Estimation of Stock Price Variances and Serial Covariances from Discrete Observations

Lawrence Harris*

Abstract

Stock price discreteness adds noise to price series. The noise increases return variances and adds negative serial correlation to return series. Standard variance and serial covariance estimators therefore overestimate the variance and serial covariance of the underlying stock values. Discreteness-induced variance and serial covariance depend on underlying volatility and on the size of the bid/ask spread. Simple formulas for approximating the effects of discreteness on variance and serial correlation are derived and presented. The approximations, which are accurate in daily data, can be used to adjust the standard variance and serial covariance estimators.

I. Introduction

Variances and serial covariances are ubiquitous in the theory and practice of finance. Variance plays a central role in risk evaluation, option pricing, information flow identification, and hypothesis testing; serial covariances arise in transaction cost and market efficiency analyses. It is therefore crucial that these moments be accurately estimated.

Stock price discreteness complicates the estimation of these moments because discreteness causes observed prices to differ from underlying stock values. (Stock prices are discrete because exchange regulations require all prices to be expressed as a multiple of some minimum tick, usually $⅛.) Observed price-change variances and serial covariances may therefore poorly estimate the variances and serial covariances of underlying stock value innovations.

Accurate characterization of the underlying value process is important to those analysts who are primarily interested in long-term fluctuations in stock value. Phenomena associated with trading, such as discreteness, cause transitory fluctuations in observed prices. Transitory volatility, if mistaken for long-term volatility, will bias analyses.

* School of Business Administration, University of Southern California, Los Angeles, CA 90089–1421. The author thanks Stephen Brown, Andrew Christie, Larry Glosten, Andrew Lo, Marc Reinganum, *JFQA* Managing Editor Paul Malatesta, and seminar participants at Cornell University, The University of Michigan, New York University, Northwestern University, University of California, Los Angeles, University of Southern California, University of Utah, Vanderbilt University, University of Pennsylvania, and Yale University for their many insightful comments.

For example, consider variance measurement in an option pricing problem. For a pure diffusion process, the accuracy of the standard variance estimator can always be improved by more frequent sampling. This suggests that variance should be measured from lots of short interval data. But if the data are obtained through a discrete filter, accuracy will not always be improved by more frequent sampling. Since discreteness causes an upward bias in the standard variance estimator, more frequent sampling can actually decrease the reliability of the pricing study.

This paper considers estimation problems related to price discreteness. The analysis is based on the assumption that observed prices are obtained by rounding underlying values to the nearest eighth. Gottlieb and Kalay (1985) first used this assumption to show that discreteness increases observed price change variance relative to the underlying value innovation variance. Below, it is also shown that rounding induces negative serial covariance into the price-change series. The standard second-moment variance estimator therefore overstates the underlying value innovation variance, and serial covariance bid-ask spread estimators, such as Roll's (1984), overstate the actual spread.[1]

Discreteness-induced estimator biases can significantly affect statistical inference. Consider, for example, the effects of these biases for a $10 stock whose underlying values follow a random walk with innovations that have a daily standard deviation of 2 percent. Discreteness causes the ratio of a five-day observed price-change variance (expressed in daily units) to a one-day observed price-change variance to have an expected value of 0.951 for this stock even though the same ratio is one for the underlying process. Discreteness also causes the first order serial correlation coefficient of observed price-changes to have an expected value of -0.031 where zero would otherwise be expected. These biases can be readily detected in daily data. They are most serious for low-priced stocks (often small firms) and for frequently sampled price series (transaction data). In both cases, underlying value innovation variance is small relative to the minimum tick.

This analysis generalizes Gottlieb and Kalay's model to take into account the bid/ask spread. The resulting model is also a generalization of Roll's model to take into account discreteness. The new model yields quantitative implications for the biases in the standard variance and in the serial covariance estimators. Both biases depend on the underlying value innovation variance and on the bid/ask spread.

A likelihood method for estimating the new model is presented. Estimation is difficult because underlying stock value and the bid-ask classification of prices are both unobserved in most price series. The observability problem can be solved by integrating the unobserved variables out of the likelihood, using diffuse prior distributions. The parameters can then be estimated (at significant cost) using maximum likelihood methods.

Fortunately, simple additive adjustments can be made to the standard variance and serial covariance estimators to remove the discreteness biases. The ad-

[1] Roll's serial covariance spread estimator is $2\sqrt{-S\text{Cov}}$, where SCov is the serial covariance of price changes. Negative bias in the price change serial covariance due to discreteness causes Roll's estimator to be upward biased.

justments depend on limiting results that are quite accurate in daily data. These limiting results also show that discreteness-induced bias in the French and Roll (1984) estimator of true variance (sample variance plus two times the serial covariance) approaches zero in daily data.

Several perspectives should be noted before proceeding. The analysis is based on an ad hoc model of discreteness, not a behavioral model. Discrete prices are simply assumed to be obtained by rounding underlying values to the nearest ⅛. The actual price formation process probably is more complex. Marsh and Rosenfeld (1986) analyze discreteness under an alternative ad hoc assumption that prices are observed only when underlying value is equal to an acceptable discrete price level. Under this assumption, discreteness does not increase variance or affect serial correlation. If the price formation process were modeled as the outcome of negotiation among rational agents, both ad hoc assumptions would probably characterize the results to some extent.

This analysis considers the effect of discreteness on estimated serial covariances using a simple model of the bid-ask spread. This simple model assumes that the spread is composed only of a transaction cost component, and not also of an adverse selection component as suggested by Bagehot (1971), Copeland and Galai (1983), and Glosten and Milgrom (1985). A generalization that includes the adverse selection component, modeled as a positive function of trading volume (as suggested by Easley and O'Hara (1985)), is presented in Glosten and Harris (1988). That analysis, however, does not focus on the effects of discreteness on serial covariances.

