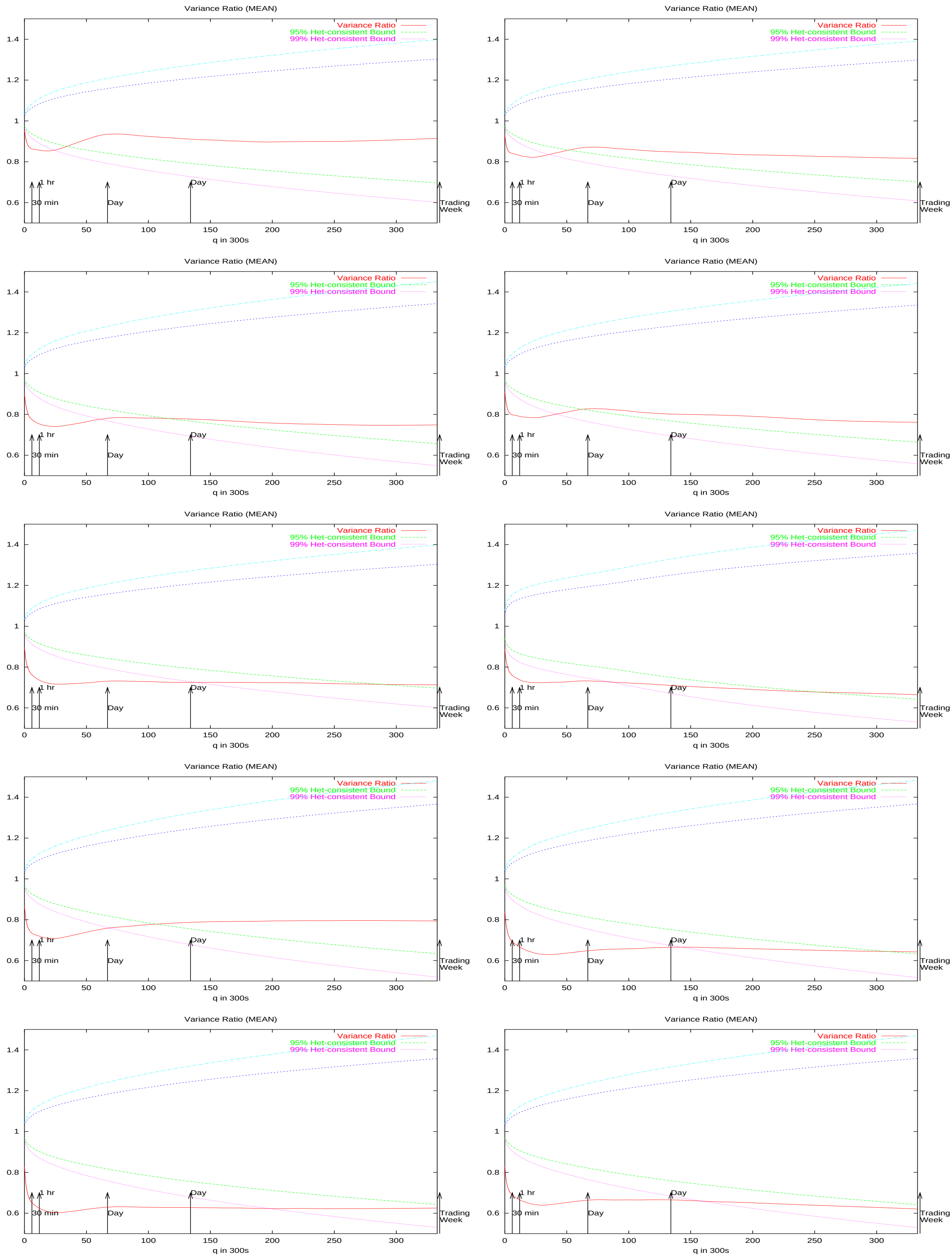


Figure 6: Cross-sectional average Variance Ratios of portfolio deciles, arranged by impact cost on an order of Rs. 10,000. Variance Ratios were calculated on 5-minute returns over the period is Mar 1999 — Feb 2001.

