Implications of rounded reported prices on financial assets on the estimation of the characteristics of the underlying variables

by

Emilio Venezian

Venezian Associates

e-mail: Emilio_venezian@yahoo.com

Telephone: 886-2-2965-2616

Fax: 886-2-2965-2616

This is a draft provided for review and comment under the condition that it will remain confidential. Please address any comments and suggestions to

emilio_venezian@yahoo.com

Implications of rounded reported prices on financial assets on the estimation of the characteristics of the underlying variables

by

Emilio Venezian

Abstract

This paper points out that the usual practice of assuming that prices are continuous needs reappraisal. It uses an approximate analysis to estimate the bias in some of the important basic statistics of rates of return when they are computed from rounded prices. The methodology is simple and does not depend on assumptions about the distribution of the underlying increments in the logarithm of continuous prices. It shows that cannot obtain valid measures about the underlying variables by simply obtaining estimates from the reported prices. Simulations based on simple models are used to validate the model, in the sense that the results show that at least when the tick sizes are small the simulated data are in close accord with the approximation.

The rounding of prices creates problems at both the academic and practical level. At the academic level, the existence of rounded prices distorts tests, such as the runs and variance ratio tests, used to examine the behavior of stock prices. The order of magnitude of the biases is determined by the ratio of the tick size to the product of stock price times the volatility over the period over which prices are sampled. Thus increasing the sampling frequency reduces the standard deviation of the estimates but increases their bias. Since the biased statistics include the variance of the rate of return and the autocorrelation coefficient, they affect the accuracy of put and call options if statistics based on trading prices are rounded. The biases generally favor the issuers of options, but the biases at current tick prices and volatilities are not large for stocks of prices above $10. Implied volatility may be a better tool for deriving fair prices but a detailed examination of this issue is needed. The existence of these biases has implications for both market behavior and for academic analysis.

Introduction

Traditionally, the statistical properties of the rates of return for financial assets have been estimated form data on recorded transaction prices. Tests regarding the process that generates the prices have been evaluated by statistics of the estimates based on a model of continuous prices. As an example, several models of price evolution assume that rates of return in successive periods are independent of each other, requiring that the correlation coefficients be zero. Correlation coefficients that are significantly different from zero are viewed as rejecting the hypothesis of independence of the increments. Such procedures would clearly be appropriate if prices were not rounded.

In practice, prices are always rounded. It is conceivable that the practice of recording discrete prices, by itself, affects the estimated serial correlation coefficients. We have an enormous corpus of literature on two issues that are closely related to the one under discussion. One is literature on the estimation of structural parameters when the data are grouped into ranges rather than as individual points, the other is related to estimation when there are errors in the variables. Rounding of transaction prices groups can be viewed as grouping the data into intervals, or as introducing error into the "true" or "underlying" prices. From the perspective of this paper, these two strands are at least marginally relevant to the problem.

However, the issue considered in that body of work is much simpler than that facing us in this problem. Two important reasons set it apart. It the literature on grouping and errors in variables the distortion is in the variables of interest. In the case discussed here the errors are in the prices, but the variables of interest are either differences or ratios of successive prices, often with a non-linear transformation, such as a logarithmic one. In view of the evidence of bias in estimation when we have errors in variables or grouped data it seems misguided to use a naive methodology without enquiring into the possible distortions that might result in our, more complicated case. It seems foolhardy to conduct estimations as though the prices were recorded as continuous variables and draw conclusions from tests based on that assumption without some form of validation that the process leads to sensible conclusions. It may be dangerous, not merely foolhardy, to use such estimates in making financial decisions.

I believe we need to consider if there might be systematic biases introduced into the estimation by the data generation process. In the following section I develop a simple first-order model and use it to approximate the relationship between statistics of the population of uncorrelated increments that are assumed to generate prices and the statistics inferred from reported prices.[1] I have concentrated on two statistics: the variance of the increment and the autocorrelation at a one period lag. The first one is of practical importance since the formulas for pricing options rely on having a valid measure

[1] I have omitted any consideration of how actively traded the issues might be. That is a potentially important consideration, but the task is complicated enough as it stands.

of this variable or of its square root, the standard deviation, if the estimates are biased the decisions will be suboptimal. The second one is important because is lies at the root of tests of the hypothesis that successive increments are uncorrelated.[2] The simple argument presented suggests that such biases might be present.

Prior Work

The problems which I will discuss have, not surprisingly, been discussed in the literature. If anything, the surprise should be that with all the papers on the closely related fields of data grouping and errors in variables the problem did not receive attention until the mid-nineteen eighties.

Gottlieb and Kalay, 1985, established that the variance estimator is upward biased in the presence of rounding of prices. Their equations are framed in terms of differences in prices rather than rates of return or differences in the logarithm of prices. The analytical work develops expected values of the even moments of the differences in prices measured at t and $t + \Delta t$. The numerical work shows ratios of the estimated to true values of the instantaneous standard deviation of the rate of return and the values of the estimated excess kurtosis in the rate of return. It is not clear whether the rate of return is a “linear return”, consistent with their equations which represent the moments of the difference in price, or a “logarithmic return”, consistent with the lognormal model of the underlying price process.

Cho and Frees, 1988, proposed an estimator of volatility based on the time of first passage out of a given price and argue that the biases in that estimator are much lower than those in the traditional estimator. The procedure of Cho and Frees requires transaction level data for implementation.

Ball, 1988, examined a model which provides more insight into the problem tackled by Gottlieb and Kalay, 1985, by developing limiting relations for the bias in the variance estimation. These are an improvement over the purely tabular results of Gottlieb and Kalay. He also developed a method of incorporating the expected limiting value of the bias into the estimation procedure and provides small sample Monte Carlo simulations of the performance of the variance estimators.

Harris, 1990, innovates in at least two ways. One is the introduction of bid-asked spreads into the relation between estimated and actual parameter values, the other is the consideration of serial covariance in the price series. Harris finds that the serial covariance at one lag has an expected value that is lower than the true value and the variance has an expected value that is higher than the true value.

Unfortunately, the results, however, appear to have been widely ignored. This may be due to the complexity of the models. A typical example is provided by Campbell, Lo, and MacKinley, 1999.

[2] Both the traditional tests of autocorrelation and the variance ratio tests would be using an incorrect null basis if the process of rounding the prices creates a bias in the estimated coefficient.

They advocate run tests to establish serial independence, though these tests are theoretically not valid with rounded prices. In fact they ignore the effect of rounding in determining the "optimal frequency" for sampling financial data. Yet the book does have a section that deals with the effects of rounding and other microstructure effects.