This analysis does not consider the clustering of stock prices, which is related to price discreteness. Niederhoffer (1965), (1966) shows that stock prices are observed most often on integers, then on halves, on quarters, and least often on eighths. It is, however, less pronounced for low-priced than for high-priced stocks, as Ball, Torous, and Tschoegl (1985) predict. This suggests that the *de facto* minimum price change is greater than ⅛ for high-priced stocks. Discreteness, therefore, may be a more important phenomenon than it otherwise would appear to be.

The remainder of this paper is organized into three sections and two appendices. Section II introduces the discrete bid-ask model, describes its implications, and suggests simple estimators that adjust for discreteness-induced volatility and serial covariance. Section III presents simulation results that rank the relative efficiency of various estimators. The final section offers a short summary of the results. Appendix A describes some unique problems encountered when using maximum likelihood to estimate the discrete model, while Appendix B presents derivations of mathematical results used in Section II.

II. A Discrete Price Model

The discrete model assumes that unobserved dividend-adjusted stock values follow a random walk with drift. Observed prices are obtained by adding (for an ask price) or subtracting (for a bid price) one-half of the bid-ask spread to the underlying stock value and then rounding the result to the nearest eighth. The model is formally stated as

(1a) $V_{t+1} = V_t - D_t + \mu + \epsilon_t$ (Underlying value process),

(1b) $P_t^o = R\left(V_t + cQ_t, d\right)$ (Observed price process), and

(1c) $\left\{\epsilon_t\right\} \sim \text{iid } N\left(0,\sigma^2\right)$ (Underlying value innovations),

where V_t and P_t^o are the underlying value and observed prices at time t, D_t is the dividend paid at t, μ is a drift term, $R(.,.)$ is the rounding function that rounds V_t to the nearest tick d (usually \$⅛), σ^2 is the value innovation variance, c is one-half of the bid-ask spread, and Q_t is a generally unobserved $\{-1,1\}$ indicator for whether the observed price is a bid or ask price. The probability of a bid is assumed equal to that of an ask, and the series $\{Q_t\}$ is assumed to be serially independent and independent of the other variables. A maximum likelihood method for estimating this model is described in Appendix A.

Gottlieb and Kalay's model and Roll's model are special cases of the above model. When c is equal to zero (no bid-ask spread), a discrete-time analog of Gottlieb and Kalay's continuous time model is obtained. When d is zero (no rounding) and the normality assumption in (1c) is not specified, Roll's model is obtained.

Implications of the model for observed price-change variances and serial covariances are apparent when it is restated for the dividend-adjusted price-change

(2a)
$$\begin{aligned}\Delta P_{t+1}^o &= P_{t+1}^o - P_t^o + D_t \\ &= \mu + c\left(Q_{t+1} - Q_t\right) + \eta_{t+1} - \eta_t + \epsilon_t ,\end{aligned}$$

where

(2b)
$$\eta_t = V_t + cQ_t - R\left(V_t + cQ_t\right)$$

is the rounding error. The discrete rounding error process increases the observed price-change variance and contributes negative serial covariance to the price series beyond that caused by the bid-ask spread process.

The implied variance of observed prices is

(3)
$$E\left(\Delta P_t^o - \mu\right)^2 = \sigma^2 + 2c^2 + E\left(\eta_{t+1} - \eta_t\right)^2 .$$

The three terms in (3) are due to value innovations, bid-ask bounce, and rounding errors. Cross-products of the square do not appear in (3) because $\{Q_t\}$ and $\{\epsilon_t\}$ are independent and because cross-products involving η_t are equal to zero when t is large. The latter result, proven in Lemmas 3 and 3.A of Appendix B, occurs because the rounding errors depend with equal weight on all past values of e_t and Q_t.

A formula for numerically evaluating the rounding variance component is derived and presented in Lemmas 5 and 6 of Appendix B. The formula displays two symmetries in $2c$: the rounding variance component evaluated at $2c$ is equal to the rounding variance component evaluated at $2c$ plus or minus any integer multiple of the minimum price variation, d; and the rounding variance component evaluated at $2c = i \times d + k$ is equal to the rounding variance component evaluated at $2c = i \times d - k$, where i is any integer and k is any number. When c is zero, this formula is analogous to one derived by Gottlieb and Kalay ((1985), p. 143) who assume a continuous log-normal diffusion process for underlying value. The two results are numerically indistinguishable to five and more decimals for variances typically encountered in finance.

A tabulation of the general result for $d = \$1/8$ (first column of Table 1) shows that for all bid-ask spreads c, the rounding variance component approaches a limit of 26.04 (cents)^2 when the value innovation standard deviation is large (generally greater than 5 cents). Lemma 7 of Appendix B shows that this limit is $d^2/6$. Since all but the lowest priced stocks have daily value standard deviations of greater than 5 cents, this simple expression can be used to adjust the standard variance estimator to produce a more accurate estimator. The results also show that discreteness-induced variation is at a minimum when the underlying bid-ask spread is equal to zero (or to any multiple of the minimum price variation, d.)