Duration	q	Portfolios (Deciles) Cross-sectional averages of $VR(q)$'s.									
		1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}	6^{th}	7^{th}	8^{th}	9^{th}	10^{th}
10 minutes	2	0.955	0.931	0.899	0.910	0.895	0.887	0.866	0.841	0.827	0.830
15 minutes	3	0.916	0.886	0.850	0.860	0.840	0.835	0.807	0.780	0.750	0.766
20 minutes	4	0.891	0.862	0.820	0.830	0.810	0.804	0.776	0.744	0.710	0.734
25 minutes	5	0.878	0.850	0.799	0.813	0.792	0.786	0.759	0.722	0.685	0.716
30 minutes	6	0.871	0.845	0.788	0.806	0.779	0.773	0.748	0.709	0.670	0.703
35 minutes	7	0.866	0.843	0.781	0.801	0.769	0.765	0.740	0.699	0.659	0.695
40 minutes	8	0.862	0.840	0.774	0.798	0.762	0.759	0.734	0.692	0.652	0.689
45 minutes	9	0.862	0.839	0.770	0.797	0.757	0.754	0.730	0.686	0.646	0.683
50 minutes	10	0.861	0.838	0.766	0.797	0.752	0.750	0.728	0.683	0.641	0.679
55 minutes	11	0.859	0.836	0.762	0.795	0.747	0.747	0.726	0.678	0.636	0.676
1 hour	12	0.859	0.834	0.759	0.794	0.743	0.744	0.724	0.674	0.632	0.672
2 hours	24	0.854	0.822	0.742	0.784	0.718	0.724	0.707	0.640	0.604	0.645
3 hours	36	0.875	0.833	0.747	0.792	0.717	0.725	0.717	0.629	0.606	0.641
1 day	67	0.934	0.870	0.779	0.827	0.731	0.732	0.757	0.647	0.629	0.663
2 days	134	0.911	0.849	0.778	0.802	0.725	0.711	0.787	0.665	0.628	0.666
1 week	335	0.915	0.817	0.748	0.761	0.713	0.665	0.793	0.643	0.625	0.621
2 weeks	670	0.931	0.797	0.736	0.699	0.713	0.619	0.810	0.629	0.645	0.566
1 month	1675	0.953	0.746	0.765	0.625	0.706	0.551	0.702	0.602	0.703	0.525
2 months	3350	0.910	0.683	0.787	0.558	0.702	0.453	0.652	0.594	0.739	0.489

Table 6: Cross-sectional averages of the variance ratios of top ten portfolios of NSE, arranged by liquidity. The stocks are selected from all the 1382 stocks that traded in the NSE in the period Mar1999-Feb2001. The stocks are arranged by liquidity and the top hundred are selected. The portfolios are then formed from the ten deciles. Impact cost for a buy/sell order of Rs. 10,000 is the liquidity measure.

References

Andersen, T., Bollerslev, T., and Das, A. (2001). Variance-ratio statistics and high-frequency data: Testing for changes in intraday volatility patterns. *Journal of Finance*, LVI(1):305–327.

Andersen, T. G. and Bollerslev, T. (1997). Intraday seasonality and volatility persistence in financial markets. *Journal of Empirical Finance*, 4:115–58.

Andrews, D. W. K. (1991). Heteroscedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59:817–858.

ap Gwilym, O., Buckle, M., and Thomas, S. (1999). The intra-day behaviour of key market variables for liffe derivatives. In Lequeux (1999), chapter 6, pages 151–189.

Baillie, R. and Bollerslev, T. (1990). Intraday and intermarket volatility in foreign exchange markets. *Review of Economic Studies*, 58:565–585.

Brock, W. A. and LeBaron, B. (1989). Lquidity constraints in production based asset pricing models. Technical Report 3107, National Bureau of Economic Research.

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

Cecchetti, S. G. and sang Lam, P. (1994). Variance-ratio tests: Small-sample properties with an application to international output data. *Journal of Business Economics & Statistics*, 12(2):177–186.

Chandrakanta, L. M. S., Mehta, J. S., and Swamy, P. V. A. B. (1989). On the use of variance ratios on the analysis of nonstationary times series. Finance and Economics Discussion Series 97, Board Of Governors of the Federal Reserve System.

Choi, I. (1999). Testing the random walk hypothesis for real exchange rates. *Journal of Applied Econometrics*, 14:293–308.

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

Cochrane, J. H. (1988). How big is the random walk in the GNP? *Journal of Political Economy*, 96:893–920.

Cogley, T. (1990). International evidence on the size of the random walk in output. *Journal of Political Economy*, 98(3):501–518.

Cooper, J. C. B. (1982). World stock markets: some random walk tests. *Applied Economics*, 14(5):515–531.

Dacorogna, M. M., Gencay, R., Muller, U. A., Olsen, R. B., and Pictet, O. V. (2001). *An Introduction to High-Frequency Finance*. Academic Press, 525 B Street, Suite 1900, San Diego California 92101-4495.

