



ELSEVIER

Journal of Financial Markets 1 (1998) 285–319

Journal of
FINANCIAL
MARKETS

Endogenous market statistics and security pricing: An empirical investigation

Thomas J. George^{a,*}, Chuan-Yang Hwang^b

^a College of Business Administration, University of Iowa, Iowa City, IA 52242, USA

^b Hong Kong University of Science and Technology, Hong Kong

Abstract

This study examines empirically the degree to which the history of daytime and overnight price changes and order flow affects estimates of traders' beliefs about future security price changes. Estimates indicate that forecasts of the permanent component of price changes occurring after the open of trading are significantly related to past price changes and order flow; but the same is not generally true for price changes occurring after the close. These results are consistent with models of technical analysis, and models in which the process of trading facilitates price discovery. The evidence also suggests that private information is an important determinant of price movements. © 1998 Elsevier Science B.V. All rights reserved.

JEL classification: G12; G14

Keywords: Microstructure; Information; Technical Analysis

1. Introduction

Several theoretical models analyze the informational role of prices in securities markets (e.g., Grossman, 1976, 1978; Danthine, 1978; Diamond and Verrecchia, 1981; Verrecchia, 1982; Admati, 1985; Grossman and Stiglitz, 1980). These models show that equilibrium prices aggregate information that individual traders possess about a security's fundamental value, and that prices convey valuable information to those who trade securities. A large body of empirical

* Corresponding author. E-mail: tom-george@uiowa.edu

work builds on these insights by viewing changes in prices as reflective of information flow (e.g., French and Roll, 1986; Barclay et al., 1990; Jones et al., 1994; Forster and George, 1995). Each of these studies compares the variability of price changes when markets are open to times when markets are closed. Differences in variability are attributed to the flow of private information because, according to the theories, trading is necessary for prices to change in response to private information. An implicit assumption in both the theories and empirical tests is that the entire effect of private information on prices occurs contemporaneously with trading. However, if prices are not effective aggregators of information, the full impact of private or public information on prices may not be realized until after the initial trading on that information occurs. Under these circumstances, individuals might rationally condition their beliefs, and hence trading strategies, on past values of endogenous market statistics such as past price changes and measures of past trading activity.

Brown and Jennings (1989), Grundy and McNichols (1989), and Blume et al. (1994) present theoretical models in which individuals optimally use past price changes and/or volume as information in making their trading decisions. In these models, past prices and volume convey information for making investment decisions that is incremental to the current security price. Although many models of market micro-structure predict associations between volume and price variability, the models of technical analysis are unique in that market statistics affect future price changes through their effect on traders' beliefs. This point is emphasized by Blume et al. (p. 154), whose model focuses on the role of volume:

In other models ... volume is interesting for its correlation with other variables, but in itself is unimportant: traders never learn from volume nor use volume in any decision making. By contrast, in our model volume enters traders' learning problems because they use the specific volume statistic in updating their beliefs. Consequently, volume matters in our model because it affects the behavior of the market, rather than merely describes it.

In this paper, returns and order flow are modeled as a vector time-series process and variance decompositions are employed to estimate the sensitivity of traders' conditional expectations (or forecasts) about future price changes to past price changes and order flow. The sample contains four sets of fifty NYSE stocks, each set chosen from a different quartile of trading activity.

Estimates of the sensitivity of forecasts of price changes after the open of trading to past price changes and order flow are economically and statistically significant. This finding is consistent with the predictions of models of technical analysis, and suggests that opening prices do not fully reflect all information in the trading history. By contrast, the sensitivity of forecasts of price changes after

the close are not economically significant, except for the least-actively traded stocks in the sample. In addition, the information that is not incorporated into opening prices is almost fully incorporated into closing prices. This suggests that closing prices generally do reflect fully information in the trading history; and is consistent with the hypothesis that daytime trading facilitates the incorporation of information in the trading history that is not reflected in opening prices. Also consistent with this interpretation is the fact that variation in daytime order flow is more strongly associated with past changes in prices and volume than is variation in opening order flow.

Comparing opening and closing prices is useful for assessing the validity of theories that predict that the process of trading facilitates price discovery (e.g., Madhavan, 1992; and Leach and Madhavan, 1993). Other studies examine this prediction empirically by measuring trading costs, or the variance of the transitory component of security returns (e.g., Amihud and Mendelson, 1991; Kleidon and Werner, 1996; Hasbrouck, 1993; Chan et al., 1994). In those studies, the quality of price discovery is assessed by the smallness of these quantities. The results of this study complement theirs by using a different metric for evaluating the quality of price discovery – the completeness with which prices reflect value-relevant information in the trading history.

The econometric approach in this paper is similar to that of Hasbrouck (1991) and Madhavan et al. (1997). Their studies isolate and analyze the permanent component of price changes to abstract carefully from the effects of market frictions. The main difference between the approaches is that this study examines the impact of past versus current information on the permanent component of price changes. Their studies focus primarily on how the sensitivity of permanent price changes to contemporaneous order-flow shocks changes over the course of a day.

The approach in this paper to the technical analysis issue is different from recent studies that examine the profitability of specific trading rules (e.g., Neftci, 1991; Brock et al., 1992; Conrad and Kaul, 1996), or that search for optimal rules (Allen and Karjalainen, 1995). Those studies test whether rules that generate buy and sell signals as functions of past levels of a stock index earn excess returns in trading the index. In this respect, they shed considerable light on whether the *systematic* component of stock prices behaves in accordance with traditional notions of the efficient markets hypothesis. They are not, however, designed for testing the implications of theoretical models in which technical analysis is an equilibrium phenomenon. The analysis in this paper focuses on *individual* securities, and the metric for judging whether technical analysis has value reflects the criterion set forth in the theoretical work; namely whether rational forecasts of future price changes are sensitive to innovations in returns and trading activity in the recent past. Efficacy of the models cannot be assessed by computing excess returns to specific strategies because without additional information about the form of traders' demand functions (e.g., their risk tolerance,

level of uncertainty, etc.), forecasts of price changes cannot reliably be translated into the buy and sell signals that traders presumably follow in equilibrium.