The implied first order serial covariance of observed price-changes is

$$(4a)\quad E\left(\Delta P^o_{t+1} - \mu\right)\left(\Delta P^o_t - \mu\right) = -c^2 + E\left(\eta_{t+1} - \eta_t\right)\left(\eta_t - \eta_{t-1}\right),$$

and the implied higher order serial covariances are

$$(4b)\quad E\left(\Delta P^o_{t+r} - \mu\right)\left(\Delta P^o_t - \mu\right) = E\left(\eta_{t+r} - \eta_{t+r-1}\right)\left(\eta_t - \eta_{t-1}\right), \text{ for } r > 1.$$

The higher order serial covariances are not equal to zero because the rounding errors are not independently distributed. Formulas for evaluating the rounding error expectations are derived in Lemmas 5 and 6 of Appendix B. Tabulations (Table 1) show that the serial covariances are negative for all lags. For all bid-ask spreads, the first order serial covariance approaches a limit of -13.02 (cents)^2 when the value innovation standard deviation is large (again generally greater than 5 cents). The higher order serial covariances, which are uniformly small, approach a limit of zero when the value innovation standard deviation is large, or when the number of lags is large (Lemmas 5 and 7). Lemma 7 shows that the limit of the first serial covariance is $-d^2/12$, and confirms that the limit for the higher order serial covariances is zero. The tabulation shows that first order rounding serial covariances are greatest when the underlying bid-ask spread is equal to $d/2$ (or $d/2$ plus some integer multiples of d), and that the higher order serial covariances are greatest when the underlying bid-ask spread is equal to zero (or some multiple of d).

Roll's serial covariance spread estimator, $2\sqrt{-S\text{Cov}}$, where SCov is the serial covariance of price changes, is upward biased by discreteness-induced negative serial covariation. The absolute bias is greatest for low-priced stocks because they have the smallest underlying return standard deviations. For example, the

TABLE 1

Rounding Error First Difference Variance and Serial Covariances by Underlying Value Innovation Standard Deviation, σ, and Total Spread, 2*c*

σ (cents)	$E(\eta_{t+r}-\eta_{t+r-1})(\eta_t-\eta_{t-1})$, in cents² — Lag *r*: 0 (variance)	1	2	3	4	5
		$2c = i/8,\ i = 1, 2, 3, \ldots$				
0.1	0.99	−0.29	−0.05	−0.02	−0.02	−0.01
0.2	1.95	−0.58	−0.10	−0.05	−0.03	−0.02
0.5	4.74	−1.46	−0.24	−0.12	−0.08	−0.06
1	8.97	−2.92	−0.48	−0.25	−0.16	−0.11
2	15.95	−5.84	−0.96	−0.48	−0.28	−0.17
5	25.37	−12.36	−0.31	−0.01	−0.00	−0.00
∞	26.04	−13.02	−0.00	−0.00	−0.00	−0.00
		$2c = i/8 \pm 1/64,\ i = 1, 2, 3, \ldots$				
0.1	9.03	−4.42	−0.02	−0.01	−0.01	−0.01
0.2	9.50	−4.56	−0.05	−0.02	−0.02	−0.01
0.5	10.79	−4.98	−0.08	−0.04	−0.02	−0.02
1	12.85	−5.45	−0.20	−0.14	−0.10	−0.08
2	17.62	−7.08	−0.74	−0.40	−0.24	−0.14
5	25.47	−12.46	−0.26	−0.01	−0.00	−0.00
∞	26.04	−13.02	−0.00	−0.00	−0.00	−0.00
		$2c = i/8 \pm 1/32,\ i = 1, 2, 3, \ldots$				
0.1	15.14	−7.47	−0.02	−0.01	−0.01	−0.01
0.2	15.61	−7.62	−0.05	−0.02	−0.02	−0.01
0.5	16.89	−8.05	−0.12	−0.06	−0.04	−0.02
1	18.64	−8.75	−0.17	−0.07	−0.05	−0.04
2	21.26	−9.68	−0.38	−0.23	−0.14	−0.08
5	25.71	−12.69	−0.15	−0.01	−0.00	−0.00
∞	26.04	−13.02	−0.00	−0.00	−0.00	−0.00
		$2c = i/8 \pm 3/64,\ i = 1, 2, 3, \ldots$				
0.1	18.80	−9.30	−0.02	−0.01	−0.01	−0.01
0.2	19.27	−9.45	−0.05	−0.02	−0.02	−0.01
0.5	20.55	−9.89	−0.12	−0.06	−0.04	−0.03
1	22.30	−10.61	−0.23	−0.11	−0.06	−0.03
2	24.36	−11.78	−0.22	−0.08	−0.04	−0.02
5	25.94	−12.92	−0.05	−0.00	−0.00	−0.00
∞	26.04	−13.02	−0.00	−0.00	−0.00	−0.00
		$2c = i/8 \pm 1/16,\ i = 1, 2, 3, \ldots$				
0.1	20.02	−9.91	−0.02	−0.01	−0.01	−0.01
0.2	20.49	−10.06	−0.05	−0.02	−0.02	−0.01
0.5	21.77	−10.50	−0.12	−0.06	−0.04	−0.03
1	23.52	−11.23	−0.24	−0.12	−0.07	−0.04
2	25.52	−12.53	−0.20	−0.03	−0.00	−0.00
5	26.04	−13.02	−0.00	−0.00	−0.00	−0.00
∞	26.04	−13.02	−0.00	−0.00	−0.00	−0.00

Note: The expectations are evaluated using Lemmas 5 and 6 of Appendix B assuming that the minimum tick, *d*, is equal to $⅛ and the underlying value innovation mean, ϕ, is equal to zero. The results display two symmetries in $2c$: the expectations evaluated at $2c$ are equal to the expectations evaluated at $2c$ plus or minus any integer multiple of the minimum price variation, *d*, and, the expectations evaluated at $2c = i \times d + k$ are equal to the expectations evaluated at $2c = i \times d - k$, where *i* is any integer and *k* is any number.

bias can be shown (using results from Table 1) to be 15 percent of the true absolute spread for a \$3 stock with an underlying return standard deviation of 2 percent and a \$⅛ total spread. For a \$30 stock with the same return standard deviation and \$¼ spread, the bias is only 4 percent. Since small firm stocks tend to be low-priced, the negative cross-sectional correlation between estimated bid-ask spreads and firm size that Roll observes may be at least partly spurious.

