Order Arrival, Quote Behavior, and the Return-Generating Process
Author(s): Joel Hasbrouck and Thomas S. Y. Ho
Source: *The Journal of Finance*, Vol. 42, No. 4 (Sep., 1987), pp. 1035-1048
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/2328305
Accessed: 26/11/2009 08:12

THE JOURNAL OF FINANCE • VOL. XLII, NO. 4 • SEPTEMBER 1987

Order Arrival, Quote Behavior, and the Return-Generating Process

JOEL HASBROUCK and THOMAS S. Y. HO*

ABSTRACT

This paper establishes three empirical results. We find positive autocorrelation in actual intra-day stock returns, in intra-day returns computed from quote midpoints, and in the arrival of buy and sell orders. We present a model of return generation that incorporates these features via lagged adjustment of the limit-order price and positive dependence in bid and ask transactions. The return model is observationally equivalent to an ARMA process, which is consistent with the observed return behavior.

MUCH EMPIRICAL RESEARCH HAS focused on the properties of stock returns over brief intervals. Key recent studies include those of Roll [17], Harris [11], Glosten and Harris [8], Amihud and Mendelson [1], and French and Roll [7]. This body of literature is motivated by several considerations. There is overwhelming evidence that stock returns are more complex than the white-noise process associated with the random-walk price evolution.[1] There is consequently a need for a more complete description of the process. Also, the nature of departures from white noise places restrictions on admissible models of the return-generating process implied by microsctructure theory.

Prior empirical research has found that the most significant departure from serial independence of intra-day returns is strongly negative first-order autocorrelation. This paper extends this finding in several ways, the most important being confirmation of the existence of positive return autocorrelation at lags greater than one. We then propose a model of market behavior consistent with these findings.

Niederhoffer and Osborne [15] and Roll [17] show that the negative first-order autocorrelation may arise as a consequence of the bid-ask spread. The Roll model assumes that the midpoint of the bid and ask quotes follows a random walk and that market buy and sell orders arrive randomly, generating trades at observed transaction prices equal to bid or ask prices. In the absence of direct tests of these assumptions, the overall adequacy of the model is largely based on the implications for observed return behavior. In our approach, we supplement the

* Both authors from the Department of Finance, Graduate School of Business Adminstration, New York University. For comments on an earlier draft, we are indebted to Yakov Amihud, Robert Cumby, Andy Lo, John Merrick, Robert Schwartz, our colleagues in the NYU Finance and Economics workshop, and the referees. The data used in this study were supplied by the American Stock Exchange. All errors are our own responsibility.

[1] Departures from random-walk behavior have also been found for return intervals longer than the intra-day periods considered here. See Fama and French [6], Lo and MacKinlay [13], and Poterba and Summers [16].

analysis of returns with direct empirical evidence concerning our assumptions about trade generation and market behavior.

Specifically, we establish three key empirical points. First, transaction-based stock returns exhibit statistically significant positive autocorrelation at orders greater than one. Second, market buy and sell orders do not arrive independently but are in fact also characterized by positive dependence: buys tend to follow buys and sells follow sells. Third, returns based on prices constructed as the midpoint of current bid and ask quotes exhibit positive autocorrelation similar to that found for the actual returns. These results violate assumptions often invoked in the construction of simple return models. Here, we take the second and third findings as descriptive assumptions in modeling market behavior and then show that this model leads to stock return behavior consistent with the first finding.

The paper is organized as follows. Section I describes the data. The following two sections present the key empirical findings. The observed autocorrelations for actual transaction returns are presented in Section II; those for the arrivals of buy and sell orders and returns based on quote midpoints are presented in Section III. We present our model in Section IV, and, in Section V, we demonstrate its consistency with the empirical findings of the earlier sections. Finally, we draw conclusions and implications in Section VI.

I. The Data

Much of the empirical analysis in this paper is based on data that, although publicly available, are not usually collected in convenient form. The data for this study consist of all transactions and quotes for all stocks on the New York Stock Exchange for March and April of 1985 (forty-two days). For each transaction, price, volume, and time (to the nearest second) were reported. These data were collected electronically and essentially constitute a transcription of the exchange tape for these months.

To minimize the effect of price discreteness, only stocks with initial prices above $20 were included in the sample. In addition, stocks that traded very infrequently were deleted (fewer than five hundred transactions in the forty-two days). The remaining sample consisted of 684 companies, with an average market value of $2,142 million. In computing returns, a problem arises due to fragmentation of reported trades; single trades may be reported as multiple trades if different parts of a market order clear at different prices. We therefore took transactions that were separated by at least two minutes. Once this requirement was imposed, the sample comprised roughly 914,000 transactions, with an average time between transactions of 15.9 minutes.

II. Autocorrelations in Transaction Returns

This section presents the evidence of positive autocorrelation in transaction returns, a finding that is the principal motivation for our analysis of the return process.