This paper is also distinct from recent empirical studies of return predictability that examine the association between volume and security return autocorrelations (e.g., Campbell et al., 1993; Conrad et al., 1994). In those studies, current or lagged volume is modeled as a determinant of the degree to which future returns can be predicted from past returns. Those studies do not examine volume as a predictor, nor do they distinguish between changes in prices that are transitory and changes that are permanent. In this study, past returns *and* order flow are modeled as predictors of future price changes; which is important if trading activity conveys information about first moments of security returns. In addition, the variance of the *permanent* component of future price changes that is related to past innovations in returns and order flow is estimated. This helps to avoid attributing predictability of temporary components of price changes to changes in beliefs about the underlying value of securities. This is important because market frictions affect significantly the dynamics of security returns (see, e.g., Stoll, 1989; George et al., 1991; Huang and Stoll, 1994), and have been shown to explain a large portion of the variation in transaction returns over short horizons (see, e.g., Hasbrouck, 1993; Madhavan et al., 1997).

The next section outlines the estimation technique. Section 3 describes the data, and Section 4 contains the empirical estimates. A brief conclusion is provided in Section 5.

2. Empirical method

2.1. Revisions of forecasts of future returns

The goal of this analysis is to estimate the sensitivity of traders' expectations about security value to previous innovations in prices and volume. These estimates are derived from a vector auto-regressive empirical model of the dynamics of price changes and order flow: $\mathbf{x}_t \equiv (r_{dt}, r_{nt}, v_{dt}, v_{nt})'$. The variables r_{nt} and r_{dt} are the overnight and daytime security returns ending on day t , respectively. The variables v_{nt} and v_{dt} are measures of order flow at the opening and during the daytime (excluding the opening transaction) on day t , respectively. Since v_{nt} is trading associated with the overnight return, it is sometimes referred to as 'overnight' order flow for expositional convenience.

This four-variate VAR specification is convenient for working out the analytics of the variances of forecast revisions (described below). It is equivalent to a bivariate specification with state-dependent (i.e., day or night) coefficients and disturbance variances. This enables the model to capture differences between the dynamic behavior of daytime and overnight returns. Differences in return dynamics reflect potential differences in the way opening and closing prices

aggregate information. For example, if trading facilitates price discovery, then closing prices may be better aggregators of information than opening prices. In this case, the dependence of daytime returns on past returns and order flow will be different from that of overnight returns.¹ Our specification detects this by allowing for different estimates of the dynamic behavior of daytime and overnight returns, which in turn affects comparisons of the variances of forecast revisions described in Eqs. (1) and (2) below. Filtering these differences out of the data (as would be done, for example, to eliminate seasonals in an analysis of quarterly GDP) would prevent us from examining the prediction of theoretical models that trading facilitates price discovery.

Both price changes and order flow are included because the theories differ in the variable whose past realization is informative to traders – past prices in Brown and Jennings (1989) and Grundy and McNichols (1989), volume in Blume et al. (1994). These variables are correlated (both in the models and in reality) and whether past innovations to either or both are useful conditioning information is an empirical question. The estimates in the tables below provide an assessment of their relative importance.

Returns and order flow are first-differences in (log) prices and security holdings, respectively. If the time series $\{x_t\}_t$ is stationary and invertible, the technique developed by Beveridge and Nelson (1981) can be used to estimate the impact of innovations in these variables on forecasts (i.e., conditional expectations) of long-run price changes. Changes in these forecasts are used as estimates of changes in traders' beliefs.

The time line in Fig. 1 helps to illustrate this approach. Points in time marked $\{ct - 2, ct - 1, ct, \dots\}$ are the close of trading on days $\{t - 2, t - 1, t, \dots\}$; and $\{ot - 1, ot, \dots\}$ are the open of trading on days $\{t - 1, t, \dots\}$. The series of closing and opening (log) prices are given by $\{p_{ct-2}, p_{ct-1}, \dots\}$ and $\{p_{ot-1}, p_{ot}, \dots\}$, respectively.

Consider the problem of forecasting how the security's value will change from ot through the indefinite future (i.e., the permanent change in the price from its level at ot). During the 24-hour time interval $(ct - 1, ct]$, the market generates the unexpected component of four statistics captured by the VAR specification that could affect traders' beliefs: (i) the overnight return from $ct - 1$ through ot , (ii) order flow at the open ot , (iii) the daytime return from ot through ct , and (iv) order flow during day t (exclusive of the opening). The first two statistics are past information in the sense that if the security price at ot , p_{ot} , fully reflects information in the trading history, neither (i) nor (ii) would affect beliefs about permanent changes in prices that occur after ot . Therefore, whether technical

¹ For evidence that errors in transaction prices induced by market frictions differ at the open and close, see for example, Amihud and Mendelson (1987), Stoll and Whaley (1990), and Forster and George (1996). Hasbrouck (1991) presents evidence on intra-day differences in rates of information flow into security prices.

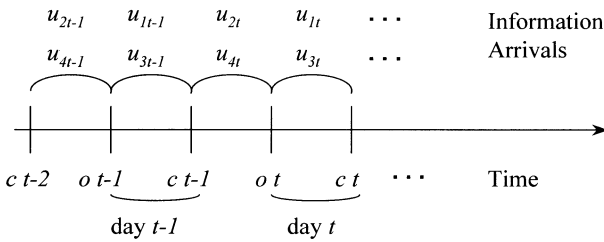


Fig. 1. The timing of information arrivals.