The French and Roll (1984) method-of-moments estimator of variance is equal to the sample variance plus two times the sample first order serial covariance. The expectation of this estimator can be obtained by manipulating (3) and (4a). It is

$$E\left(\operatorname{Var}\left(\Delta P_t\right) + 2S\operatorname{Cov}\left(\Delta P_t\right)\right) = \sigma^2 - E\left(\eta_{t+1} - \eta_t\right)^2 - 2E\left(\eta_{t+1} - \eta_t\right)\left(\eta_t - \eta_{t-1}\right). \tag{5}$$

Substituting limiting values for the expectations causes the last two terms to cancel. In addition to removing variation due to the bid-ask spread, the French and Roll estimator also removes variation due to price discreteness when σ is large. When σ is small, discreteness-induced error in this estimator is still small because the two expectations tend to cancel.

III. Estimator Efficiency

A simple simulation study demonstrates that the volatility and bid-ask spread estimators that take discreteness into account are more accurate than those that do not. Four price volatility estimators and three bid-ask spread estimators are examined. The volatility estimators include the sample standard deviation of observed price-changes, the adjusted sample standard deviation obtained by subtracting $d^2/6$ from the sample variance, the French and Roll estimator, and the maximum likelihood estimator described in Appendix A. The bid-ask spread estimators include Roll's estimator, an adjusted serial covariance estimator computed by subtracting $-d^2/12$ from the sample serial covariance before applying Roll's formula, and the maximum likelihood estimator.

The simulated sample consists of 168 price series of 252 observations each (one year of trading days) generated using the discrete model ($1a-c$). The parameters used to create the simulated series are maximum likelihood estimates of σ and c obtained from actual daily stock price time series for 1983. Each simulated series was initialized at the first stock price for the corresponding stock in the actual sample.

A stock qualified for inclusion in the actual sample if it was continuously listed on the NYSE or AMEX in 1983, if it traded every market day in 1983, if it only traded on eighths, and if shareholders received no distributions other than normal taxable cash dividends. Approximately 1694 stocks meet these criteria. Due to the high computational cost of the maximum likelihood analyses, only a subsample of these qualifying stocks was further considered for inclusion in the actual sample. This subsample consists of all 208 securities for which the predicted bias in the sample standard deviation estimator of the underlying value innovation standard deviation is greater than 1 percent. In addition, 46 other randomly chosen qualifying securities are included so that estimation properties might be explored for securities for which the predicted bias is inconsequential.[2] The final sample of 168 actual stocks is obtained by eliminating all subsample stocks for which the sample serial covariance estimate is not positive.[3] The data are obtained from the CRSP Daily Stock Master File.

The various estimators are ranked by cross-sectional mean-squared error after dividing by price level to control heteroskedasticity (Table 2). The maximum likelihood estimators are the best performing estimators in their respective classes. They have lower mean-squared errors and smaller mean biases than all of the other estimators. Their good performance is probably due to their greater use of prior information about the discreteness and the bid-ask spread processes.

Among the variance estimators, the method-of-moments estimator is next most efficient, followed by the adjusted estimator and finally by the sample variance. Since the latter two estimators do not take into account the bid-ask spread, they overestimate underlying value innovation standard deviations. The adjusted variance is less biased than the sample variance because it adjusts for discreteness.

Among the spread estimators, the adjusted-Roll estimator performs better than the unadjusted estimator, as expected. The latter is significantly upward biased by discreteness-induced negative serial correlation.

TABLE 2

Estimator Efficiency in Simulated Data

Estimator	Root Mean-Squared Error	Bias	t: Bias = 0	Mean-Adjusted Root Mean-Squared Error
Underlying Value Innovation Variance Estimators of σ				
Maximum Likelihood	0.185%	−0.025%	−1.75	0.184%
French-Roll	0.257	−0.054	−2.78	0.251
Adjusted Sample Variance	0.369	0.153	5.88	0.336
Sample Variance	0.513	0.266	7.83	0.439
Bid-Ask Spread Estimators of c				
Maximum Likelihood	0.358%	0.011%	0.43	0.358%
Adjusted Serial Covariance	0 406	0.039	1.25	0.404
Roll Serial Covariance	0.520	0.211	5.73	0.475

Note: The simulated data consist of 168 price series of 252 observations each generated by the discrete model of Section II, with parameters given by maximum likelihood estimates obtained from 168 actual low-priced daily price series. The volatility estimates are expressed as standard deviations as a percentage of price. The spread estimates are expressed as a percentage of price.

IV. Summary

The discreteness of stock prices increases price-change variances and adds

[2] The number 46 is a consequence of the process used to choose the additional securities. Every twentieth security in the set of qualifying securities is included, regardless of the predicted variance bias.

[3] The maximum likelihood estimate of c is zero for 24 of the 168 actual price series with negative serial covariances. As pointed out in Section II, discreteness causes negative serial covariance in observed price series even when c is zero.

negative serial covariance to price-change series. This paper shows how these effects are related to the variance of underlying value changes and to the bid-ask spread. Simple limiting expressions for the discreteness-induced variance and serial covariance are derived. They are equal to $d^2/6$ and $-d^2/12$, respectively, where d is the minimum discrete price change. These limits, which are generally attained in daily data, can be used to adjust standard variance and serial covariance estimators to produce more accurate estimators.