A. Methodology

Our autocorrelation estimates were formed as follows. To begin with, we computed autocovariances for each stock over forty-two days in our sample. Let $r_{t,s}$ denote the return on day t between transaction s and $s-1$, and let N_t denote the number of transactions on day t (subject to the two-minute minimum discussed above). The estimated autocovariance at lag k was then computed as

$$\hat{\gamma}_k = [\textstyle\sum_{t=1}^{42} (N_t - k)]^{-1}[\sum_{t=1}^{42}\sum_{s=1}^{N_t-k} r_{t,s}r_{t,s-k}] \quad k = 0, \cdots, 12. \tag{1}$$

This is a standard computation except for two features. First, we assume for computational purposes that the expected return is zero. Any reasonable expected return value over such a brief interval is very small compared with the magnitude of the error in its estimation. (See Merton [18].) Second, we wish to limit our attention to intra-day behavior and so do not compute the autocovariances over trading-day boundaries: the overnight return is presumed to follow a different process and is ignored.[2] The autocorrelation estimates were computed in the usual fashion:

$$\hat{\rho}_{k,j} = \hat{\gamma}_{jk}/\hat{\gamma}_{j0} \quad k = 0, \cdots, 12; \quad j = 1, \cdots, 684, \tag{2}$$

where a subscript, j, has been introduced to refer to the particular stock.

For individual stocks, these autocorrelation estimates were highly variable. Accordingly, much of the subsequent analysis is based on the autocorrelation function averaged over the sample. Under the null hypothesis of zero autocorrelations at all orders, the sample variace of $\hat{\rho}_{k,j}$ is approximately the inverse of the number of observations (denoted M_j). The averages were therefore weighted by $w_j = M_j/\sum_j M_j$:

$$\bar{\rho}_k = \textstyle\sum_j w_j \hat{\rho}_{k,j}, \tag{3}$$

where the index runs over the 684 stocks in the sample.[3]

B. Results and Implications

As indicated above, the autocorrelations were computed for individual stocks.[4] The twenty-five percent, fifty percent, and seventy-five percent quartiles are reported for these estimates in Table I along with the weighted means and their standard errors. These means are plotted in Figure 1 as boxes, with the vertical

[2] Limiting the computation to within-day returns imposes a limit on the number of autocovariances that can be successfully computed. The least frequently traded stock in the sample had an average intertransaction time of thirty-eight minutes. For such stocks, the computed autocovariances at long lags may be based on relatively few observations.

[3] Computation of the standard error in this fashion also relies on independence of the observations (the individual estimates). Since all autocorrelations are computed from the same forty-two days of data, this assumption will not be appropriate if strong dependencies exist between individual stock returns. The importance of the market factor over brief return intervals is relatively minor, however, and so these dependencies will be ignored.

[4] With approximately 1,300 observations per stock, the asymptotic standard error of an autocorrelation estimate is roughly 0.028. The first-order autocorrelations were, for individual stocks, almost invariably of a negative sign and magnitude several times this value. Autocorrelations at higher orders individually evinced no consistent pattern and could not be considered individually significant.

Table I

Return Autocorrelations

Lag	Weighted-Mean Autocorrelation	Percentiles 25%	50%	75%	Predicted Autocorrelation
1	−26.364 (0.468)	−31.953	−24.922	−15.501	−26.364
2	0.720 (0.158)	−2.082	0.557	3.574	0.721
3	0.780 (0.140)	−1.637	1.124	3.553	0.748
4	0.307 (0.126)	−1.961	0.311	2.730	0.403
5	0.342 (0.148)	−1.896	0.464	3.365	0.212
6	−0.077 (0.147)	−2.770	−0.132	2.973	0.112
7	0.232 (0.153)	−2.524	0.076	3.041	0.059
8	−0.090 (0.163)	−2.829	−0.310	2.995	0.031
9	0.139 (0.173)	−2.752	0.	3.176	0.016
10	0.109 (0.176)	−2.788	0.	2.920	0.009
11	−0.101 (0.196)	−3.143	−0.398	2.650	0.005
12	0.402 (0.193)	−2.695	0.419	3.469	0.002

Notes: All autocorrelations and related measures are reported in percentages (i.e., × 100). Computation of the weighted means is discussed in Section II; standard errors of these means are given in parentheses. "Predicted autocorrelations" are those generated by the best-fit ARMA(2, 2) process. (See Section V.)