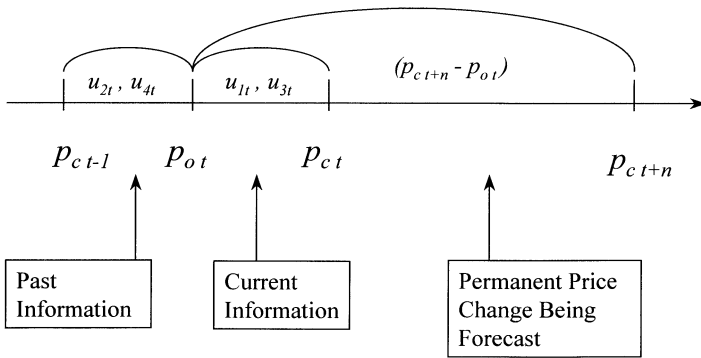


Fig. 2. Forecasts based on 24 h of current and past information.

analysis has value can be tested by assessing the significance of the impact of (i) and (ii) on changes in expectations of permanent price changes.

The revision in expectations of the permanent price change following the open is given formally by

$$\Delta_{ot}(ct - 1, ct) \equiv \lim_{n \rightarrow \infty} \{E_{ct}[p_{ct+n} - p_{ot}] - E_{ct-1}[p_{ct+n} - p_{ot}]\}, \quad (1)$$

where $E_{ct}[\cdot]$ denotes expectation conditional on information available at time ct . To estimate how informative are overnight returns and order flow for forecasts of permanent changes in prices after the open of trading, the variance of $\Delta_{ot}(ct - 1, ct)$ is decomposed into components related to (i) – (iv) above, then examined for whether the components relating to (i) and (ii) are large in economic terms and statistically significant (see Fig. 2).

Similarly, the degree to which daytime returns and order flow are informative for forecasts of permanent changes in security prices after the close of trading is

examined by decomposing the variance of

$$\Delta_{ct-1}(ot - 1, ot) \equiv \lim_{n \rightarrow \infty} \{E_{ot}[p_{ot+n} - p_{ct-1}] - E_{ot-1}[p_{ot+n} - p_{ct-1}]\}, \quad (2)$$

and examining whether the components relating to the return and order flow during $(ot - 1, ct - 1]$ are large in economic terms and statistically significant. If p_{ct-1} fully reflects all value-relevant information in the trading history, neither the return nor order flow during $(ot - 1, ct - 1]$ should affect beliefs about permanent price changes occurring after $ct - 1$.²

It may seem strange that we work with forecast revisions that occur during periods that straddle the price in question, rather than a period preceding this price. For example, in examining whether technical analysis has value at the close, we work with $\Delta_{ct-1}(ot - 1, ot)$ instead of $\Delta_{ct-1}(ot - 1, ct - 1)$. Information that arrives during $(ct - 1, ot]$ will be highly relevant to forecasts of permanent price changes starting at $ct - 1$; though its relevance says nothing about the value of technical analysis. We do not interpret variation attributable to information in the interval $(ct - 1, ot]$ as evidence that technical analysis has value, however. The only reason we use it is to gauge the *economic* significance of information in the earlier interval $(ot - 1, ct - 1]$ relative to information in the later period. Specifically, tables report the proportion of the variance of $\Delta_{ct-1}(ot - 1, ot)$ attributable to the early $(ot - 1, ct - 1]$ interval, and the proportion attributable to the later $(ct - 1, ot]$ interval for comparison. We believe that cross-sectional averages of these numbers are easier to interpret than averages of raw estimates of variance components. The statistical tests do not use proportions. Inferences are drawn from tests of whether the component of the variance (in variance units) attributable to the early period is zero. Tables report rejection rates from such tests at the individual security level.

2.1.1. Estimation

It suffices to describe the estimation procedure for the case where $\mathbf{x}_t \equiv (r_{dt}, r_{nt}, v_{dt}, v_{nt})'$, is assumed to be auto-regressive of order one,

$$A\mathbf{x}_t = B\mathbf{x}_{t-1} + \mathbf{u}_t, \quad (3)$$

because models with higher order dependence can be re-written as first-order models by stacking the variables (see Hamilton (1994) (p. 259)). A and B are 4×4 matrices, and $\mathbf{u}_t = (u_{1t}, u_{2t}, u_{3t}, u_{4t})'$ is an iid random vector of disturbances—unexpected returns and order flow. In this specification, the matrix A captures contemporaneous dependence among the variables, and B captures

²To minimize the potential impact of micro-structure frictions on the starting points from which permanent price changes are computed (p_{ot} in Eq. (1) and p_{ct-1} in Eq. (2)), midquotes are used in estimation rather than transaction prices. Sampling is discussed further in Section 3.

lagged dependence. The expression for Δ_{ot} in terms of elements of the \mathbf{u}_t vector is

$$\begin{aligned} \Delta_{ot}(ct - 1, ct) = & \left\{ \sum_{k=0}^{\infty} \beta_{11}(k)u_{1t} + \sum_{k=0}^{\infty} \beta_{12}(k)u_{2t} + \sum_{k=0}^{\infty} \beta_{13}(k)u_{3t} \right. \\ & \left. + \sum_{k=0}^{\infty} \beta_{14}(k)u_{4t} \right\} + \left\{ \sum_{k=1}^{\infty} \beta_{21}(k)u_{1t} + \sum_{k=1}^{\infty} \beta_{22}(k)u_{2t} \right. \\ & \left. + \sum_{k=1}^{\infty} \beta_{23}(k)u_{3t} + \sum_{k=1}^{\infty} \beta_{24}(k)u_{4t} \right\}, \quad (4) \end{aligned}$$

where $\beta_{ij}(k)$ is the moving average coefficient at lag k corresponding to the j th disturbance in the equation for x_{it} (see Appendix A). The intuition behind this formula is as follows. A shock to daytime (overnight) returns, $u_{1t}(u_{2t})$, affects expectations of the security's value through its effect on expectations of current and future daytime returns – the first (second) sum in the first set of curly brackets – and through its effect on expectations of future overnight returns – the first (second) sum in the second set of curly brackets. If p_{ot} fully reflects all value-relevant information in the trading history, then u_{2t} , the shock to the *previous* overnight return, has no impact on beliefs about changes in value during day t and during future daytime and overnight periods.