Dunis, C., editor (1996). *Forecasting Financial Markets: Exchange Rates, Interest Rates and Asset Management*. Wiley Series in Financial Economics and Quantitative Analysis. John Wiley & Sons, Baffins Lane, Chichester, West Sussex PO19 1UD, England.

Dunis, C. and Zhou, B., editors (1998). *Nonlinear Modelling of High Frequency Financial Time Series*. Wiley Series In Financial Economics And Quantitative Analysis. John Wiley & Sons, Baffins Lane, Chichester, West Sussex PO19 1UD, England.

Durlauf, S. N. (1991). Spectral testing of the martingale hypothesis. *Journal of Econometrics*, 50:355–376.

Fama, E. F. and French, K. R. (1987). Dividend yields and expected stock returns. Technical Report 3-88, Anderson Graduate School of Management, University of California at Los Angeles.

Faust, J. (1988). The variance ratio test: Statistical properties and implementation. Working Paper 88-08, Federal Reserve Bank of Kansas City.

Faust, J. (1992). When are variance ratio tests for serial dependence optimal? *Econometrica*, 60(5):1215–1226.

Fong, W. M. and Ouliaris, S. (1995). Spectral tests of the martingale hypothesis for exchange rates. *Journal of Applied Econometrics*, 10(3):255–271.

Gallant, A. R. (1981). On the bias in the flexible functional forms and an essentially unbiased form: The fourier flexible form. *Journal of Econometrics*, 15:211–245.

Gavridis, M. (1998). Modelling with high frequency data: A growing interest for financial economists and fund managers. In Dunis and Zhou (1998), chapter 1, pages 3–22.

Goodhart, C. A. E. and Figliuoli, L. (1991). Every minute counts in financial markets. *Journal of International Money and Finance*, 10:23–52.

Goodhart, C. A. E. and O'Hara, M. (1997). High frequency data in financial markets: Issues and applications. *Journal of Empirical Finance*, 4:73–114.

Granger, C. W. J. and Ding, Z. (1994). Stylized facts on the temporal and distributional properties of daily data from speculative markets. Technical report, University of California, San Diego and Frank Russell Company, Tacoma, Washington.

Guillaume, D. M., Dacorogna, M. M., Dave, R. R., Muller, U. A., Olsen, R. B., and Pictet, O. V. (1994). From the bird's eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets. Technical report, Olsen & Associates Research Group.

Harris, L. (1986). A transactions data study of weekly and intradaily patterns in stock returns. *Journal of Financial Economics*, 16:99–117.

Hasbrouck, J. A. (1991). Measuring the information content of stock trades. *Journal of Finance*, 46:179–207.

Kemp, A. G. and Reid, G. C. (1971). The random walk hypothesis and the recent behaviour of equity prices in Britain. *Economica*, 38:28–51.

Kim, M. J., Nelson, C. R., and Startz, R. (1991). Mean reversion in stock prices? a reapprisal of the empirical evidence. *Review of Economic Studies*, 58:515–528.

Krol, R. (1992). Trends, random walks and persistence: An empirical study of disaggregated u.s industrial production. *Review of Economics and Statistics*, 74(1):154–159.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, 53:1315–1335.

Lequeux, P., editor (1999). *Financial Markets Tick by Tick*. John Wiley & Sons, Baffins Lane, Chichester, West Sussex PO19 1UD, England.

Lippi, M. and Reichlin, L. (1991). Trend-cycle decompositions and measures of persistence: Does time aggreagation matter? *Economic Journal*, 101(405):314–323.

Liu, C. and He, J. (1991a). Do exchange rates follow random walks: A heteroscedasticity-robust autocorrelation test. *International Economics Journal*, 5(3):39–48.

Liu, C. and He, J. (1991b). A variance-ratio test of random walks in forein exchange rates. *Journal of Finance*, 46(2):773–785.

Lo, A. W. and MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies*, 1(1):44–66.

Lo, A. W. and MacKinlay, A. C. (1989). The size and the power of the variance ratio test in finite samples: a monte carlo investigation. *Journal of Econometrics*, 40:203–238.

Lockwood, L. J. and Linn, S. C. (1990). An examination of stock market return volatility during overnight and intraday periods, 1964-1989. *Journal of Finance*, 45:591–601.

Low, A. and Muthuswamy, J. (1996). Information flows in high frequency exchange rates. In Dunis (1996), chapter 1, pages 3–32.