The seriousness of the discreteness problem should be kept in perspective. Observed price variance due to the rounding process (ignoring stock price clustering) is small compared to underlying value innovation variance and bid-ask bounce in daily data for all but the lowest price and lowest variance stocks. For most purposes, discreteness in these data can be ignored, although researchers should always be aware of its possible influence. However, when data are measured at short intervals (such as transaction intervals), when studies are done of stocks that happen by design or by accident to be low-priced stocks (such as small firm studies), when sample sizes are large so that small variance and serial covariance components can be readily identified (such as pre/post split volatility studies and variance ratio studies), discreteness must be considered as a source of variation and serial covariation.

Appendix A: Maximum Likelihood Estimation

Estimation of the discrete bid-ask model is difficult because underlying stock values are not observed and because most data sets do not classify observed prices as bids and asks. The estimation procedure described here solves the unobserved variables problem by integrating them out of the conditional likelihood function. The resulting unconditional likelihood is then maximized to obtain point estimates of μ, σ^2, and c. The integration is carried out over independent uniform priors for the roundoff errors, $\{\eta_t\}$,[4] and over the independent Bernoulli distributions assumed for the bid-ask indicators, $\{Q_t\}$. These priors make the integrated likelihood function an equal weighted average of the conditional likelihood function evaluated over all stock value paths and bid-ask classifications that map (via the rounding function) into the observed price path.

Since the time-series model is stationary and ergodic, the maximum likelihood estimator will be consistent and asymptotically normally distributed. (Gordin (1969) provides a proof of the central limit theory for stationary ergodic processes with dependent observations.) Ergodicity means that the dependence between two observations declines to zero as the interval between them increases. When a process is ergodic, the information in a sample about the process parameters increases without bound when the sample size increases. Ergodicity for this process is verified by noting that its autocorrelations approach zero as the lag gets large (Lemmas 5 and 7).

[4] The uniform prior is a useful way of characterizing prior ignorance. (See Zellner (1971) or Jeffreys (1966).) Gottlieb and Kalay also show that, given any initial distribution for the rounding error at time 0, the distribution of the error at time t approaches the uniform distribution as t becomes large.

Exact analytic computation of the integrated likelihood function is impossible because it involves a $T+1$-fold summation over a $T+1$-fold integral. The summation is over the $T+1$ bid-ask indicators $\{Q_t\}$, which take discrete $(-1,1)$ values with equal probability. The integral is over the $T+1$ roundoff errors, $\{\eta_t\}$, which take continuous values over $[-d/2, d/2]$ with uniform density $1/d$. The function is

$$\sum_{R_0}\sum_{R_1}\cdots\sum_{R_T}(1/2)^{T+1}\int_{\Theta_0}\int_{\Theta_1}\cdots\int_{\Theta_T}(1/d)^{T+1}$$
$$\times\ \mathscr{L}_{1,0}\ \mathscr{L}_{2,1}\cdots\ \mathscr{L}_{T,T-1}\,d\eta_0\,d\eta_1\ldots d\eta_T\ ,$$

where R_t is the discrete range of Q_t, H_t is the continuous range of η_t, and $\mathscr{L}_{t+1,t}$ is the normal probability density of ΔP^{o}_{t+1}, given η_{t+1}, η_t, Q_{t+1}, and Q_t, which is implied by (2a) and (1c). Numeric evaluation of the multiple integral can be achieved by approximating Θ_t by a discrete subset of its continuous range. This study uses a lattice of m equally spaced points: $\{d/m(k-(m+1)/m) \mid k = 1, 2, \ldots, m\}$. This approximation reduces the likelihood function to a $T+1$-fold summation over the prior distributions of the $T+1$ state variable vectors (η_t, Q_t), each of which appears in the state-conditional likelihoods of two adjacent observations. The summation can be evaluated recursively using a formula derived in the next paragraph. In practice, the model is accurately estimated for m greater than 5. This study uses 15 points.[5]

A recursion for evaluating the likelihood function is apparent when the $T+1$-fold summation is reordered. The unconditional likelihood function, reexpressed in a simplified notation, is

$$\mathscr{L} = \sum_{0}\sum_{1}\cdots\sum_{T}P_0P_1\ldots P_T\,\mathscr{L}_{1,0}\,\mathscr{L}_{2,1}\cdots\mathscr{L}_{T,T-1}\ ,$$

where the summations are understood to be over all possible values of the state variables, η_t and Q_t, and where the prior probability density function of the state variable vector is represented by P_t. This can be reordered as

$$\mathscr{L} = \sum_{T}P_T\Bigg[\sum_{T-1}P_{T-1}\,\mathscr{L}_{T,T-1}\Bigg(\cdots$$
$$\Bigg\{\sum_{2}P_2\,\mathscr{L}_{3,2}\Bigg[\sum_{1}P_1\,\mathscr{L}_{2,1}\Bigg(\sum_{0}P_0\,\mathscr{L}_{1,0}\Bigg)\Bigg]\Bigg\}\cdots\Bigg]\ .$$

This expression can be evaluated recursively by computing summation 0 over all values of the state variables at $t = 0$ for each value of the state variables at $t = 1$, then summation 1 (which depends on the results of summation 1) over all values of the state variables at $t = 1$ for each value of the state variables at $t = 2$ and so on, until summation T is computed over all values of the state variables at $t = T$.

[5] Parameter estimates change considerably from $m = 1$ (discreteness ignored) to $m = 2$, and from $m = 2$ to $m = 3$. After $m = 4$, they quickly converge. The estimates for $m = 5$ and $m = 15$ are equal to greater than four significant decimals.