In this paper I present approximations based on a simple first order Taylor expansion of the relations found when rounding is present. The virtue of this approach is that it can handle logarithmic rates of return in the context of rounding based on linear prices. The approach uses basically the same assumptions as those of Ball and Harris. In particular, I assume as they do, that in the absence of prior information the real price of the security is uniformly distributed over the interval of rounding, that the logarithm of price is a random walk, and I use an ad hoc assumption about the rounding process.[3] On the other hand, I make no assumptions about the distribution of the increments in real price, allowing any distributional form and serial interdependence. My results for the sign and magnitude of the bias in the variance are identical with those of Ball and Harris; for the serial covariance at a lag of one period my results also agree with those of Harris. However, the serial correlation coefficient of the rate of return has a smaller bias than the serial covariance (because of the bias in the variance partly counteracts the difference) and under if the increments are not independent the bias in the correlation becomes positive when the underlying price increment process has correlation coefficients below -0.50.

Simulation results show that the first order Taylor approximation does not do very well, in some cases, especially when the tick size is roughly the same as the product of the stock price times the variance of the increments over one period. Higher order approximations do not appear to improve the performance. Two factors appear to account for this seemingly paradoxical effect. One is that if the tick size is much greater that the product of price times standard deviation then most trades will exhibit no change in reported price and the assumption that the distribution of prices is uniform breaks down; in the limit, all trades will be restricted within one tick and the distribution of underlying prices will be determined by the form of the underlying process. The other is that for series of 1,000 or more observations the range in prices will differ substantially among the simulated series and the initial price times the standard deviation of the innovations may be a poor proxy. Dealing with these problems requires additional research.

A model of the data generating process and its application

Suppose that the "real", but unknown, price of the security is $P(t)$ and, as often postulated in financial models, the price is generated through a process that generates the price is given by

[3] I assume all transactions are real, so I do not take into consideration bid-asked spreads. Taking those into account would, in my view, require modeling of the streams of market and limit orders over the course of trading and non-trading periods.

$$p(t) = p(t-1) + \varepsilon(t) \tag{1}$$

The variable $p(t)$ is often called "the log price process" and is not further defined. The variable $\varepsilon(t)$ is a random change or increment in the pricing process. Two possible more detailed definitions appear to make sense: one is that $p(t)$ is the log of the reported price of the asset at time *t*, the other is that it represents the underlying or true price at that time. Here *t* is viewed as a discrete variable, presumably representing the end of some calendar period; specification in continuous time are sometimes used.

In the first interpretation the random variable must be such that the change in the price is an integer multiple of the "tick size," the smallest recorded price change. This implies that any decision by the market or a market regulator to change the value of the tick size will have a real effect on the economy. Even if we were to accept this notion, however, another objection would exist. In the presence of rounded prices, if Equation 1 is to hold it is necessary that the distribution of the errors cannot be stationary. This can be shown by example; if the tick size is one dollar, the change in the log price to bring the price from 100 to 101 is just not the same as that needed to bring the price from 101 to 102 or from 99 to 100. This may also impose correlation between the expectation of the random variable in consecutive periods because an increase in price from 100 to 101, will change the admissible levels of error in the opposite direction of that which would result from a change from 100 to 99. Further, since the admissible changes in log price change we need to specify how the probability of these changes would be affected. I will not devote attention to this interpretation.

In the second case the economy is not dependent on the decision of how precisely prices are to be recorded, but the data that we have does not represent "the log price process" exactly, since the two recorded prices are rounded to within a tick and therefore measure the log price process with error. In that case we need to determine how the rounding affects any proposed test or estimate.

The relation between the underlying price and its logarithm may be expressed as

$$P(t) = e^{p(t)} = \tau[q(t) + \theta(t)] \tag{2}$$

where τ is the "tick size" or unit to within which prices are rounded, and $q(t)$ is chosen so that $\tau q(t)$ is the largest integer less than the price. It follows that $0 \le \theta(t) < 1$. Then we can express the log of the real price as $p(t) = \ln[\tau q(t) + \theta(t)]$, and re-write it as:

$$p(t) = \ln\left\{\tau q(t)\left(1 + \frac{\theta(t)}{q(t)}\right)\right\} = \ln(\tau q(t)) + \ln\left(1 + \frac{\theta(t)}{q(t)}\right) \tag{3}$$

whereas the log of the reported price will be

$$p_R(t) = \ln\left\{\tau \mathrm{INT}\left(\frac{e^{p(t)}}{\tau} + \frac{1}{2}\right)\right\} = \ln(\tau q(t)) + \ln\left(1 + \frac{INT[\theta(t) + 0.5]}{q(t)}\right) \tag{4}$$

From these relationships we find, after some algebra, that:

$$p_R(t) = p(t) + \left[\ln\left(1 + \frac{\tau INT[\theta(t) + 0.5]}{\tau q(t)}\right) - \ln\left(1 + \frac{\tau\theta(t)}{\tau q(t)}\right)\right] \tag{5}$$

If the tick size is small relative to the reported price, which is known, is given by:

$$P_R(t) = \mathrm{INT}\left(e^{p(t)}\right) = \tau\mathrm{INT}[q(t) + \theta(t)] \tag{6}$$

I have used the notation $\mathrm{INT}(x)$ to represent the integer part of the continuous quantity x.

In practice we are interested in situations in which the price is much larger than the tick size, so the quantity $\tau q(t)$ is very close to both the true or reported price, so we can write.

$$\begin{aligned} p_R(t) &\cong \ln\left\{\tau\mathrm{INT}\left(\frac{e^{p(t)}}{\tau} + \frac{1}{2}\right)\right\} = \ln\left\{\tau q(t)\left(1 + \frac{\mathrm{INT}[\theta(t) + 0.5]}{q(t)}\right)\right\} \\ &= \ln(\tau q(t)) + \ln\left(1 + \frac{\mathrm{INT}[\theta(t) + 0.5]}{q(t)}\right) \end{aligned} \tag{7}$$

Define $\nu(t) = \frac{\tau I\{NT[\theta(t)+0.5]-\theta(t)\}}{P(t)}$, so that we have, approximately, $-\frac{\tau}{2P(t)} \le \nu(t) \le \frac{\tau}{2P(t)}$

Using a Taylor series expansion we get, as a first approximation:

$$p_R(t) \cong \ln\left[\tau[q(t) + \theta(t)]\right] \cong \ln[\tau q(t)] + \nu(t) \tag{8}$$

The return for the period beginning at t-1 and ending at t is

$$r_R(t,1) \cong p_R(t) - p_R(t-1) = p(t) - p(t-1) + [\nu(t) - \nu(t-1)] \tag{9}$$

It follows that if we approximate the local distribution of $\nu(t)$ by a uniform distribution between and $-\frac{\tau}{2P(t)} \le \nu(t) \le \frac{\tau}{2P(t)}$, and assume that successive values of ν are independent,[4] then:

$$E[r_R(t)] \cong E[r(t)] \tag{10}$$

and

$$V[r_R(t)] \cong V[r(t)] + 2V[\nu(t)] = \sigma^2\left(1 + \frac{1}{6}\left(\frac{\tau}{P\sigma}\right)^2\right) \tag{11}$$

This can also be written as:

$$\frac{V[r_R(t)]}{V[r(t)]} \cong \left(1 + \frac{1}{6}\left(\frac{\tau}{P\sigma}\right)^2\right) \tag{12}$$

---

4 Evidence validating these assumptions is presented in the appendix.

It is worth noting that the development of this approximation has made no assumptions about the distribution of the increments in the logarithm of price. From the above we find that:

$$E[r_R(t)r_R(t-1)] = E\big[\big(r(t)+\nu(t)\big)\,\big(r(t-1)+\nu(t-1)\big)\big] \tag{13}$$

Thus to this level of approximation the covariance is given by

$$\mathrm{Cov}[r_R(t)r_R(t-1)] = \mathrm{Cov}[r(t)r(t-1)] - \frac{1}{12}\left(\frac{\tau}{P\sigma}\right)^2 \tag{14}$$

This represents a clear negative bias; the expected value of the covariance based on reported rates of return will be lower than that based on the real rate of return.