The impact of past market statistics on beliefs at the open of trading is quantified by the component of the variance of $\Delta_{ot}(ct - 1, ct)$ that is related to *past* shocks u_{2t} and u_{4t} . If the covariance matrix of \mathbf{u}_t is diagonal, these components are

$$\left\{ \sum_{k=0}^{\infty} \beta_{12}(k) + \sum_{k=1}^{\infty} \beta_{22}(k) \right\}^2 \sigma_2 \quad \text{and} \quad \left\{ \sum_{k=0}^{\infty} \beta_{14}(k) + \sum_{k=1}^{\infty} \beta_{24}(k) \right\}^2 \sigma_4, \quad (5)$$

where $\sigma_h \equiv \text{Var}[u_{ht}]$.

The formula for $\Delta_{ct-1}(ot - 1, ot)$ is similar, as is the intuition. If p_{ct-1} fully reflects information in the trading history, then u_{1t-1} and u_{3t-1} – the shocks to the daytime return and daytime order flow *prior* to $ct - 1$ – have no impact on beliefs about changes in security value during future overnight and daytime periods. The importance of these elements of the trading history are quantified by the components of the variance of $\Delta_{ct-1}(ot - 1, ot)$ attributable to u_{1t-1} and u_{3t-1} . If the covariance matrix of \mathbf{u}_t is diagonal, these components are

$$\left\{ \sum_{k=1}^{\infty} \beta_{11}(k) + \sum_{k=1}^{\infty} \beta_{21}(k) \right\}^2 \sigma_1 \quad \text{and} \quad \left\{ \sum_{k=1}^{\infty} \beta_{13}(k) + \sum_{k=1}^{\infty} \beta_{23}(k) \right\}^2 \sigma_3. \quad (6)$$

Since these measures are defined in terms of changes in expectations of *long-horizon* price changes, they capture the *permanent* effect of shocks to returns and order flow, and therefore abstract from any temporary effects these shocks might have on future prices (perhaps because of market structure imperfections). This can readily be seen in the formula for $\Delta_{ot}(ct - 1, ct)$ in

Eq. (4), which depends on sums of *signed* moving-average coefficients. Since temporary effects of shocks reverse themselves over time, they have no impact on $\Delta_{oi}(ct - 1, ct)$. Hasbrouck (1991) uses this aspect of variance decompositions to abstract from market structure frictions in estimating the impact of individual trades on permanent changes in bid and ask quote midpoints within the day.

Individual moving-average coefficients, β_{ih} , in Eqs. (5) and (6) can be expressed analytically in terms of the auto-regressive parameters A and B [see, e.g., Judge et al. (1985) (p. 657)]. Variance components are typically estimated by approximating an infinite sum with a truncated sum of its first several terms, then substituting estimates of the auto-regressive parameters into the analytical expressions for the moving-average coefficients that constitute the truncated sum [see, e.g., Hodrick, 1982; Hasbrouck, 1991, 1993; Campbell and Ammer, 1993]. This is justified because in order for their infinite sum to be convergent, the moving-average coefficients must decline rapidly.³

In this study, the infinite sums are not approximated. Instead, exact analytical expressions for the limits to which these sums converge are used. For the first-order model in Eq. (3), these are given by

$$\sum_{k=a}^{\infty} \beta_{ih}(k) = \begin{cases} e'_i(A - B)^{-1}e_h & \text{if } a = 0, \\ e'_i((A - B)^{-1} - A^{-1})e_h & \text{if } a = 1, \end{cases} \quad (7)$$

where e_i is the four-dimensional vector that has zeros in all but the i th row, which contains unity (see Appendix A). Auto-regressive parameter estimates are then substituted into these expressions. Since the expressions are very simple, they are easier to work with than truncated sums of individual moving-average coefficients. As mentioned below, statistics for assessing their significance also take a simple form.

2.1.2. Assessing economic and statistical significance

The variance components in Eqs. (5) and (6) measure the sensitivity of conditional expectations to individual elements of u_t . In particular, they can be used to compare the relative importance of past innovations to price changes versus order flow. As mentioned above, their economic significance is assessed by how large they are as a *proportion* of the variance of $\Delta_{oi}(ct - 1, ct)$ and $\Delta_{oi-1}(ot - 1, ot)$. These variances measure the sensitivity of conditional expectations to *all* information that arrives during $(ct - 1, ct]$ and $(ot - 1, ot]$, respectively. Since the latter variances are likely to differ across securities, cross-sectional averages of ratios measure the relative importance of current and past price changes and order flow in a manner that abstracts from security-specific

³ Whether a particular approximation is good or not is an empirical question that is addressed briefly at the end of Section 4.2.

differences in the total sensitivity of conditional expectations to information arrivals.

Statistical significance of the impact of past information on permanent price changes is assessed for each security individually. Standardizing by the total variance of Δ is not necessary in this context, so the significance of the raw components in Eqs. (5) and (6) is evaluated. The strategy behind the tests is as follows. Since the (two-stage least squares) estimators of the time-series parameters are asymptotically normal, estimates of Eqs. (5) and (6) are also asymptotically normal because they are (non-linear but) smooth functions of the time-series parameter estimates. Using a consistent estimate of the asymptotic variance of a raw component, a statistic can be constructed that is asymptotically standard normal under the null that the raw component is truly zero. The gradients needed to compute such statistics take a simple form when exact expressions for the variance components are used. Detailed calculations are contained in Appendix B.