The recursion is best expressed in matrix notation. To do so, let the $k \times k$ square matrix $\underline{L}_t$ have rows that correspond to each of the k possible values of the current state vector, columns that correspond to each of the k possible values of the lagged state vector, and (i,j) elements that are equal to the conditional likelihood function at time t evaluated at the ith current state value and jth lagged state value. Likewise, let $\underline{D}_t$ be a diagonal matrix whose diagonal is given by the vector of probabilities associated with each possible state vector. The recursion can then be written as

$$\begin{aligned} \mathscr{L} &= \underline{\iota}' \underline{D}_T \underline{A}_T \\ \underline{A}_t &= \underline{L}_t \underline{D}_{t-1} \underline{A}_{t-1} \\ \underline{A}_0 &= \underline{L}_1 \underline{D}_0 \underline{\iota} , \end{aligned}$$

where $\underline{\iota}$ is $k \times 1$ the unit vector.

Appendix B: Mathematical Results

This appendix derives the mathematical results used in the text. For convenience, the discrete bid-ask model is restated as

$$\begin{aligned} \Delta P^o_{t+1} &= P^o_{t+1} - P^o_t + D_t & &\text{(Observed price change)} , \\ &= \mu + c\left(Q_{t+1} - Q_t\right) + \eta_{t+1} - \eta_t + \epsilon_t & & \\ \eta_t &= V_t + cQ_t - R\left(V_t + cQ_t\right) & &\text{(Rounding error process)} , \\ \epsilon_t &\sim \text{iid}\, N\left(0,\sigma^2\right) & &\text{(Underlying value innovations), and} \\ Q_t &= \begin{bmatrix} 1 \text{ with probability } 1/2 \\ -1 \text{ with probability } 1/2 \end{bmatrix} & &\text{(Bid-ask indicator)} , \end{aligned}$$

where D_t is the dividend paid in period t, c is one-half of the bid-ask spread, μ is a drift parameter, and $R(.,.)$ is the rounding function that rounds its argument to the nearest tick, d (usually d is ⅛). The bid-ask indicator, Q_t, is assumed to be independent of the other variables and serially independent.

Lemma 1. A recursion for η_t.

$$\begin{aligned} \eta_t = {} & \eta_{t-1} + cQ_t - cQ_{t-1} + \mu + \epsilon_{t-1} \\ & - R\left(\eta_{t-1} + cQ_t - cQ_{t-1} + \mu + \epsilon_{t-1}\right) . \end{aligned}$$

Proof.

$$\begin{aligned}
\eta_t &= V_t + cQ_t - R\left(V_t + cQ_t\right) \\
&= V_{t-1} + \mu + \epsilon_{t-1} + cQ_t - R\left(V_{t-1} + \mu + \epsilon_{t-1} + cQ_t\right) \\
&= V_{t-1} + \mu + \epsilon_{t-1} + cQ_t + cQ_{t-1} - cQ_{t-1} - R\left(V_{t-1} + cQ_{t-1}\right) \\
&\qquad - R\left(V_{t-1} + \mu + \epsilon_{t-1} + cQ_t + cQ_{t-1} - R\left(V_{t-1} + cQ_{t-1}\right)\right) \\
&\qquad\qquad \text{since } R(x + R(y)) = R(x) + R(y) \text{ for all } x,y \\
&= \eta_{t-1} + cQ_t - cQ_{t-1} + \mu + \epsilon_{t-1} \\
&\qquad - R\left(\eta_{t-1} + cQ_t - cQ_{t-1} + \mu + \epsilon_{t-1}\right) .
\end{aligned}$$

Corollary 1.A. An extended recursion for η_t.

$$\begin{aligned}
\eta_t &= \eta_s + cQ_t - cQ_s + (t-s)\mu + \sum_{i=s}^{t-1} \epsilon_i \\
&\quad - R\left(\eta_s + cQ_t - cQ_s + (t-s)\mu + \sum_{i=s}^{t-1} \epsilon_i\right) ,
\end{aligned}$$

for all $t > s$.

Proof. Repeated application of Lemma 1.

Lemma 2. A useful expectation.

$$E(k - R(\eta + k)) = 0 ,$$

where η is a random variable uniformly distributed on $[-d/2, d/2]$, and k is a constant.

Proof. Let $x = k - R(k)$.

$$\begin{aligned}
E(k - R(\eta + k)) &= E(x - R(\eta + x)) \\
&= x - ER(\eta + x) .
\end{aligned}$$

$$\text{For } x > 0, R(\eta + x) = \begin{bmatrix} d & \text{if } \eta + x > d/2 \\ 0 & \text{otherwise,} \end{bmatrix}$$

$$\text{so that } E(R(\eta + x) | x > 0) = \int_{d/2-x}^{d/2} 1/d\, d\, d\eta + \int_{-d/2}^{d/2-x} 1/d\ 0\, d\eta = x .$$

$$\text{Likewise, } E(R(\eta + x) | x < 0) = x ,$$

$$\text{so that } ER(\eta + x) = x$$

$$\text{and } E(x - R(\eta + x)) = x - x = 0 .$$

Lemma 3. Cross product expectations.

$$Ee_s \eta_t = 0 \text{ for all } s \text{ and } t .$$

Proof. The proof for $s \geq t$ is trivial since the series $\{\epsilon_s\}$ is serially independent.