Taking into account the distortion of the tick size on the estimated variance of the rate of return, the correlation coefficient should be approximately

$$\rho\big(r_R(t), r_R(t+1)\big) = \frac{\rho[r(t)r(t-1)]}{1+\frac{1}{6}\left(\frac{\tau}{P\sigma}\right)^2} - \frac{\frac{1}{12}\left(\frac{\tau}{P\sigma}\right)^2}{1+\frac{1}{6}\left(\frac{\tau}{P\sigma}\right)^2} \tag{15}$$

I have used the ratio of the expected value of the covariance to the expected value of the variance. Jensen's inequality implies that the denominator should be smaller than the one used in this approximation. If the continuous data have zero first order correlation the magnitude of the correlation coefficient given above is on the low side.[5]

Note that the effect of the denominator on the first term is to bias the correlation coefficient of the log of true price towards zero whereas the second term imposes a bias that is unambiguously negative. The bias on the covariance is unambiguously negative. The bias on the correlation coefficient, however, is positive if $\rho[r(t)r(t-1)] < -\frac{1}{2}$ and, negative if $[r(t)r(t-1)] > -\frac{1}{2}$; if the true correlation is exactly minus one half there is no bias. The usual null hypothesis is that the increments are uncorrelated, so a negative bias should be anticipated for all sufficiently long series.[6]

The model leads to the conclusion that the third moment is not biased and the autocorrelation coefficients at lags 2 and 3 are biased towards zero. If the increments are independent this would lead to no bias in the estimates.

---

5 If the process is stationary and the number of observations used in estimating the variance is large, the effect of Jensen's inequality will be small.

6 It is interesting to note that Fama, 1975, found that among the 30- Dow Jones stocks, the autocorrelation at a lag on one trading day was **positive** for 22 and negative for 8. Among the 11 stocks for which the correlation coefficient was more than twice its standard deviation, he found 9 with positive signs and 2 with negative signs. The ones with positive sign, of course, would differ from the expected value after allowing for the bias due to a tick size of $0.125 by more than two standard deviations. For the two with significantly negative signs the best estimate I can develop based on the published data is that for American Can the ratio of the correlation coefficient minus the expected bias was about 2.5 standard deviations and for Goodyear the ratio was about 4.

It is worth pointing out that the model indicates that the absolute magnitude of the tick is not the key parameter, nor is the ratio of tick size to asset price. The standard deviation of the increments is also important in the determination of the bias in the volatility estimate. Thus two socks of the same price trading in a market with the same rounding rules may have substantially different biases in the estimates if one has a higher standard deviation than the other.

Testing the simple approximation

Simulations

To test the simple approximations given above (and attempted extensions of it which I will discuss later) I simulated a fairly large number of conditions. For each condition I generated 100 series of 2,500 observations. This is roughly equivalent to following 100 companies of similar initial price for 10 years of daily prices. For each set I generated independent increments from either a normal or uniform distribution to apply to the underlying price of the previous period, these were generated with zero mean and standard deviations of 0.005, 0.01, 0.02, and 0.03. The resulting series of increments was then fixed and used to generate prices of assets whose starting price was 1, 5, 10, 20, and 50. These were then "grouped" at a variety of tick sizes to generate series of reported prices and rates of return, defined as the logarithm of the ration of the reported prices were computed.

The statistic of these rates of return could then be compared with the model's predictions. The comparisons can be done simply because the approximation gives the results in terms of a single relevant variable, $\left(\frac{\tau}{P\sigma}\right)$.

Figure 1 plots the ratio of the estimated variance to the variance used in the simulation as a function of $\left(\frac{\tau}{P\sigma}\right)$ based on the initial price of the asset. Alternatives such as using the average price, the average of the reciprocal of price were explored but they do not have a major impact on the comparison.

The approximation does reasonably is several respects:

1. The one relevant parameter that emerges from a first order approximation is $\left(\frac{\tau}{P\sigma}\right)$, and the results show that this does a fairly good job at organizing the data.

2. The conclusion that the result is not sensitive to the distribution of the increments appears to be correct.

3. It predicts the ratio of the variance estimated from the reported data to the variance used in the simulation quite well for parameter values up to one. For parameter values between one and two it does reasonably well, though there is an indication then the variance of the series of increments has a separate influence on the bias. For parameter values above two the prediction in not so good; the difference between the prediction and the simulated value may differ by roughly 0.5, but that is not bad, considering that empirically the average factor in this range is about 3.

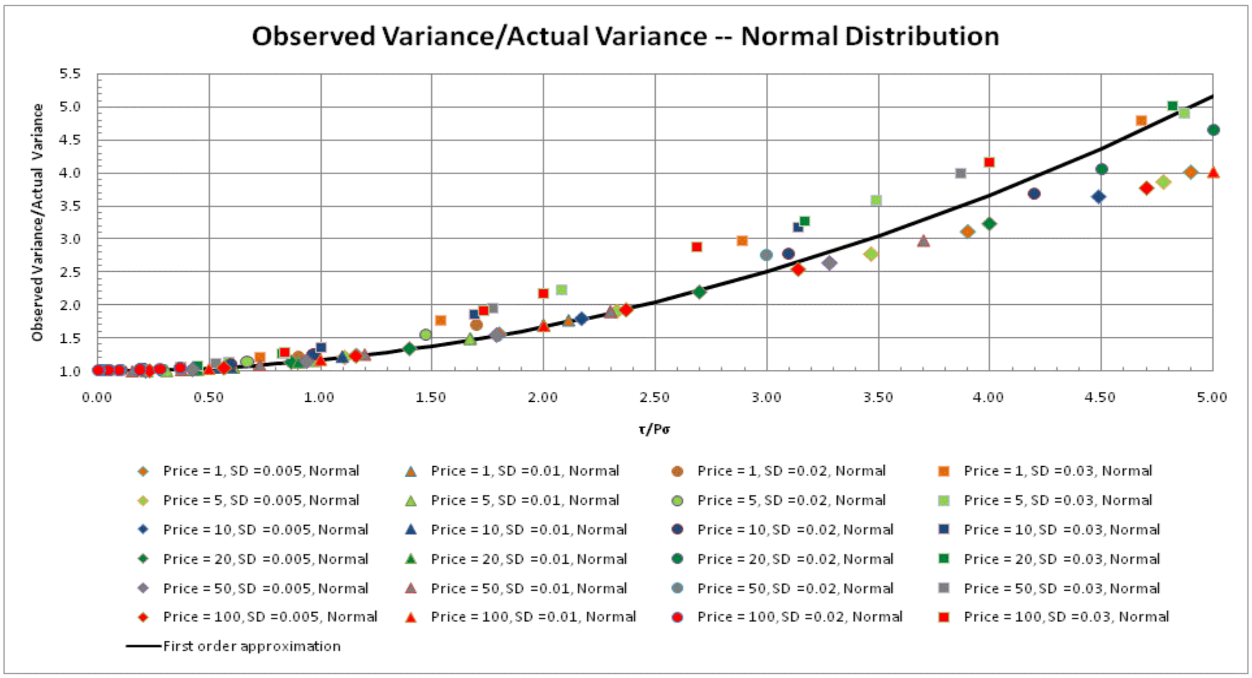

Figure 1a. Effect of tick size on the estimated variance (volatility) of the rate of return when the underlying increments are normally distributed