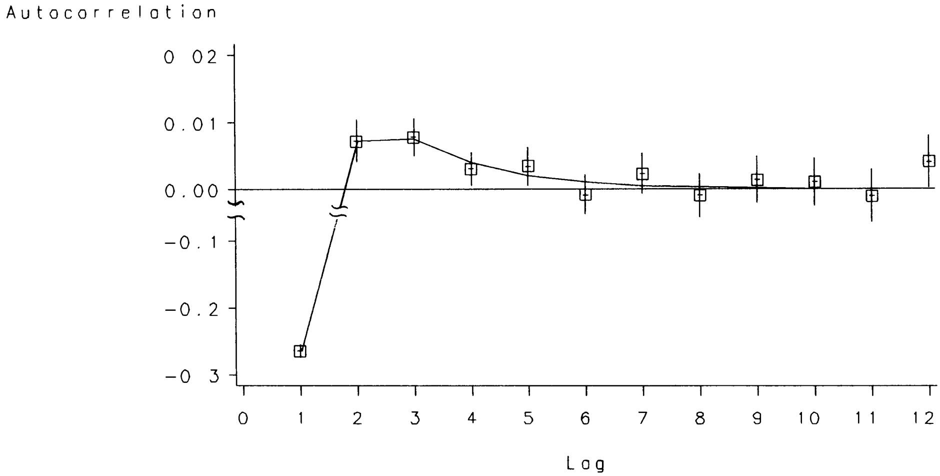

Figure 1. Average autocorrelations of NYSE stock returns; figure depicts plot of values from Table I. The vertical axis is split to show all points.

span marking a range plus or minus two standard errors. The pattern consists of a large negative autocorrelation at the first lag, followed by positive autocorrelations of decreasing magnitude that are statistically significant (at least two standard errors removed from zero) through the fifth lag.[5] The negative first-order autocorrelation in transactions data is consistent with the findings of other studies. The positive autocorrelations at higher lags, however, are (in transactions data) a new finding.[6]

The positive autocorrelations are admittedly small in magnitude, certainly well below the threshold required for a technical trading strategy to be profitable net of transaction costs. Nevertheless, these findings have broad implications for admissible microstructure models. The existence of any nonzero autocorrelation is, of course, inconsistent with martingale price models. The martingale property was first asserted by Samuelson [18] in a framework of perfect competition and no informational asymmetries. It is worth noting, however, that this property also obtains in the Glosten and Milgrom [9] model, which allows for informed traders, and the Kyle [12] model, which permits an informed trader to act cognizant of his or her effect on prices.

Existence of first-order negative correlation is consistent with bid-ask models in which the spread is at least in part due to transactions costs. (See Roll [17] and Glosten and Milgrom [9].) These models cannot, however, account for the presence of positive autocorrelation.

III. Autocorrelation in Buy and Sell Orders and Quote-Midpoint Returns

The basic bid-ask model in which the market maker is assumed to be a simple provider of liquidity relies in part on the assumption that buyers and sellers arrive independently and with equal probability. Hence, a transaction at the bid price is equally likely to be followed by a transaction at the (possibly revised) bid price or ask price. Since dependence in buys and sells may lead to dependence in observed stock returns, an investigation of buy/sell arrival is important.

Since our data include quotes, it is feasible to construct a buy/sell indicator series as follows. We classified as a "buy" any order that transacted at a price greater than the midpoint of the current quotes and assigned the indicator a value of +1. Similarly, a sell was any transaction at a price lower than the midpoint, and the indicator was assigned a value of −1. An "intermediate" value of zero was assigned if the transaction occurred at the midpoint.

The autocorrelations of this series were found to be strongly positive and declining. (The average first-order autocorrelation was 0.160, with a standard error of 0.002.) There are several plausible explanations for this positive auto-

[5] In analysis of subsamples constructed on the basis of market values, the smallest market-value subsample was found to exhibit more pronounced positive autocorrelation.

[6] Using daily data, French and Roll [7] report significant autocorrelations (generally negative) at lags extending through fifteen days. Our estimates of intra-day return autocorrelations suggest that much of the daily return behavior stems from the overnight return, which we do not attempt to model.

correlation. For example, in the process termed "working an order," a trader may distribute purchases or sales over time. In any event, the assumption of random equiprobable buy and sell orders in the simple dealer model is quite suspect. The dependence in buys and sells may lead to the observed dependence in stock returns, and this effect will be explored more fully in Section IV.

Another common assumption in the simple bid-ask spread model is that the bid-ask quotes bracket a "perfect-market" price. If there is no information asymmetry and if the market maker acts as a simple provider of liquidity, the bid and ask quotes will be set symmetrically about this "perfect-market" price. From efficiency considerations, the "perfect-market" price is assumed to follow a martingale. The autocorrelation properties of returns constructed from bid-ask quote midpoints therefore provide an insight into the additional assumption of the simple bid-ask-spread model. Our data include quotes, and, since quote changes are reported immediately, the most recent quotes can be considered current. The timing of these returns is kept identical to the actual returns, and the average autocorrelations are computed for the quote-midpoint returns as described in Section II. The average autocorrelations and standard errors are reported in Table II and exhibit a pattern strikingly similar to that of the actual returns.