For $s < t$, at least two different proofs can be made. The first proof assumes that in the distant history of the process, some η_o is uniformly distributed on its range $[-d/2, d/2]$. (This is the Gottlieb and Kalay uniform assumption.) Apply the law of iterated expectations to get

$$Ee_s\eta_t = E_Q E_\epsilon E_{\eta_o}\left(\epsilon_s \eta_t\right) = E_Q E_\epsilon\left(\epsilon_s E_{\eta_o}\left(\eta_t \mid \{\epsilon_o \ldots \epsilon_s\}, \{Q_o \ldots Q_t\}\right)\right),$$

where E_Q is the expectation operator over the joint distribution of $\{Q_o \ldots Q_t\}$, E_ϵ is the expectation operator over the joint distribution of $\{e_o \ldots \epsilon_s\}$, and E_{η_o} is the expectation operator over the uniform distribution of η_o. Expanding η_t using Corollary 1.A yields

$$Ee_s\eta_t = E_Q E_\epsilon\left(\epsilon_s E_{\eta_o}\left(\eta_o + G - R\left(\eta_o + G\right)\right)\right),$$

$$\text{where } G = cQ_t - cQ_o + \left(t - t_o\right)u + \sum_{i=t_o}^{t-1} \epsilon_i .$$

The proof is completed by applying Lemma 2 to the innermost expectation and by noting that $E\eta_o = 0$.

A second proof of this proposition can be made without assuming that η_o is uniformly distributed. Since a similar proof can be found in Gottlieb and Kalay, only a quick description of the method is presented here.

Let $H = G - \epsilon_s + \eta_o$, where G is given above and let $J = G - R(G)$. The limiting distribution of J is uniform on $[-d/2, d/2]$ as the history of the series becomes infinitely long. $E\epsilon_s\eta_t$ can be expressed in terms of J. Using the law of iterated expectations and applying Lemma 2 to J, the result is obtained.

Lemma 3.A. A similar result for Q_s.

$$EQ_s\eta_t = 0 \text{ for all } s \text{ and } t .$$

Proof. Same as for Lemma 3.

Lemma 4. Rounding error serial cross-products.

$$E\eta_t\eta_s = d^2/12 - E\left(\eta_s R\left(\eta_s + k_{ts}\right)\right)$$
$$= d^2/12 - E(\eta R(\eta + K(t-s))) ,$$

$$\text{where } k_{ts} = cQ_t - cQ_s + (t-s)\mu + \sum_{i=s}^{t-1} \epsilon_i, \text{ for } t > s ,$$

η is a random variable uniformly distributed on $[-d/2, d/2]$,

$K(t-s) = c\Delta Q + (t-s)\mu + e(t-s)$,

$$\Delta Q = \begin{cases} 2 \text{ with probability } 1/4 \\ 0 \text{ with probability } 1/2 \\ -2 \text{ with probability } 1/4, \end{cases}$$

and $e(t-s) \sim N\left(0, (t-s)\sigma^2\right)$.

Proof.

$$\begin{aligned} E\eta_t\eta_s &= E\eta_s^2 + E\eta_s k_{ts} - E\left(\eta_s R\left(\eta_s + k_{ts}\right)\right) && \text{by Lemma 1} \\ &= \int_{-d/2}^{d/2} (1/d)\eta_s^2 d\eta_s - E\left(\eta_s R\left(\eta_s + k_{ts}\right)\right) && \text{by Lemmas 3 and 3.A} \\ &= d^2/12 - E(\eta R(\eta + K(t-s))) \ , \end{aligned}$$

since η_s is distributed as η, $Q_t - Q_s$ is distributed as ΔQ because $\{Q_t\}$ are iid, and $\Sigma_{i=s}^{t-1}\epsilon_i$ is distributed as $e(t-s)$.

Lemma 5. Rounding error first difference variance and serial covariances.

$$\begin{aligned} E\left(\eta_{t+1} - \eta_t\right)^2 &= 2E(\eta R(\eta + K(1))) \\ E\left(\eta_{t+1} - \eta_t\right)\left(\eta_t - \eta_{t-1}\right) &= -2E(\eta R(\eta + K(1))) \\ &\quad + E(\eta R(\eta + K(2))) \\ E\left(\eta_{t+r} - \eta_{t+r-1}\right)\left(\eta_t - \eta_{t-1}\right) &= -2E(\eta R(\eta + K(r))) \\ &\quad + E(\eta R(\eta + K(r+1))) \\ &\quad + E(\eta R(\eta + K(r-1))) \quad \text{for } r > 1 \ , \end{aligned}$$

with all notation as defined in Lemma 4.

Proof. Expand each product and then apply Lemma 4.

Lemma 6. An expression for evaluating $E(\eta R(\eta + K(r)))$ found in Lemma 5.

$$E(\eta R(\eta + K(r))) =$$

$$\sum_{R_{\Delta Q}} \frac{1}{4} \sum_{j=-\infty}^{\infty} \int_{jd - r\mu - c\Delta Q - d/2}^{jd - r\mu - c\Delta Q + d/2} \left(d/2 \,|K(r) - jd| - 1/2(K(r) - jd)^2\right) \times \frac{1}{\sqrt{2\pi r\sigma^2}} \exp\left\{-\frac{e^2}{2r\sigma^2}\right\} de \ ,$$

where $R_{\Delta Q}$ is $\{-2,0,0,2\}$, the set of possible outcomes of ΔQ,
$K(r) = c\Delta Q + r\mu + e(r)$,
and $e(r) \sim N(0, r\sigma^2)$.