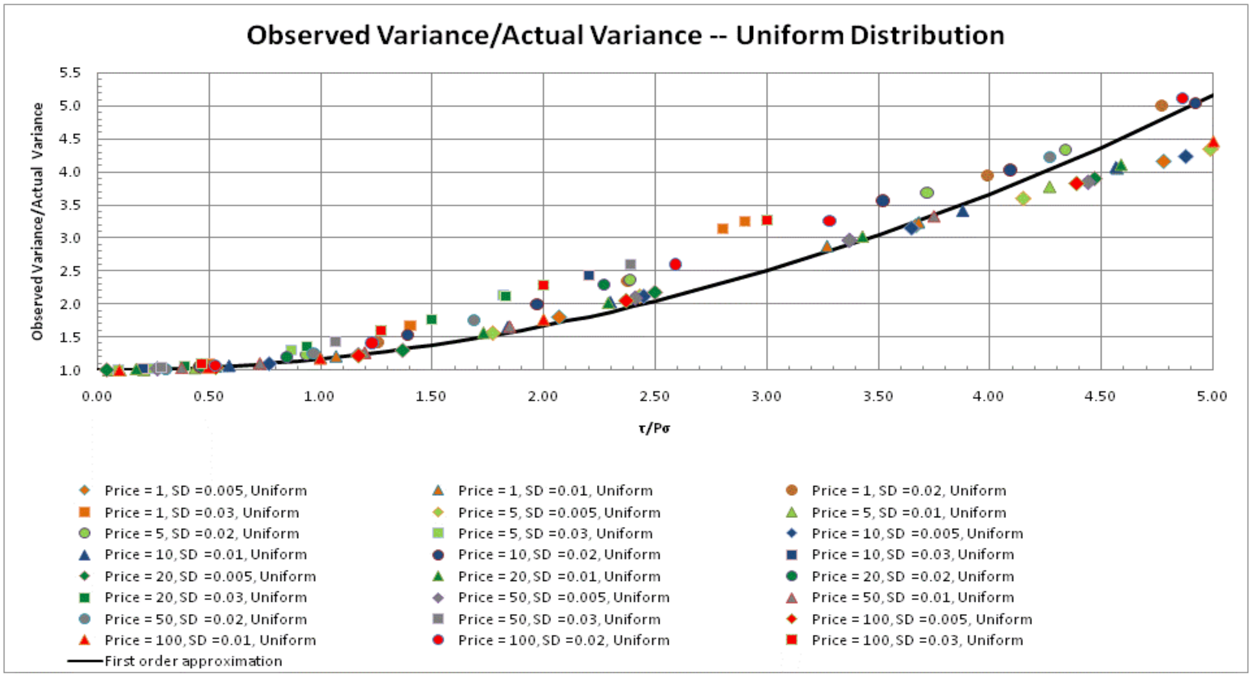

Figure 1b. Effect of tick size on the estimated variance (volatility) of the rate of return when the underlying increments are uniformly distributed

Figure 2 shows the results for the autocorrelation coefficient.

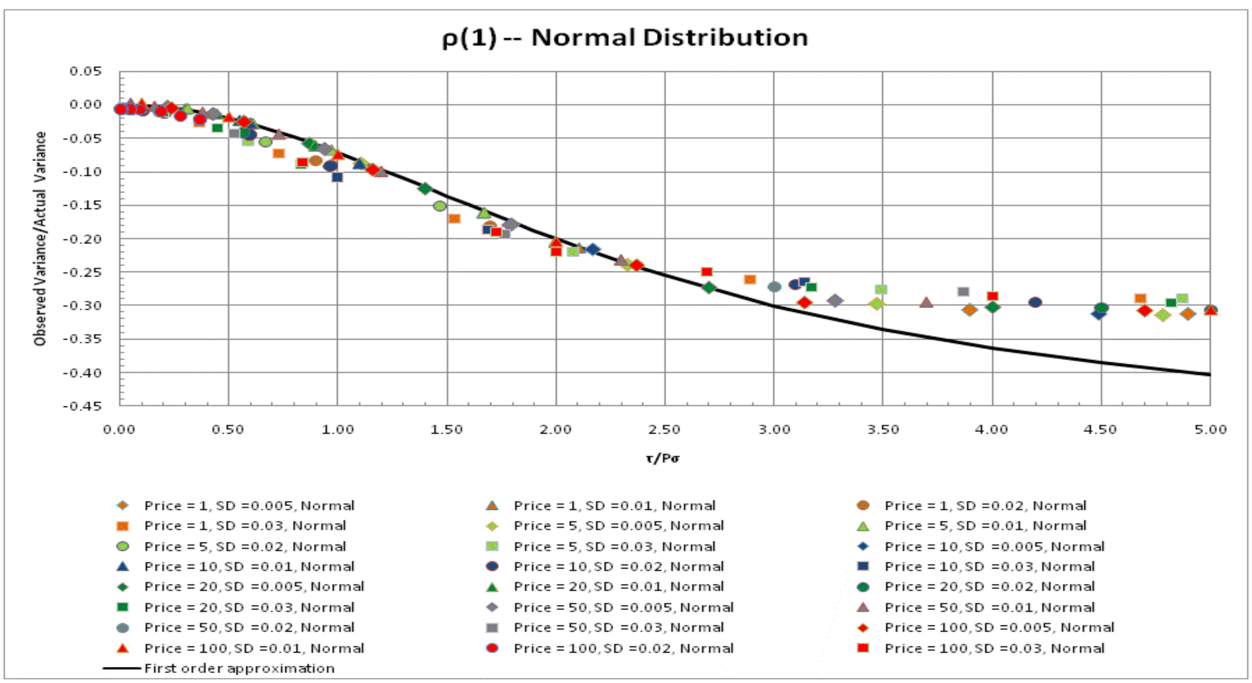

Figure 2a. Effect of tick size on the autocorrelation coefficient at lag 1 of the rate of return when the underlying increments are normally distributed