2.1.3. Generalizations

The estimates and inferences in this paper are based on a vector autoregression of order two. Table 2 reports the proportions of $\text{Var} [\Delta_{ot}(ct - 1, ct)]$ and $\text{Var} [\Delta_{ct-1}(ot - 1, ot)]$ that are related to each of the disturbances, and rejection rates of significance tests of the raw components, for this model. However, focusing on 24 h time periods, and their partitioning at the open and the close, are not necessary elements of the methods used here. For example, it is possible to look further into the past for evidence that endogenous market statistics are important determinants of beliefs. This is done by estimating quantities similar to those described above, but for revisions in expectations over 48 h periods for which ‘the past’ contains both a daytime and an overnight period (see Fig. 3). In accordance with the notation above, these expectational

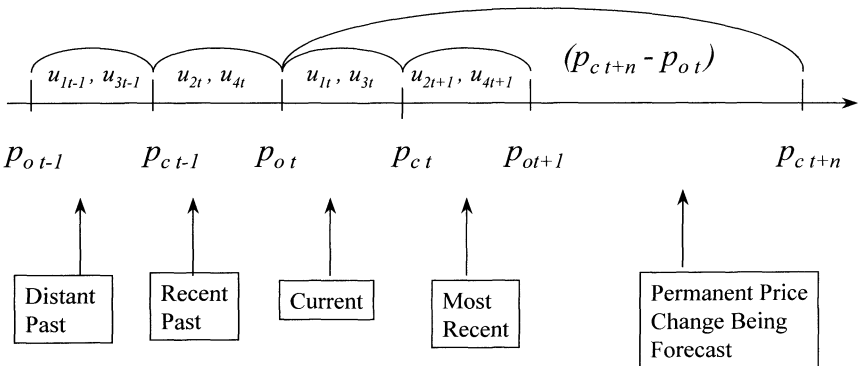


Fig. 3. Forecasts based on 48 h of current and past information.

revisions are defined as $\Delta_{ot}(ot - 1, ot + 1)$ and $\Delta_{ct-1}(ct - 2, ct)$. The components of their variances that relate to past and current shocks to returns and order flow are reported in Tables 3 and 4.

2.2. Decomposing the variance of order flow

If the updating of beliefs resulting from public information leads to trading, then not only will traders' beliefs be sensitive to past market statistics, but so will their trading strategies. In fact, the way Brown and Jennings (1989) define whether technical analysis has value is by whether past market statistics enter traders' demand functions. To examine this, the total variances of the order flow variables v_{nt} and v_{dt} are decomposed into components relating to each of the four shocks in u_t . The variance of forecasts is not decomposed, as was done in Section 2.1, because the present concern is not with beliefs about long-run changes in holdings; but whether holdings themselves change in response to past shocks to prices and order flow. The strategy for estimation and testing is the same as above, but the details of the calculations are different; so they are described briefly here.

From the infinite moving-average representations of $x_{3t} \equiv v_{dt}$ and $x_{4t} \equiv v_{nt}$ [see Eq. (A.2)], the component of the variance of v_{dt} attributable to u_{2t} , a shock to the previous overnight return, is

$$\left\{ \sum_{k=0}^{\infty} \beta_{32}(k)^2 \right\} \sigma_2. \quad (8)$$

Similarly, the component of the variance of v_{nt} attributable to u_{1t-1} , a shock to the previous daytime return, is

$$\left\{ \sum_{k=1}^{\infty} \beta_{41}(k)^2 \right\} \sigma_1. \quad (9)$$

These differ from Eqs. (5) and (6) in that the total variances here are functions of sums of squared moving average coefficients. Comparing components such as these, calculated for each disturbance, sheds light on the relative importance of contemporaneous versus past shocks to returns and order flow in explaining observed trading activity.

Components of the variance of order flow are estimated by substituting auto-regressive parameter estimates into exact expressions for the infinite sums of squared moving-average coefficients like those in Eqs. (8) and (9). Computational formulas for these sums are [see Hamilton (1994) (p. 265)]

$$\sum_{k=a}^{\infty} \beta_{ih}(k)^2 = \begin{cases} e'_i \mathcal{S}_h e_i & \text{if } a = 0, \\ e'_i (\mathcal{S}_h - A_h) e_i & \text{if } a = 1, \end{cases} \quad (10)$$

where,

$$\text{Vec}[\mathcal{S}_h] = [I - \Theta \otimes \Theta]^{-1} \text{Vec}[A_h], \quad A_h = A^{-1} e_h e_h' (A')^{-1} \quad \text{and} \quad \Theta = A^{-1} B,$$

and e_i is the four-dimensional vector that has zeros in all but the i th row, which contains unity.

Economic significance of the components of the variance of order flow is assessed by computing cross-sectional averages of the *proportion* of the variance of v_{dt} and v_{nt} explained by each component:

$$\frac{\{\sum_{k=0}^{\infty} \beta_{32}(k)^2\} \sigma_2}{\text{Var}[v_{dt}]} \quad \text{and} \quad \frac{\{\sum_{k=1}^{\infty} \beta_{41}(k)^2\} \sigma_1}{\text{Var}[v_{nt}]} \quad (11)$$

When averaged cross-sectionally, these ratios have a clearer interpretation than the raw components (e.g., Eqs. (8) and (9)) because they abstract from differences across securities in the total variance of daytime or overnight order flow. Statistical significance is assessed for each security individually, and the strategy behind the tests is the same as it was in connection with the raw components of permanent returns. However, the functional form of the components is different (see Appendix B).

3. Sample selection and data

The data for this study consist of the quotations, transaction prices and order quantities obtained from the Institute for the Study of Security Markets (ISSM) files for 1986–1989. The sample is selected by first ranking all NYSE stocks that appear in the files by their average daily trading volume (in dollars) over all four years, then selecting the top fifty stocks in each quartile that are not traded in Britain or Tokyo.⁴ Stocks are classified into groups based on dollar trading volume in order to control for differences in the speed of price responses to information that might exist between stocks that are highly visible and those that are less well-known. For each stock, the time series of daytime and overnight returns are calculated using the midpoints of the first and last NYSE quotes of each day adjusted for cash distributions, stock dividends and splits. Order flow at the opening and during the rest of the day are computed using the volume (in shares) of individual transactions, and the algorithm suggested by Lee and Ready (1991) to determine whether each order is a purchase or a sale (see Appendix C). Order sizes are re-scaled so that stock dividends and splits do not affect the order flow measure.

⁴ See Forster and George (1995, 1996) for comparisons of cross-listed and non-cross-listed stocks.