Table II

Quote-Midpoint Return Autocorrelations

Lag	Weighted-Mean Autocorrelation	Percentiles		
		25%	50%	75%
1	−2.062 (0.316)	−7.026	−1.893	3.686
2	1.189 (0.228)	−2.255	1.391	5.397
3	1.282 (0.172)	−1.641	1.255	4.732
4	0.605 (0.155)	−2.091	0.783	3.617
5	0.602 (0.153)	−1.735	0.755	3.785
6	−0.116 (0.153)	−2.679	0.	3.235
7	0.226 (0.150)	−2.178	0.344	3.006
8	0.201 (0.166)	−2.533	0.332	3.049
9	0.263 (0.174)	−2.440	0.219	3.253
10	0.142 (0.205)	−2.772	0.170	3.321
11	0.344 (0.194)	−2.568	0.377	3.196
12	0.007 (0.212)	−3.087	0.151	3.066

See notes to Table 1.

IV. The Model of the Return-Generating Process

The evidence presented in the last section establishes the limitations of the simple model. Our intent here is to arrive at a model that is consistent with the empirical evidence. Specifically, we incorporate lagged adjustment in quote-midpoint prices and serial dependence in market orders.

A. Assumptions

As a starting point, we consider the price that would be observed in a perfect market. Such a market is characterized by atomistic competition, costless computation of individual demand functions, costless revelation of these functions to the market, costless trade execution, no constraints on holding or transaction sizes, and no barriers to instantaneous adjustment of market price. This price is assumed to follow (in the logarithm) a random walk. Denoting this as w_t, we have

$$w_t = w_{t-1} + u_t, \tag{4}$$

where u_t is an i.i.d. stochastic process with $\mathrm{E}[u_t] = 0$ and $\mathrm{var}[u_t] = \sigma_u^2$, and the subscript "$t$" indexes transactions.

Suppose now that the market organization permits limit orders, reflecting the current trading propensity of the agents. Cohen, Maier, Schwartz, and Whitcomb [5] show that a bid-ask spread will arise in such a framework with transaction costs. These costs in effect eliminate orders that would otherwise be placed between the current quotes. The limit orders may, of course, be revised in the absence of any transactions.

Denote by p_t the midpoint of the current bid and ask prices. The position of the quotes should certainly reflect the underlying "perfect-market" price w_t, and we would expect that $\mathrm{E}[w_t - p_t] = 0$. Due to transactions costs, however, the limit orders may reflect the actual propensity to trade only with a delay, and p_t may therefore lag w_t. To model the adjustment process, we use a first-order autoregressive approximation given by

$$p_t = p_{t-1} + \alpha(w_t - p_{t-1}) \tag{5}$$

for some adjustment parameter α, $0 < \alpha < 1$.[7] If $\alpha = 1$, the adjustment is instantaneous and w_t and p_t are always identical.

The transaction price that we actually observe will coincide with either the current bid or ask price, as small shifts in the demands of marginal traders induce transactions at these prices. Denote the actual transaction price as π_t:

$$\pi_t = p_t + \varepsilon_t, \tag{6}$$

where ε_t is a binomial random variable serving to "shift" p_t to either the bid or ask price. The sign of ε_t is obviously determined by whether the order is a buy or sell; serial dependence in these events will lead to autocorrelation in the ε_t series.

[7] This specification is quite similar to that proposed by Goldman and Beja [10] and Beja and Goldman [3]. However, their p_t is defined as the clearing price, while ours has the simpler role of quote midpoint. They suggest the addition of a noise term to (5). We omit this in the interest of simplicity; no essential results are changed.

To model this dependence in a simple fashion, we assume that the $\{\varepsilon_t\}$ process behaves as follows. Assume that the bid-ask spread is $2c$ and therefore that the feasible values for ε_t are $+c$ and $-c$. We specify probabilities of these two outcomes as depending on the previous outcome:

$$\varepsilon_t = \begin{cases} +c & \text{with probability} \quad P(\varepsilon_{t-1}) \\ -c & \text{with probability} \quad 1 - P(\varepsilon_{t-1}), \end{cases} \tag{7}$$

where the probability function is given as

$$P(x) = \begin{cases} (1+\delta)/2 & \text{if} \quad x = +c \\ (1-\delta)/2 & \text{if} \quad x = -c \quad \text{for} \quad 0 < \delta < 1. \end{cases}$$

Thus, if the most recent transaction is a market buy order, there is a probability of $(1 + \delta)/2$ that the next transaction will be a market buy order. This specification for $\{\varepsilon_t\}$ is, in essence, a first-order autoregressive scheme, and it may be shown that the autocorrelation at lag k is δ^k.[8]

In summary, equation (4) is a standard efficiency assumption, and equation (6) is definitional. The key equations of our model are (5) and (7), motivated by the empirical evidence in the preceding sections. Equation (5) reflects the observed lagged adjustment in the quote midpoints, which is hypothesized to be a consequence of delays in order revision. Equation (7) captures the serial dependence in buy and sell orders.