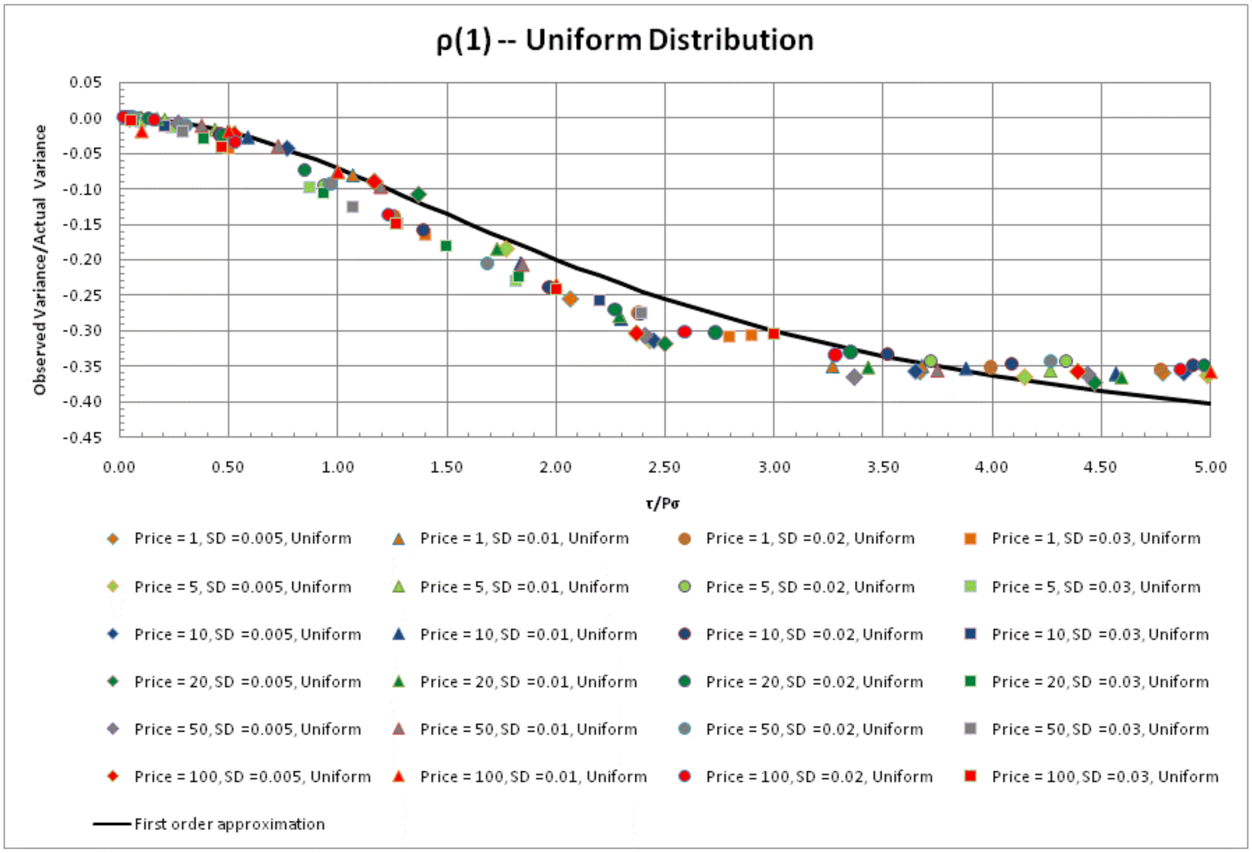

Figure 2b. Effect of tick size on the autocorrelation coefficient at lag 1 of the rate of return when the underlying increments are uniformly distributed

The remarks that were made in relation to Figure 1 apply here as well, except that the evidence for a separate effect by the variance of the increments is weaker here, and that for the normal case there appear to be systematic departures between empirical and model results when the parameter is above three.

These approximations rely on the price remaining approximately constant over time, since otherwise the question of what price should be used becomes a serious question; if the price changes by a factor of two, the range of the estimate is roughly four, depending on whether we use the lowest price or the highest price. The analysis also suggests that caution should be used in interpreting the autocorrelation of stocks that have undergone splits during the period of analysis.

The autocorrelation coefficient estimated above is at a single lag. Since the null hypothesis is that the true correlation is zero, then to the degree of approximation used here the covariance of the reported returns would be zero because no common error term in the relation.

Some implications

Can we reduce the bias?

One way in which we might seek to reduce the bias is by taking periods longer than one day. If the returns are obtained by differences in log prices k trading days apart, and the correlation of returns over not-overlapping periods is calculated, then under the null hypothesis that the increments are independent, corresponding relations would be:

$$\frac{V[r_R(t_k)]}{V[r(t_k)]} \cong \left(1 + \frac{1}{6k}\left(\frac{\tau}{P\sigma}\right)^2\right) \quad (16)$$

$$\rho\big(r_R(t_k), r_R(t_k + k)\big) \cong -\frac{\frac{1}{12k}\left(\frac{\tau}{P\sigma}\right)^2}{1 + \frac{1}{6k}\left(\frac{\tau}{P\sigma}\right)^2} = -\frac{\frac{1}{12}\left(\frac{\tau}{P\sigma}\right)^2}{k + \frac{1}{6}\left(\frac{\tau}{P\sigma}\right)^2} \quad (17)$$

Both are declining functions of the number of days over which the rate of return is taken. Thus the bias is reduced as the number of observations is reduced. This would tend to make the problem disappear because both reductions tend to make the test less powerful. The largest reduction in bias, however, occurs in going from one day to two or three days, hence a major benefit can be obtained at a relatively minor cost in terms of the loss of degrees of freedom. Going in the other direction the bias increases but the precision of the estimator increases. It would be possible to determine an optimal period so as to minimize the joint effects of bias and residual uncertainty, but that would require more faith in the approximations than I believe they deserve. It should be clear, however, that in a market with rounded prices the solution advocated by Campbell, Lo, and MacKinlay, 1999, of going to the shortest possible time increment, is not likely to be effective.

An unsettled theoretical question

It is important to remember that the approximations given above are based on expectations at one time, the expectation taken over the states of nature, rather than over a period of time. In the physical sciences the first is referred to as “ensemble” statistics and the second as “temporal” statistics. The results do not always match, and in this case the match is questionable. In an ensemble sense the price is known, and could be used in the estimation. In the analysis of a time series, the price varies over time and its distribution may not be the same over time; certainly the expected value and standard deviation are not likely to both be the same over time. This in temporal analysis we need to keep in mind that P is not constant and the expected value of $1/P$ is not the reciprocal of the expected value of P. This may be particularly important if we are interested in correlation coefficients at large lags,

Usefulness of variance ratio tests

Variance ratio tests[7] are considered useful in testing alternatives to independent or uncorrelated increments and increments determined by some ARIMA process. As usually described, the variance ratio tests all have expected ratios of one under the assumption of independence. The art is then involved in choosing the periods so the alternate model has expected ratios as different as possible to one. I leave to further analysis the task of determining how the existence of rounding affects the expected value of the ratio under these ARIMA hypotheses. It is, however, useful to point out that fact that the bias (at least in the first order approximation) is limited to lag 1 is important is considering variance ratio tests.