Although they are computed in a similar manner, opening and daytime order flow variables have different economic interpretations. Daytime order flow is computed by cumulating the signed individual transactions that occur throughout the day, excluding (including) the first transaction if the stock opens with a trade (the posting of quotes). Assuming that the signing algorithm is correct, this variable is a precise measure of net trading activity. Opening order flow is computed similarly. If the stock opens with a trade, it is signed using the algorithm described in Appendix C; if it opens with the posting of quotes, opening order flow is zero. For actively traded stocks, opening transactions are likely to involve the crossing of several orders, the extent of which cannot be estimated. Consequently, this variable is not necessarily a measure of net order flow at the opening, but a signed measure of total trading at the opening for active stocks. Since opening and daytime order flow are treated as separate variables, the difference in their construction does not create stationarity problems that would occur if an alternating sequence of these variables were treated as a single order flow series. Moreover, we believe that these variables correspond to the market statistics that are readily observable to traders, and for that reason are well-suited to this study.

Descriptive statistics are contained in Table 1. The equity capitalization of stocks in each quartile is approximately four times that of stocks in the next (lowest) quartile of dollar trading volume. This is similar to the sample that would have resulted from initially ranking securities based on equity capitalization rather than dollar volume. Return volatility is similar across quartiles; but varies across stocks within quartile 3 by about twice as much as within the other quartiles. The standard deviation of daytime order flow declines approximately by a factor of two across quartiles of lesser trading activity. The same is true of opening order flow across quartiles 2–4; but the standard deviation of the opening order flow of quartile 1 is more than three times that of quartile 2. This could be due to the fact that our measure of opening order flow for active stocks is more likely to reflect the aggregation of several orders than it is for less active stocks. The last two rows of Table 1 indicate that the average absolute value of daytime order flow is between twelve and fifteen times as large as that of opening order flow for stocks in quartiles 2–4; but only slightly above seven for stocks in quartile 1.

4. Model specification and empirical results

4.1. Specification

Results reported in the tables are based on a four-variate VAR(2) involving daytime and overnight returns, and daytime and overnight order flow. This specification is defined to include two lags of each of the four variables, and to

Table 1
Descriptive statistics

The statistic named in the first column is computed using data from 1986–1989 for each security in the sample. The cross-sectional means and standard deviations are reported by quartile category. Each quartile category contains the top 50 stocks in each quartile of dollar trading volume for NYSE stocks during the 1986–1989 period. Equity capitalization is computed as of 31 December, 1987. Order-flow variables are reported in 100-share units

	Quartile 1		Quartile 2		Quartile 3		Quartile 4	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Equity capitalization (in millions)	4942.84	3544.86	1136.22	570.15	389.869	346.208	151.563	206.869
Average daytime return (%)	0.03	0.07	0.03	0.07	0.00	0.16	0.00	0.09
Std. Dev. of daytime return (%)	2.05	0.99	1.90	0.59	2.59	1.69	2.05	0.83
Average overnight return (%)	-0.01	0.07	-0.03	0.06	-0.02	0.09	-0.03	0.07
Std. Dev. of overnight return (%)	1.97	1.26	1.54	0.86	1.99	1.56	1.51	0.75
Average daytime order flow	176.485	223.873	23.241	79.411	-9.678	39.188	-8.607	27.950
Std. Dev. of daytime order flow	2345.902	1474.793	1364.880	856.061	729.316	843.962	386.184	584.879
Average opening order flow	21.219	19.683	3.003	3.548	1.097	2.846	0.250	1.265
Std. Dev. of opening order flow	301.907	143.351	84.720	42.501	42.208	49.037	27.412	55.810
Average absolute value of daytime order flow	896.630	1466.111	355.484	877.617	166.038	848.707	79.786	585.331
Average absolute value of opening order flow	123.720	131.114	28.589	40.779	10.980	48.740	6.087	55.235

respect the dependence among the variables that occurs because observations with the same time subscript are ordered in clock time. The same model is estimated for all the stocks in the sample to facilitate cross-sectional aggregation and comparisons. This specification is relatively parsimonious, yet its lag structure should be sufficient to capture the effects of micro-structure frictions on returns.

In accordance with the notation of Section 2, the model is

$$A\mathbf{x}_t = B\mathbf{x}_{t-1} + C\mathbf{x}_{t-2} + \mathbf{u}_t, \quad (12)$$

where $\mathbf{x}_t = (r_{dt}, r_{nt}, v_{dt}, v_{nt})'$, $\mathbf{u}_t = (u_{1t}, u_{2t}, u_{3t}, u_{4t})'$ is a vector of random disturbances assumed to be iid, B is an arbitrary 4×4 matrix, and

$$A = \begin{bmatrix} 1 & a_{12} & a_{13} & a_{14} \\ 0 & 1 & 0 & a_{24} \\ a_{31} & a_{32} & 1 & a_{34} \\ 0 & a_{42} & 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} c_{11} & 0 & 0 & 0 \\ c_{21} & c_{22} & c_{23} & 0 \\ 0 & 0 & c_{33} & 0 \\ c_{41} & 0 & c_{43} & c_{44} \end{bmatrix}.$$

The matrix C is defined so that two lags of each variable appear in each equation.⁵ The matrix A is defined to allow daytime returns and order flow to depend on each other and on previous overnight returns and opening order flow; but overnight returns do not depend on the daytime returns and order flow that succeed them.⁶ Contemporaneous correlation is captured by parameters in the A matrix; so in estimation and hypothesis testing, the covariance matrix of the disturbances is assumed to be diagonal.

Following Gallant et al. (1992), we account for mean return and order-flow effects associated with day-of-the-week, turn-of-the-year and the October 1987 crash by estimating regressions of each variable on an intercept and indicators for each of Tuesday through Friday, an indicator for whether the date is between December 15 and January 15, and an indicator for October 1987. The residuals from these regressions (estimated separately for each security) are used in

⁵ This is selected in favor of a first-order model (as used in Campbell and Ammer, 1993, for example) because spot checks indicate that significant second-order relations exist for some securities. Even if VAR(1) is the correct specification for some securities, inferences will not be affected because the test statistics are constructed in a manner that accounts for the precision with which the VAR parameters are estimated. Parameter estimates that are insignificantly different from zero receive little weight in testing the significance of variance components.