B. Derivation of the Return Process

Observed returns may be computed in the usual fashion as the first difference of the transaction prices. Denoting these by r_t and letting L denote the lag operator, we have

$$r_t = (1 - L)\pi_t = (1 - L)p_t + (1 - L)\varepsilon_t.$$

From equation (5), $p_t = (1 - (1 - \alpha)L)^{-1}\alpha w_t$. Therefore,

$$\begin{aligned} r_t &= (1 - (1 - \alpha)L)^{-1}\alpha(1 - L)w_t + (1 - L)\varepsilon_t \\ &= (1 - (1 - \alpha)L)^{-1}\alpha u_t + (1 - \delta L)^{-1}(1 - L)\eta_t, \end{aligned} \tag{8}$$

where $\eta_t = (1 - \delta L)\varepsilon_t$ possesses zero autocovariances at all orders of one or greater. (See footnote 6.) This equation specifies the dynamics of the stock return process. It contains as special cases the random-walk model ($\alpha = 1$, $\delta = \sigma_\eta^2 = 0$) and the Roll bid-ask-spread model ($\alpha = 1$, $\delta = 0$, $\sigma_\eta^2 \neq 0$).

C. Implied Autocorrelations

The autocovariances implied by model (8) for the observed return series are

$$\gamma_0 = \frac{1-\beta}{1+\beta}\sigma_u^2 + \frac{2}{1+\delta}\sigma_\eta^2 + 2(1-\beta)\frac{1-\beta}{1-\delta\beta}\sigma_{u\eta}, \tag{9}$$

[8] A representation equivalent to the one given in the text is $\varepsilon_t = \delta\varepsilon_{t-1} + \eta_t$ where, if $\varepsilon_{t-1} = c$, then $\eta_t = (1 - \delta)c$ or $-(1 + \delta)c$ with probabilities $(1 + \delta)/2$ and $(1 - \delta)/2$, and, if $\varepsilon_{t-1} = -c$, then $\eta_t = -(1 - \delta)c$ or $(1 + \delta)c$ with probabilities $(1 + \delta)/2$ and $(1 - \delta)/2$. By enumeration, it is clear that $E\eta_t\eta_{t-i} = 0$ for $i \neq 0$, and, hence, η_t is serially uncorrelated, although not serially independent.

where $\beta = 1 - \alpha$, and

$$\gamma_k = \beta^k \frac{1-\beta}{1+\beta}\sigma_u^2 - \delta^{k-1}\frac{1-\delta}{1+\delta}\sigma_\eta^2$$

$$+ (1-\beta)\left\{\beta^k \frac{1-\beta}{1-\delta\beta} - \delta^{k-1}\frac{1-\delta}{1-\delta\beta}\right\}\sigma_{u\eta} \quad \text{for} \quad k > 0, \tag{10}$$

where, if δ happens to be zero, the convention $\delta^0 = 1$ is used.[9] If $\alpha = 1$ (instantaneous adjustment), then $\beta = 0$, and the autocovariances are

$$\gamma_0 = \sigma_u^2 + \frac{2}{1+\delta}\sigma_\eta^2 + 2\sigma_{u\eta} \tag{11}$$

and

$$\gamma_k = -\delta^{k-1}\left[\frac{1-\delta}{1+\delta}\sigma_\eta^2 + (1-\delta)\sigma_{u\eta}\right]. \tag{12}$$

The implications of (9) through (12) for admissible autocovariance structures will be discussed in Section V.

D. An Equivalent ARMA Representation

While Sections II through IV have presented much evidence in support of the model, many of the parameters in (8) are not amenable to direct measurement. It is therefore useful to develop an alternative form of (8) that is, from the viewpoint of returns, observationally equivalent. This form can be shown to be second-order autoregressive and second-order moving average, conventionally denoted ARMA(2, 2).

To see this, note that the return equation (8) may be rearranged as

$$(1 - (1-\alpha)L)(1-\delta L)r_t = (1-\delta)\alpha u_t + (1 - (1-\alpha)L)(1-L)\eta_t. \tag{13}$$

[9] Equations (9) and (10) may be established as follows. Defining $\beta = 1 - \alpha$, the return autocovariances may be computed as

$$\begin{aligned}\gamma_k &= \mathrm{E}r_t r_{t-k}\\ &= (1-\beta)^2\mathrm{E}[(1-\beta L)^{-1}u_t(1-\beta L)^{-1}u_{t-k}]\\ &\quad + \mathrm{E}[(1-\delta L)^{-1}(1-L)\eta_t(1-\delta L)^{-1}(1-L)\eta_{t-k}]\\ &\quad + (1-\beta)\mathrm{E}[(1-\beta L)^{-1}u_t(1-\delta L)^{-1}(1-L)\eta_{t-k}]\\ &\quad + (1-\beta)\mathrm{E}[(1-\delta L)^{-1}(1-L)\eta_t(1-\beta L)^{-1}u_{t-k}].\end{aligned} \tag{10}$$

If $|\beta| < 1$ and $|\delta| < 1$, we may use power-series expansions to obtain

$$(1-\beta L)^{-1} = 1 + \beta L + \beta^2L^2 + \beta^3L^3 + \cdots.$$

$$(1-\delta L)^{-1}(1-L) = 1 + (\delta+1)L + \delta(\delta-1)L^2 + \delta^2(\delta-1)L^3 + \cdots.$$

Equation (9) results from using these expansions in (16), taking expectations, and summing the resulting geometric series. Equation (10) follows similarly.