MacGregor, P. (1999). The sources, preparation and use of high frequency data in the derivative markets. In Lequeux (1999), chapter 11, pages 305–311.

Madhavan, A. (1992). Trading mechanisms in securities markets. *Journal of Finance*, 47:607–642.

Malliaris, A. G. and Urrutia, J. L. (1990). How big is the random walk in macroeconomic time series. *Economic Letters*, 34(2):113–116.

Marinelli, C., Rachev, S., and Roll, R. (2001). Subordinated exchange-rate models: Evidence for heavy-tailed distributions and long-range dependence. *Mathematical and Computer Modelling*, 34(9-11):1–63.

McNish, T. A. (1993). A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. *Journal of International Money and Finance*, 12(4):413–438.

McNish, T. H. and Wood, R. A. (1990a). An analysis of transactions data for the toronto stock exchange. *Journal of Business & Finance*, 14:458–491.

McNish, T. H. and Wood, R. A. (1990b). A transaction data analysis of the variablity of common stock returns during 1980-1984. *Journal of Business & Finance*, 14:99–112.

McNish, T. H. and Wood, R. A. (1991). Hourly returns, volume, trade size, and number of stocks. *Journal of Financial Research*, pages 303–315.

McNish, T. H. and Wood, R. A. (1992). An analysis of intraday patterns in bid/ask spreads for nyse stocks. *Journal of Finance*, 47:753–764.

Mikkelsen, H. O. (1992). A note of the superiority of variance ratio tests for random walks against long memory alternatives. University of Aarhaus Economics Memo 1992-6, University of Aarhaus.

Nelson, C. R. and Plosser, C. I. (1982). Trends and random walks in macroeconomic time series. *Journal of Monetary Economics*, 10:139–162.

Noreen, E. W. (1989). *Computer-intensive methods for testing hypotheses: An introduction.* John Wiley & Sons, New York.

O'Hara, M. (1995). *Market Microstructure Theory.* Basil Blackwell, Oxford.

Pan, M.-S., Chan, K. C., and Fok, R. C. W. (1997). Do currency futures follow random walks? *Journal of Empirical Finance*, 4:1–15.

Patnaik, T. C. and Shah, A. (2002). An empirical characterization of the national stock exchange, using high-frequency data. Technical report, IGIDR, Mumbai.

Poterba, J. and Summers, L. (1988). Mean Reversion in Stock Returns: Evidence and Implications. *Journal of Financial Economics*, 22:27–60.

Priestly, M. B. (1982). *Spectral Analysis and Time Series.* Academic Press.

Proietti, T. (1996). Persistence of shocks on seasonal processes. *Journal of Applied Econometrics*, 11(4):383–398.

Richardson, M. and Smith, T. (1991). Tests of financial models in the presence of overlapping observations. *Review of Financial Studies*, 4(2):227–254.

Richardson, M. and Stock, J. H. (1989). Drawing inferences from statistics based on multiyear asset returns. *Journal of Financial Economics*, 25:323–348.

Roll, R. (1984). A simple measure of the implicit bid-ask spread in an efficient market. *Journal of Finance.*

Ronen, T. (1997). Tests and properties of variance ratios in microstructure studies. *Journal of Financial and Quantitative Research*, 32:183–204.

Smith, M. G. and Naik, N. Y. (1997). Analytic small sample bias and standard error calculations for tests of serial correlation in market returns. Working Paper 97/06, Monash University Department of Econometrics and Business Statistics.

Stoll, H. R. and Whaley, R. E. (1990). Stock market structure and volatility. *Review of Financial Studies*, 3:37–71.

Szroeter, J. (1978). Generalized variance-ratio tests for serial correlation in multivariate regression models. *Journal of Econometrics*, 8(1):47–59.

Tse, Y. K. and Zhang, X. B. (2002). The variance ratio test with stable paretian errors. *Journal of Time Series Analysis*, 23(1):117–126.

Wilson, D. P. and Tonascia, J. (1971). Tables for Shortest Confidence Intervals on the Standard Deviation and Variance Ratio from Normal Distributions. *Journal of the American Statistical Association*, 66(336):909–912.

Wood, R. A., McNish, T. H., and Ord, J. K. (1985). An investigation of transaction data for NYSE stocks. *Journal of Finance*, 40:723–741.

Wright, J. H. (2000). Alternative variance-ratio tests using ranks and signs. *Journal of Business Economics & Statistics*, 18(1):1–9.