⁶ This is crucial because ignoring the temporal ordering of variables with the same time subscript would lead us to systematically mis-classify the source of variation in forecast revisions. For example, if the dependence of daytime returns on (preceding) overnight returns were ignored, the effect of the overnight return shock on the revision in beliefs about price changes after the opening would be understated.

estimating the vector auto-regressions. Eq. (12) is estimated using two-stage least squares separately for each of the two-hundred securities in the sample.⁷

In principle, parameter estimates should be obtained jointly for all securities in a manner that accounts for the correlation of \mathbf{u}_t vectors across securities; but the number of parameters involved in such an exercise with just a few securities exceeds what can feasibly be estimated with available computing resources. Since the significance tests at the individual security level are not likely to be independent, both rejection rates and the average value of the test statistic are reported in the tables below. The average test statistic conveys a sense of whether the (potentially correlated) rejections are strong or weak.

4.2. Empirical results

4.2.1. Permanent changes in prices

Table 2 presents the results of the variance decomposition of the permanent component of returns described in Section 2.1. Each security's estimated VAR is checked for invertibility before computing variance decompositions. The VARs of all but three securities in the sample are invertible. Each of the cross-sectional averages that are reported in the tables by quartile are based on at least 48 components estimates. Each panel contains within-quartile cross-sectional means of the standard deviations of $\Delta_{ot}(ct - 1, ct)$ and $\Delta_{ct-1}(ot - 1, ot)$, and the proportions of the variance of each Δ attributable to each of the four elements of the disturbance vector, \mathbf{u}_t . The cross-sectional means of the standard deviations of $\Delta_{ot}(ct - 1, ct)$ and $\Delta_{ct-1}(ot - 1, ot)$ are between 2.5% and 14.2%.⁸ These measure the extent to which forecasts of permanent price changes are revised based on information that arrives within a 24 h period. The estimates suggest that non-trivial revisions in conditional expectations occur over such periods.

The forecast revisions $\Delta_{ot}(ct - 1, ct)$ and $\Delta_{ct-1}(ot - 1, ot)$ are defined so that part of the information causing revisions in beliefs is available in the recent past, and the rest is available only at present. The decomposition of the variance of $\Delta_{ot}(ct - 1, ct)$ for stocks in quartile 1, for example, indicates that on average

⁷ By the way opening order flow is defined, the posting of opening quotes follows the realization of opening order flow in clock time. So although the overnight return and opening order flow both reflect events of the period subsequent to the prior closing, opening order flow is observable to liquidity providers at the time that opening quotes are posted. Consequently, in estimation, the overnight return is treated as an endogenous variable in estimating the equation for opening order flow; but opening order flow is treated as a pre-determined variable in estimating the equation for overnight returns. The same treatment is applied to daytime returns and order flow because even if the stock closes with a transaction, the last quotes are posted after almost all of daytime order flow has been realized.

⁸ The large mean of quartile 4 is due to a single security whose standard deviation estimates are in excess of 500%. The averages without this security are 3.7% and 2.6%.

4.46% of the variation in the revision in agents' beliefs about changes in security value *after* the open of trading is associated with the shocks to opening order flow; and 10.45% relates to the shock to the price change during the previous overnight period. These estimates imply that, on average, nearly 15% of the variation in estimates of agents' beliefs about security value is associated with past market statistics; the proportions for quartiles 2–4 are 12.55%, 13.16% and 7.55%, respectively. The remainder of the variation is associated with current shocks to security returns and order flow.⁹

Also reported in Table 2 are cross-sectional averages of z -statistics computed for individual securities; and the percentage of securities in each quartile for which the z test rejects the null hypothesis that the variance component is equal to zero. These numbers indicate that past market statistics are statistically significant determinants of belief revisions for between 74% and 83% of securities in quartiles 1–3; rejection rates are 66% and 60% for quartile 4. By way of comparison, rejection rates for current information are between 84% and 98% for quartiles 1–3; and 78% and 86% for quartile 4. These inferences are robust to controlling for possible correlation in the test statistics across securities. The average z -statistics for past innovations all exceed six, which implies that the vector of z -statistics for each quartile falls well outside the Bonferroni bound of 3.5 for a parameter vector of dimension fifty at the 1% significance level.¹⁰

The implication of these estimates is that, on average, opening prices do not fully reflect information in the trading history. Rather, past market statistics convey information that is incremental to the current price for the purpose of forecasting changes in the security's value. This effect is stronger for more active than less active stocks, and is significant in both economic and statistical terms for stocks in quartiles 1–3. These findings are consistent with the predictions of models of technical analysis in which agents condition their beliefs on past market statistics in arriving at their assessment of the security's true value. The magnitudes of the estimates suggest that past innovations in price changes are

⁹ The discussion in the text focuses on means to highlight the relative contributions of the variance components. Medians do not account for all the variation because they need not add to one across components. However, medians are reported because they convey a sense of how dependent is the forecast revision of the 'typical' stock in a quartile to an *individual* source of variation. For instance, forecast revisions after the opening for the typical stock in quartile 1 are insensitive to unexpected opening order flow (0.434%), barely sensitive to the unexpected overnight return (2.471%) and moderately sensitive to current daytime order flow (14.4%). Looking across quartiles, the typical security in each quartile is less sensitive to these three components than are securities on average. This is not entirely a reflection of the fact that these components are bounded below by zero. The typical security is *more* sensitive to the current-day unexpected return than are securities on average.

¹⁰ As a sensitivity check, we also conducted inferences using heteroskedasticity-consistent standard errors for quartiles 1 and 4; and rejection rates varied only slightly from those reported in the table.

somewhat more important than past innovations in order flow. However, the differences are small enough that a clear case does not emerge for whether models based on past volume or past prices are more realistic.