The left-hand side of (13) is autoregressive of order 2 and may be written as

$$r_t - \phi_1 r_{t-1} - \phi_2 r_{t-2}, \quad \text{where} \quad \phi_1 = (1 - \alpha) + \delta \quad \text{and} \quad \phi_2 = -(1 - \alpha)\delta.$$

Since u_t and η_t are serially uncorrelated processes, the right-hand side of (13) will therefore exhibit positive autocovariances at lags 0, 1, and 2, and this remains true even if u_t and η_t are contemporaneously correlated. Ansley, Spivey, and Wrobleski [2] show that a stochastic process with zero autocovariances beyond order q possesses a moving-average representation of order q. The right-hand side may therefore be written as $e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$, where the e_t are serially uncorrelated disturbances, and, in summary,

$$r_t - \phi_1 r_{t-1} - \phi_2 r_{t-2} = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}, \tag{14}$$

which describes an ARMA(2, 2) process. It is clear that representation (13) is richer in the sense of capturing the various component contributions to returns, but, if only returns are observable, (14) is an equally valid description.

The ARMA(2, 2) model (14) has five parameters: ϕ_1, ϕ_2, θ_1, θ_2, and σ_e^2. With respect to the original representation in (13), it is clear that two convenient parameter relationships exist, namely those noted above for ϕ_1 and ϕ_2. Other relationships are unfortunately more complex and will not be further developed here. In particular, we emphasize that there is no simple transformation relating $\{e_t\}$ and $\{u_t, \eta_t\}$: the expression for e_t will generally be a linear function of infinite order in prior values of $\{u_t, \eta_t\}$.[10]

The autocovariances for the ARMA(2, 2) representation are needed in the subsequent estimation procedure. These may be computed following Box and Jenkins [4], pp. 74–75. Define the cross-covariance function between the driving shocks and the process, $\gamma_{re,k} = \mathrm{E} r_{t-k} e_t$. Taking $\sigma_e^2 = 1$ for convenience, we then have $\gamma_{re,0} = 1$; $\gamma_{re,-1} = \phi_1 - \theta_1$; $\gamma_{re,-2} = \phi_1 \gamma_{re,-1} + \phi_2 - \theta_2$. The autocovariances may then be computed from the relationships.

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \gamma_{re,0} - \theta_1\gamma_{re,-1} - \theta_2\gamma_{re,-2}$$

$$\gamma_1 = \phi_1\gamma_0 + \phi_2\gamma_1 - \theta_1\gamma_{re,0} - \theta_2\gamma_{re,-1}$$

$$\gamma_2 = \phi_1\gamma_1 + \phi_2\gamma_0 - \theta_2\gamma_{re,0}$$

$$\gamma_k = \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2} \quad \text{for} \quad k = 3, \tag{15}$$

The autocorrelations may then be computed in the usual fashion, $\rho_k = \gamma_k/\gamma_0$.

[10] The representation in (13) is in the general case characterized by five parameters $\{\alpha, \delta, \sigma_u^2, \sigma_\eta^2, \sigma_{u\eta}\}$. With the exception of the relationships indicated above between $\{\phi_1, \phi_2\}$ and $\{\alpha, \delta\}$, no closed-form relationships exist between the two sets of parameters. Numerically, the transformation may be effected by solving the nonlinear system of equations obtained by equating the two sets of expressions for the return autocovariances: (9) and (10) for representation (13) and (15) for representation (14).

V. The Evidence in Support of the Model

In this section, we develop the observable implications of the model and analyze them in light of the empirical results presented in Section I. Our primary goal is to establish the consistency of the model with the observed behavior of actual stock returns.

A. Analysis of the Return Model

The actual return autocorrelation pattern was found to be negative at lag one, followed by decaying positive autocorrelations at higher lags. It is easily verified from (9) and (10) that the return-generating process associated with our model (13) can imply this behavior if δ is small relative to ϕ and σ_η^2 is large relative to σ_u^2. (For example, take $\alpha = .5$, $\delta = 0$, $\sigma_u^2 = 1$, $\sigma_\eta^2 = 10$, $\sigma_{u\eta} = 0$.)

We now turn to some estimation results. Our return model (13) was shown to be observationally equivalent to an ARMA(2, 2) model of the form (14). Estimation can proceed as follows. Given the parameters, ϕ_1, ϕ_2, θ_1, and θ_2 for the ARMA(2, 2) process, implied autocorrelations may be computed from (9) and (10). By varying these parameters, a least-squares fit to the actual sample autocorrelations may be obtained; this is essentially the method proposed by Box and Jenkins [4] to obtain "preliminary" parameter estimates. This calculation was done using the IMSL ZXSSQ subroutine to perform the nonlinear optimization. While this procedure is useful in obtaining the ARMA representation of the process, there are no convenient standard errors for the estimated parameters. Accordingly, inference regarding the exact values of the adjustment parameters is limited.[11] Autocorrelations at all twelve lags reported in Table I were used in the estimation, a generous number since most of the estimates at higher lags are statistically indistinguishable from zero. To test the sensitivity, various lag lengths from eight to twelve were used, and no significant changes in the estimated parameter values were found.