The general version involves the ratio of the variance of the return over q periods to q times the variance of the return over one period. If prices were continuous and the returns were uncorrelated, the expected value of the variance ratio under the null hypothesis would one for all values of q.

The variance ratio using the variance over q periods divided by the variance over a single period can be expressed, under the assumption of continuous prices, as:

$$VR(q) = 1 + 2\sum_{j=1}^{q-1}\left(1 - \frac{j}{q}\right)\rho(j) \tag{18}$$

Under the usual null hypothesis of uncorrelated increments and using only the simple approximation, $\rho(j) = 0$ for all j and the expected value of the ratio would be one. If prices are not continuous but reported only within intervals, the simple approximation I have used suggests that $\rho(j) = 0$ for all j greater than one, but for that value it would be

$$\rho(1) \cong -\frac{\frac{1}{12}\left(\frac{\tau}{P\sigma}\right)^2}{1 + \frac{1}{6}\left(\frac{\tau}{P\sigma}\right)^2} \tag{19}$$

Hence we can write

$$VR(q) = 1 - 2\frac{q-1}{q}\frac{\frac{1}{12}\left(\frac{\tau}{P\sigma}\right)^2}{1 + \frac{1}{6}\left(\frac{\tau}{P\sigma}\right)^2} \tag{20}$$

[7] See, for example, Campbell, J.Y, A.W. Lo, and A.C. MacKinlay, 1997, “The Econometrics of Financial Markets”, Princeton University Press, Princeton, NJ.

Thus as the number of lags increases the expected variance ratio would decrease from $VR(2) = 1 - \frac{\frac{1}{12}\left(\frac{\tau}{P\sigma}\right)^2}{1+\frac{1}{6}\left(\frac{\tau}{P\sigma}\right)^2}$ to $VR(\infty) = 1 - 2\frac{\frac{1}{12}\left(\frac{\tau}{P\sigma}\right)^2}{1+\frac{1}{6}\left(\frac{\tau}{P\sigma}\right)^2}$.

The situation may be ameliorated by using returns over periods longer than one day but, as always, at the corresponding cost of having fewer points in the analysis.

Other implications

The model also makes it clear that trying to get estimates of the parameter in a period in which the stock has experienced a price shock, such as might result from a stock split, is dangerous. It might be better to get one estimate before the shock and a different one after the shock; and then combine them appropriately. This would involve estimating the expected biases in the intervals, correcting the estimates for these biases, and then pooling the results. Similar considerations would apply when large changes in price are due to unexpected market events or by regulatory revisions in the tick size.

It should be clear that the results of analyses that ignore the effect of tick size need to be reassessed, especially if the data used derives from the period before tick sizes were reduced to one cent. Several areas need attention. The usual practice has been to adjust the price series to reflect substantial stock dividends, stock splits, or reverse splits. Even if the market effect were limited to the time of the announcement of the decision, the adjustment would not work well in a world of rounded prices because of the rounding effect itself. We need to find ways of overcoming that problem. Even if that problem did not exist, however, we would need to reassess the issue of "fat tails" given that all the analyses indicate that rounding will give an upward biased estimate of kurtosis. That applies equally well to attempts to measure the fatness of tails by fitting stable Pareto distributions without taking into account the effect of rounding. Preliminary results show that the "range test" used by Fama, 1965, as one method of estimating the characteristic exponent, is not very useful.[8]

[8] One problem is that with ticks of the order of $0.125, as existed in the data used by Fama, a sample of rates of return from rounded prices and using the 75th percentile has a large probability of giving a result of less than 1.9 for the exponent. Another is that the test is based on highly non-linear estimators but substitutes the functions of the expected value for the expected value of the functions. A third one is that the method assumes independence of the part with the whole.

Extensions of the analysis

It is tempting to extend the analysis to higher orders in an attempt to obtain better approximations or to extend the range of approximations. Expanding the Taylor series to include additional terms is a tedious, though straightforward task. The predictions for the variance and the autocorrelation coefficient do not improve the correspondence between simulation and approximation. They do make clear, however, that the parameter $\frac{\tau}{P\sigma}$ is not the only relevant one; powers of σ^2 also appear in these approximations.

The approximations suggest that autocorrelation coefficients beyond that at lag one might also be negative, at least under the null hypothesis that the increments themselves are not autocorrelated. But the higher approximation does poorly when it is used to predict the fourth moment of the distribution of observed returns at high values of the basic parameter.

The first order approximation to the kurtosis is shown, along with the simulated values, in Figure 3. The approximation does indicate an upward bias to the fourth moment, and does reasonably well at values of $\frac{\tau}{P\sigma}$ below one, it underestimates the bias at larger values and underestimates the effect of large standard deviations. These have a dominant effect for values of $\frac{\tau}{P\sigma}$ over two. This approximation also underestimates distortions in kurtosis, as shown in Figure 4.

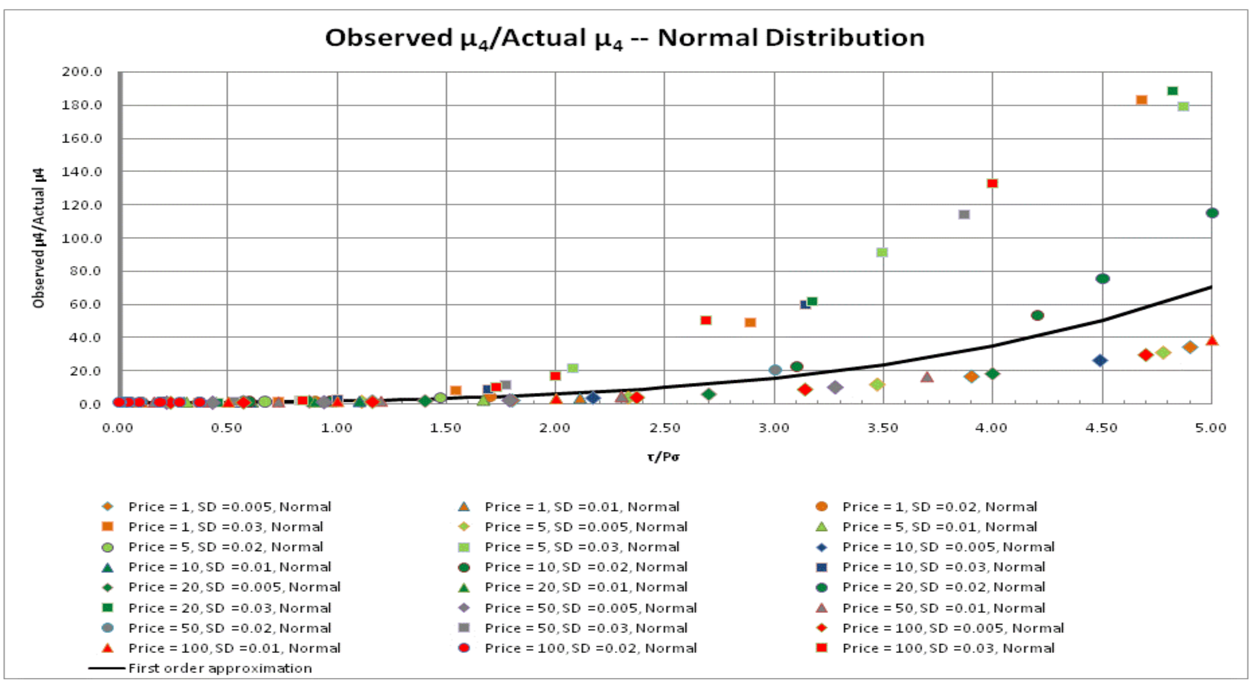

Figure 3a. Effect of tick size on the estimated fourth moment of the rate of return with normally distributed increments