Proof. First take the expectation with respect to η conditional on $K(r)$.

$$\begin{aligned} E(\eta R(\eta + K(r))) &= E\eta(R(K(r)) + R(\eta + K(r) - R(K(r)))) \\ &= 0 + E(\eta R(\eta + x)), \end{aligned}$$

where $x = K(r) - R(K(r))$ since $E\eta = 0$.

$$\text{For } x > 0,\ R(\eta + x) = \begin{cases} d & \text{if } \eta + x > d/2 \\ 0 & \text{otherwise,} \end{cases}$$

$$\begin{aligned} \text{so that } E(\eta R(\eta + x) \mid x > 0) &= \int_{d/2 - x}^{d/2} 1/d \ d\eta \ d\eta + \int_{-d/2}^{d/2 - x} 1/d \ 0 \ \eta \ d\eta \\ &= d/2 \ x - x^2/2 \ . \end{aligned}$$

Likewise, $E(\eta R(\eta+x) \mid x<0) = d/2\, x - x^2/2$,
so that $E(\eta R(\eta+x)) = d/2\,|x| - x^2/2$.

The result is obtained by taking the expectation of this expression with respect to ΔQ and $e(r)$. The expectation is taken over the distribution of ΔQ whose joint density function is ¼. The index j counts the discrete levels that $R(K(r))$ may assume. Those levels are given by jd so that $x = K(r)-jd$. There is one integral for each of the discrete levels. The bounds of integration are the end points of the discrete interval. The integral is over the normal distribution of $e(r)$.

There are two symmetries in this formula. The expectation evaluated at $2c$ is equal to the expectation evaluated at $2c$ plus or minus any integer multiple of the minimum price variation, d, and the expectation evaluated at $2c = i\times d+k$ is equal to the expectation evaluated at $2c = i\times d-k$, where i is any integer and k is any number. These symmetries result because c appears in the integral only when $\Delta Q = -2$ or 2, and because everywhere that $c\Delta Q$ appears in the integral, jd is subtracted from it.

Lemma 7. Limiting values for $E(\eta R(\eta+K(r)))$.

$$\lim_{\sigma\to\infty} E(\eta R(\eta+K(r))) = d^2/12\,,$$

$$\text{and}\ \lim_{r\to\infty} E(\eta R(\eta+K(r))) = d^2/12\,.$$

Proof. Transform the index of the integral in Lemma 6 from ϵ to

$$y = \epsilon + r\mu + c\Delta Q - jd\,,$$

so that $E(\eta R(\eta+K(r))) =$

$$\sum_{R_{\Delta Q}} 1/4 \sum_{j=-\infty}^{\infty} \int_{-d/2}^{d/2} \left(d/2\,|y| - y^2/2\right) \frac{1}{\sqrt{2\pi r\sigma^2}} \exp\left\{-\frac{1}{2r\sigma^2}(y - r\mu - c\Delta Q + jd)^2\right\} dy.$$

Exchange the order of the infinite summation with that of the integral and multiply by d/d,

$$\sum_{R_{\Delta Q}} 1/4 \int_{-d/2}^{d/2} \left(d/2\,|y| - y^2/2\right) 1/d \sum_{j=-\infty}^{\infty} \frac{d}{\sqrt{2\pi r\sigma^2}}$$

$$\exp\left\{-\frac{1}{2r\sigma^2}(y - r\mu - c\Delta Q + jd)^2\right\} dy\,.$$

The infinite summation is an expression for the numeric evaluation of the integral of a normal density function with mean $(r\mu + c\Delta Q - y)/d$ and variance $r\sigma^2/d^2$ over its entire range (method of rectangles). The larger σ or r, the better this numeric approximation. The limit of this summation as σ or r goes to infinity is therefore equal to one. Thus the expression reduces to

$$\sum_{R_{\Delta Q}} 1/4 \int_{-d/2}^{d/2} \left(d/2\,|y| - y^2/2\right) 1/d\; dy = d^2/12\,.$$

References

Bagehot, W. "The Only Game in Town." *Financial Analysts Journal,* 27 (March/April 1971), 12–14.

Ball, C. A.; W. A. Torous; and A. E. Tschoegl. "The Degree of Price Resolution: The Case of the Gold Market." *The Journal of Futures Markets,* 5 (Spring 1985), 29–43.

Copeland, T. E., and D. Galai. "Information Effects on the Bid-Ask Spread." *Journal of Finance,* 38 (Dec. 1983), 1457–1469.

Easley, D., and M. O'Hara. "Prices, Trade Size and Information in Securities Markets." *Journal of Financial Economics,* 19 (Sept. 1987), 69–90.

French, K. R., and R. Roll. "Stock Return Variances: The Arrival of Information and Reaction of Traders." *Journal of Financial Economics,* 17 (Sept. 1986), 5–26.

Glosten, L. R., and L. Harris. "Estimating the Components of the Bid/Ask Spread." *Journal of Financial Economics,* 21 (May 1988), 123–142.

Glosten, L. R., and P. R. Milgrom. "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." *Journal of Financial Economics,* 14 (March 1985), 71–100.

Gordin, M. I. "The Central Limit Theorem for Stationary Processes." *Soviet Mathematics—Doklady,* 10 (Sept.–Oct. 1969), 1174–1176.

Gottlieb, G, and A. Kalay. "Implications of the Discreteness of Observed Stock Prices." *Journal of Finance,* 40 (March 1985), 135–153.

Jeffreys, H. *Theory of Probability* (3rd ed.). Oxford: Clarendon (1966).

Marsh, T. A., and E. R. Rosenfeld. "Nontrading, Market Making, and Estimates of Stock Price Volatility." *Journal of Financial Economics,* 15 (March 1986), 359–372.

Niederhoffer, V. "Clustering of Stock Prices." *Operations Research,* 13 (April 1965), 258–265.

__________________. "A New Look at Clustering of Stock Prices." *Operations Research,* 14 (April 1966), 309–313.

Roll, R. "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market." *Journal of Finance,* 39 (Sept. 1984), 1127–1139.

Zellner, A. *An Introduction to Bayesian Inference in Econometrics.* New York: John Wiley & Sons (1971).