For the actual return series, we estimated $\phi_1 = 0.55$, $\phi_2 = -0.01$, $\theta_1 = 0.83$, and $\theta_2 = -0.18$. The autocorrelations associated with these values are plotted as a connected line in Figure 1. These values fall well within the two-standard-deviation bounds, suggesting that the ARMA(2, 2) model fits the data well. By

[11] While it would in principle be possible to estimate the ARMA model directly via maximum likelihood on the total pooled sample of stock returns, substantial problems arise. One such problem involves the presumed distribution of the disturbances. The average coefficient of kurtosis for the returns is 11.9, substantially higher than the value of 3 associated with the normal distribution. The departure is certainly substantial enough to render extremely dubious any standard errors based on a normality assumption. There are also computational problems; one likelihood computation would require a pass through the nearly one million observations.

The empirical analysis has been based on autocorrelation functions averaged over samples and subsamples of stocks. The question may arise as to whether the distinctive shape of this function may be a consequence of the averaging. Suppose each individual autocorrelation function is (beyond the first lag) something closer to a step function (strongly positive, say, at lags two, three, and four and zero thereafter). If the location of the "step" varies across stocks, aggregation may produce the illusion of a gradual decay. While the existence of this effect cannot be ruled out entirely, we note that, if it were significant, histograms of the autocorrelations would be skewed, with a mode at zero. The sample histograms displayed no such characteristics.

comparing (13) and (14), it is clear that $\phi_1 = (1 - \alpha) + \delta$ and $\phi_2 = -(1 - \alpha)\delta$. If $(1 - \alpha)$ and δ are not near unity, as suggested by the size of ϕ_1, it is not surprising that their product, ϕ_2, is small.

If ϕ_2 is taken to be zero, the model reduces to an ARMA(1, 2). This reduced form is somewhat noteworthy because no more parsimonius ARMA representation is capable of giving rise to autocorrelations that are negative at one lag and positive at all higher lags. Since the average return autocorrelations in Table I are reliably negative at lag 1 and then positive through lag 5, we may reject models such as AR(1), AR(2), MA(1), MA(2), and ARMA(1, 1).[12] That is, no return model simpler than an ARMA(1, 2) can be consistent with our observed return autocorrelations.

B. The Buy/Sell and Quote-Midpoint Return Behavior

As noted in the last section, the autocorrelation pattern suggested by the model (7) is that of a first-order autoregressive series; the autocorrelations should decay at rate δ. While the observed autocorrelations do diminish with lag, the rate of decay is less than one would expect. Since the autocorrelation at lag one is about 0.15, that at lag two would be expected to be $(0.15)^2$, which is substantially less than the observed 0.08. The specification (7) is therefore best considered an approximation consistent with the implications for actual return autocorrelations.

The implications of the buy/sell autocorrelation for the return-generating function are worthy of note. First, the existence of positive autocorrelation in buys and sells by itself cannot generate the pattern observed in return autocorrelations. If the adjustment of p_t toward w_t is instantaneous ($\alpha = 1$), then, from (12), it is apparent that (since $\delta > 0$) all autocorrelations of lag one and above must have the same sign. A switch from negative first-order to positive higher order autocorrelations cannot arise. Second, if we wish only to account for the qualitative features of the return autocorrelations, the dependence we built into our buy/sell disturbances is superfluous. Even if $\delta = 0$, it is clear from (10) that, for $0 < \alpha < 1$ and σ_η^2 sufficiently large, the observed return autocorrelation pattern can arise. The observed autocorrelation in the buy/sell indicator series, however, is quite substantial. Since this autocorrelation enters the autoregressive part of the return-generating model (14), the autoregressive coefficients, particularly ϕ_1, will reflect this contribution.

The quote-midpoint returns shown above exhibit autocorrelation behavior substantially similar to that of the actual returns. Under plausible assumptions, this is also consistent with our model.[13]

[12] Nonzero autocorrelations at lags greater than two rule out MA(1) and MA(2) models (Box and Jenkins [4], p. 68). AR(1) and ARMA (1, 1) processes may exhibit atuocorrelation functions that have both negative and positive values, but, if so, the signs must alternate (Box and Jenkins [4], pp. 57 and 78). The autocorrelation functions permitted with an AR(2) process are more complex, but a pattern of negative first-order and positive second-, third-, and fourth-order autocorrelations is inadmissible.