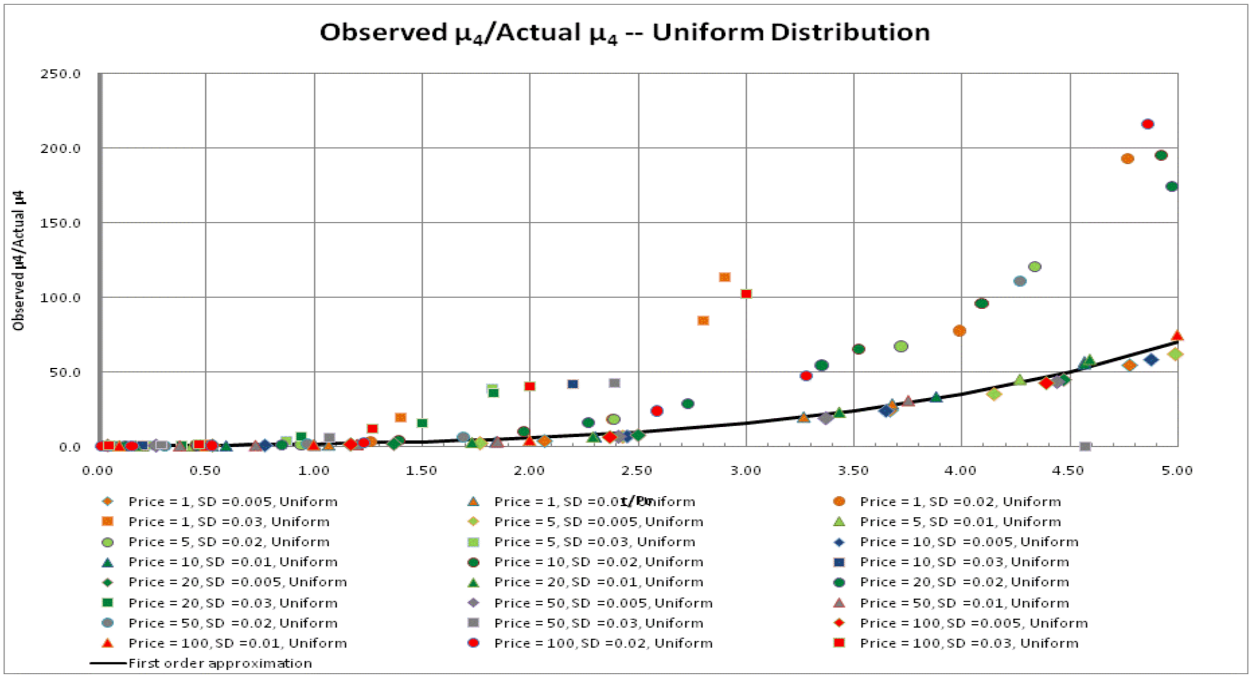

Figure 3b. Effect of tick size on the estimated fourth moment of the rate of return with uniformly distributed increments

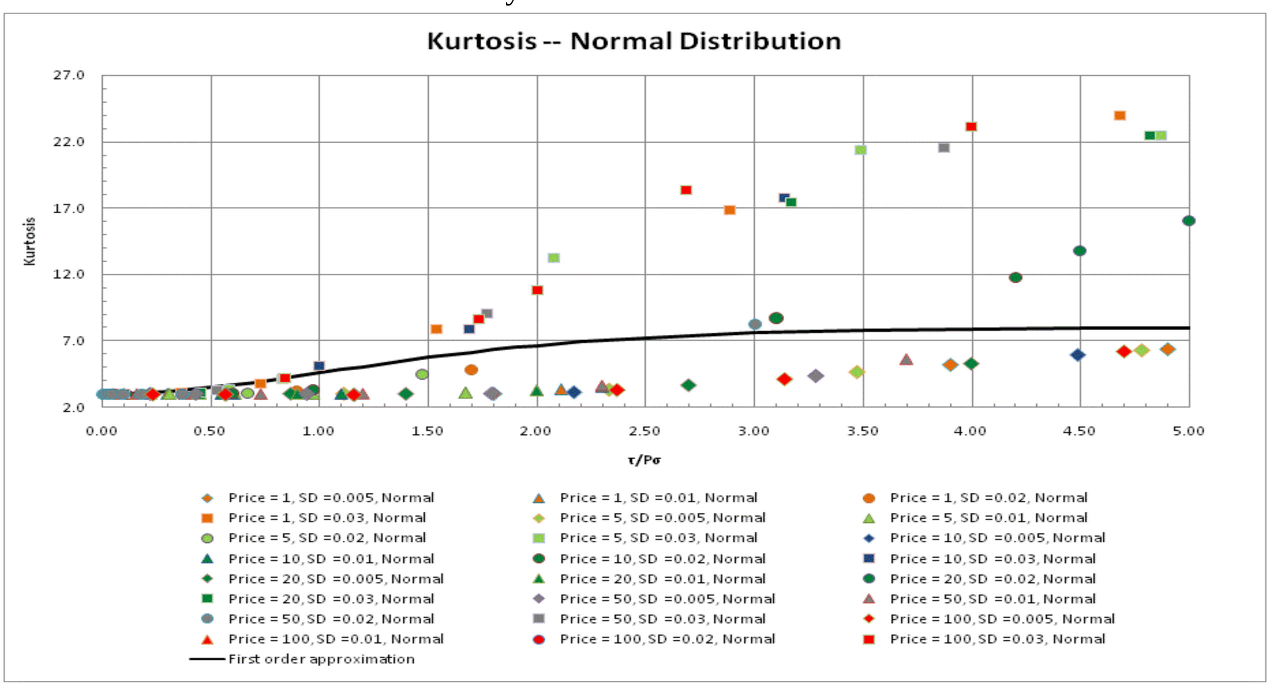

Figure 4a. Effect of tick size on the estimated fourth moment of the rate of return with normally distributed increments