[13] In the model, the quote-midpoint returns may be computed as

$$s_t = (1 - L)p_t = (1 - (1 - \alpha))^{-1}\alpha u_t.$$

VI. Implications and Conclusions

This paper has described three new empirical findings. First, we have established the existence of positive autocorrelation in actual returns. Second, our construction and analysis of a market-order series establish positive serial dependence in the arrival of buys and sells. Third, the presence of positive autocorrelation in returns computed from quote midpoints suggests that the positive autocorrelation in actual returns is not solely due to positive dependencies in order arrival but reflects lagged adjustment in quotes. Thus, while most prior literature has been based solely on the properties of transaction returns, we have extended the analysis to incorporate relevant properties of the quote series as well.

These findings contradict certain assumptions of prior models of the return-generating process and motivate the development of an alternative framework. By virtue of our analysis of order arrival and quote-midpoint return behavior, this model is based on assumptions that have been more completely established. The proposed model implies an ARMA(2, 2) structure for transaction returns and, under plausible parameter estimates, is quite close to an ARMA(1, 2) model, the most parsimonious ARMA structure consistent with the observed return behavior. These findings for an aggregated sample should stimulate research on the precise estimation of the return-generating process for individual stocks. They should also have broad implications for market microstructure research, computation of stock return variance, and intra-day event studies.

This specification is equivalent to an AR(1) repesentation and, as noted above, cannot yield a pattern consisting of negative first-order, positive higher order autocorrelations. If, however, our quote midpoints are subject to observational error, v_t, we have

$$s_t = (1 - L)(p_t + v_t) = (1 - (1 - \alpha)L)^{-1}\alpha u_t + (1 - L)v_t$$

or

$$(1 - (1 - \alpha)L)s_t = \alpha u_t + (1 - (1 - \alpha)L)(1 - L)v_t.$$

This is of the same form as the actual return specification (8) or (13) with $\delta = 0$, which leads to an ARMA(1, 2) model capable of generating the observed autocorrelation pattern. Observation error can arise quite easily, most likely as a consequence of price rounding. (Although an observational error could be added to the actual return specification in (8), it would be indistinguishable from η_t.)

REFERENCES

1. Yakov Amihud and Haim Mendelson. "Trading Mechanism and Stock Returns: An Empirical Investigation." Working Paper, University of Rochester, 1986.
2. Craig F. Ansley, W. Allen Spivey, and William J. Wrobleski. "On the Structure of Moving Average Processes." *Journal of Econometrics* 6 (July 1977), 121–34.
3. Avraham Beja and Barry Goldman. "On the Dynamics of Behavior of Prices in Disequilibrium." *Journal of Finance* 35 (May 1980), 235–48.
4. George E. P. Box and Gwilym M. Jenkins. *Time Series Analysis: Forecasting and Control.* San Francisco: Holden-Day, 1976.
5. K. Cohen, S. Maier, R. Schwartz, and D. Whitcomb. "Transactions Costs, Order Placement Strategy, and Existence of the Bid/Ask Spread." *Journal of Political Economy* 89 (April 1981), 287–305.

6. Eugene F. Fama and Kenneth R. French. "Permanent and Temporary Components of Stock Prices." Working Paper, University of Chicago, 1986.
7. Kenneth R. French and Richard Roll. "Stock Return Variances: The Arrival of Information and the Reaction of Traders." *Journal of Financial Economics* 17 (September 1986), 5–26.
8. Lawrence Glosten and Lawrence Harris. "Estimating the Components of the Bid/Ask Spread." Working Paper, Northwestern University, 1985.
9. Lawrence Glosten and Paul Milgrom. "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." *Journal of Financial Economics* 14 (March 1985), 71–100.
10. Barry Goldman and Avraham Beja. "Market Prices vs. Equilibrium Prices: Return's Variance, Serial Correlation, and the Role of the Specialist." *Journal of Finance* 34 (June 1979), 595–607.
11. Lawrence Harris. "Estimation of 'True' Price Variances and Bid-Ask Spreads From Discrete Observations." Working Paper, University of Southern California, 1985.
12. Albert S. Kyle. "Continuous Auctions and Insider Trading." *Econometrica* 53 (November 1985), 1315–36.
13. Andrew W. Lo and A. Craig MacKinlay. "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test." Rodney L. White Center Working Paper 5-87, Wharton School, 1987.
14. Robert Merton. "Estimating the Expected Rate of Return." *Journal of Financial Economics* 8 (December 1980), 323–62.
15. Victor Niederhoffer and M. F. M. Osborne. "Market Making and Reversal on the Stock Exchange." *Journal of the American Statistical Association* 61 (December 1966), 897–916.
16. James Poterba and Lawrence Summers. "Intertemporal Reversion in Stock Returns: Evidence and Implications." Working Paper, M. I. T., 1987.
17. Richard Roll. "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market." *Journal of Finance* 39 (September 1984), 1127–39.
18. Paul Samuelson. "A Proof That Properly Anticipated Prices Fluctuate Randomly." *Industrial Management Review* 6 (Spring 1965), 41–49.