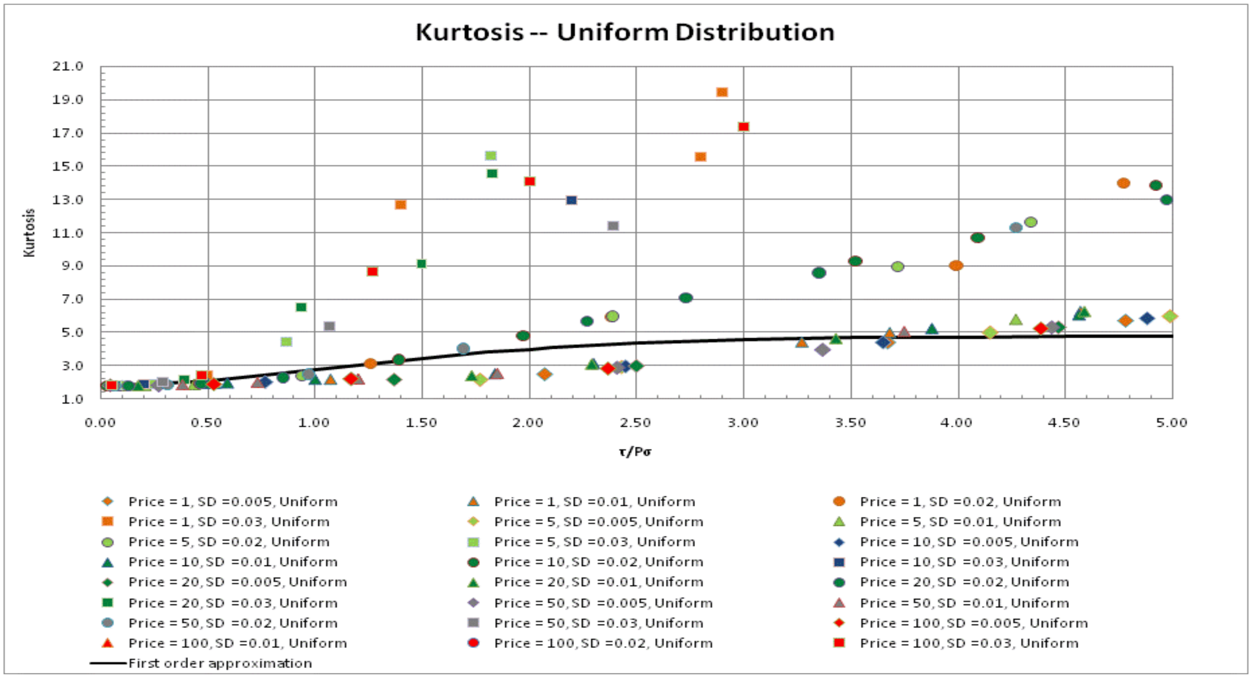

Figure 4b. Effect of tick size on the estimated kurtosis of the rate of return with normally distributed increments

If better approximations are needed they cannot be obtained merely by adding terms to the expansion. As explained in the appendix the major causes for this "failure to converge" is that two approximations made in deriving the first approximation break down as $\frac{\tau}{P\sigma}$ increases: the local distribution is no longer uniform and the error terms at *t* and *t*+1 are not uncorrelated.

This is not unexpected. Large values of $\frac{\tau}{P\sigma}$ imply that consecutive prices are very likely to fall within the same tick. If that is so, then as the parameter increases the local distribution should approach the distribution of the increments. Along with this, even if the overall distribution within a tick is uniform, the distribution of the error at time *t* will not be independent of the value of the error at time *t*-1. If the error at time *t*-1 is near zero the probability the error at time *t* will be close to zero is high. On the other hand if the error at time *t*-1 is near $-\frac{\tau}{P\sigma}$, then the probability that it will be close to $-\frac{\tau}{P\sigma}$ at time t is not much greater than one half, since any negative increment will bring the price down to the next lower tick and the error to the high end of the range. Thus a better description of these effects is needed before a better approximation can be obtained.

Practical significance

What I have shown above is mostly theory and simulation. It does, however, have practical significance at two levels. At the academic level it matters because it suggests that we are not entitled to draw conclusions about the variance and autocorrelation of increments in price (or in its log) by merely estimating these statistics for the reported prices or returns. At the market level it may matter because the pricing of options is intimately related to knowledge of the volatility, and I have shown that the volatility estimated from stock returns is biased upwards.

In fact, at the academic level this analysis raises the question of whether any of the econometric tests and estimation procedures that ignore rounding apply to analysis of real data. Given the transformations involved and the problems of using modular arithmetic to represent rounding this amounts to a major effort to reanalyze or analytical tools.

The ***

At the market level, the upward bias in the estimates of volatility would lead investors who base their decisions on such estimates to pay too much for both puts and calls. This would be to the advantage of institutions that write the options. It is also relevant, though less directly, to investors that use implied volatility in their decisions: the estimation of implied volatility requires knowledge of the price of the stock and of the option at the same time for recent trades. Both those prices are rounded to the nearest tick, and we have no definitive analysis of the biases that might be involved in that process. Preliminary explorations suggest that the biases involved are much lower.

Moreover, the option contract makes the settlement at the reported effective closing price, not at the underlying martingale price, which is unknown. The models, on the other hand, are based on the relation that should exist between underlying price of the stock and the underlying price of the option. Thus when expiration is very close (when the ratio of the stock's tick size to the price times the standard deviation of the increments between purchase and expiration tends to infinity) we might expect the market to focus its attention to the reported price rather than the underlying price. I would expect this to lead to "peculiar" behavior of option prices close to maturity. But this also implies that implied volatilities may have biases that need to be explored more carefully.

References

Ball, C., 1988, “Estimation Bias Induced by Discrete Security Prices,” *Journal of Finance*, vol. 43, pp. 841-865.

Campbell, J.Y, A.W. Lo, and A.C. MacKinlay, 1997, “The Econometrics of Financial Markets”, Princeton University Press, Princeton, NJ.

Cho, D. and E. Frees, 1988, “Estimating the Volatility of Discrete Stock Prices,” *Journal of Finance*, vol. 43, pp. 451-466.

Fama, E. F., 1965, “The behavior of stock market prices”, *Journal of Business*, vol. 38, 34-105.

Gottlieb, G. and A. Kalay, “Implications of the Discreteness of Observed Stock Prices,” *Journal of Finance*, vol. 40, pp. 135-154.

Greene, W.H., 2000, “Econometric analysis”, fourth edition, Prentice Hall, Upper Saddle River, NJ.

Gujarati, D., 1992, “Essential of Econometrics”, McGraw Hill, Inc., New York, NY.

Gujarati, D.N., 1995, “Basic Econometrics”, third edition, McGraw Hill, Inc., New York, NY.

Hamilton, J.D., 1994, “Time Series Analysis”, Princeton University Press, Princeton, NJ.

Harris, L. 1990, Estimation of Stock Variances and Serial Covariances from Discrete Observations,” *Journal of Financial and Quantitative Analysis*, vol. 25, pp. 291-306.

Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lütkepohl, and T.C. Lee, 1985, “The Theory and Practice of Econometrics”, second edition, John Wiley and sons, New York, NY.

Wooldridge, J.M., 2006, “Introductory Econometrics”, third edition, Thomson South-Western, Mason, OH.