Table 2 also contains estimates of the proportion of the variance of $\Delta_{ct-1}(ot - 1, ot)$ attributable to each disturbance. In this decomposition, forecasts relate to price changes after the closing, and unexpected *daytime* returns and order flow are the information in the trading history. The estimates imply that, on average, 8.09%, 9.85%, 7.57% and 13.2% of the variation in estimates of agents' beliefs about the value of stocks in quartiles 1–4 is associated with past market statistics. These proportions are statistically significant for 74% or more of the securities in each of quartiles 1–3; with the lowest rejection rate at 64% for past daytime order flow in quartile 4. For all quartiles, the statistical significance of these estimates is similar to those above. Economic significance is less than above for all but quartile 4.

These results indicate that the closing prices of stocks in quartiles 1–3 reflect information in the trading history more fully than opening prices on average. Only quartile 4 stocks exhibit a proportional forecast revision variance that exceeds 10%, on average. The differences across quartiles suggest that information generated by the trading process is better incorporated into closing prices for stocks that trade more actively than stocks that trade less actively. This is consistent with the hypothesis that trading facilitates price discovery, a view advanced by Madhavan (1992) and Leach and Madhavan (1993). The nature of our findings is similar to evidence presented in Madhavan et al. (1997). Their estimates imply that the volatility of the component of returns that is not related to market frictions decreases throughout the day; and the permanent impact of orders on prices decreases more for stocks of large than small firms.

The results in Table 2 also show that order flow is a significant source of value-relevant information. Daytime order flow accounts for between 14.9% and 22.7% of the variance of forecast revisions occurring after the opening, and is statistically significant for 78% or more of the securities in the sample. The measure of opening order flow is less precise than daytime order flow because opening orders are aggregated, but still accounts for between 24.1% and 34.8% of the variance of forecast revisions occurring after the close. This supports Hasbrouck (1991) finding that information about trading is an important source of variation in the permanent component of intra-day revisions of NYSE quotes. He interprets trade-related variation as the incorporation of private information into prices. Using his interpretation, the results in Table 2 suggest that as much as one-third of the value-relevant information that arrives over the course of an average trading day could be private.

Table 3 reports the variance decomposition of $\Delta_{ot}(ot - 1, ot + 1)$, which measures the revision in expectations between the open of trading on day $t - 1$ and the open on day $t + 1$ concerning long-horizon returns to occur after the open on day t . This definition of Δ reflects information that arrives in the 24 h

Table 3
Variance decomposition of 48 h revisions in forecasts of price changes occurring after the opening

Parameters of the VAR model are estimated for each security using two-stage least squares from daily data over the period 1986–1989. Variance decompositions are computed for each security individually. The numbers reported in the table are cross-sectional means and medians of the proportional variance components; and cross-sectional means of z -statistics and rejection rates ($\alpha = 1\%$) relating to security-by-security tests of the hypothesis that a variance component is zero. Δ_o and Δ_c are revisions of forecasts of price changes occurring after the opening and closing, respectively. u_1 and u_2 are unexpected daytime and overnight returns; u_3 and u_4 are unexpected order flow during the day (exclusive of the opening) and at the opening, respectively. Each quartile category contains the top fifty stocks in each quartile of dollar trading volume for NYSE stocks during the 1986–1989 period. Columns in boldface type relate to the most current information (u_{2t+1} and u_{4t+1}), and information in the distant past (those labeled u_{1t-1} and u_{3t-1})

		Percent of $\text{Var}[\Delta_o(\text{ot} - 1, \text{ot} + 1)]$ attributable to							
		u_{1t-1}	u_{2t}	u_{3t-1}	u_{4t}	u_{1t}	u_{2t+1}	u_{3t}	u_{4t+1}
Quartile 1	Means	0.778	6.356	0.923	1.165	37.815	28.079	12.148	12.736
	Medians	0.095	1.088	0.172	0.257	37.384	21.259	7.114	7.426
	Rejection rates	76%	80%	71%	76%	94%	96%	94%	96%
	Average z -value	6.061	34.435	6.908	8.480	60.779	39.498	33.097	30.949
Quartile 2	Means	1.772	2.229	1.300	1.527	37.666	33.896	9.341	12.269
	Medians	0.441	0.651	0.076	0.203	36.740	28.106	3.923	6.711
	Rejection rates	70%	74%	56%	80%	96%	84%	84%	86%
	Average z -value	8.751	10.977	5.557	9.019	69.936	22.813	24.701	20.933
Quartile 3	Means	0.614	5.684	0.579	1.935	46.895	26.383	9.351	8.559
	Medians	0.135	0.687	0.052	0.656	49.463	17.930	5.234	3.723
	Rejection rates	71%	77%	71%	83%	98%	94%	94%	92%
	Average z -value	7.143	26.698	6.777	10.991	77.419	34.989	23.039	17.675
Quartile 4	Means	1.384	1.882	1.271	1.553	44.361	22.242	13.725	13.582
	Medians	0.340	0.554	0.195	0.278	44.075	15.343	7.431	6.713
	Rejection rates	54%	66%	54%	60%	86%	80%	78%	72%
	Average z -value	4.974	9.478	5.438	6.167	54.111	19.051	25.837	14.078

periods preceding and following the opening, rather than just the previous overnight and the subsequent daytime periods captured by $\Delta_{or}(ct - 1, ct)$. These decompositions are used to examine the degree to which variation in forecast revisions is attributable to information that arrives in both the recent and distant past.

The boldface columns on the left-hand side of Table 3 show that shocks to returns and order flow from the preceding day (i.e., the distant past) explain, on average, 3.2% or less of the variance of forecast revisions (with smaller rejection rates for the larger estimates). This compares with percentages between 3.3% and 7.5% for information in the preceding night (i.e., recent past). This means that the earlier finding that opening prices do not reflect fully information in the trading history is limited to information in the *recent* past.

Similar conclusions emerge from the decomposition of the variance of $\Delta_{ct-1}(ct - 2, ct)$ reported in Table 4. The boldface columns on the left side of the table indicate that, on average, 2.9% or less of the variance of forecast revisions of price changes to occur after the close are explained by shocks to returns and order flow from the previous overnight period. The results in Table 2 indicate that opening prices do not fully reflect information from the previous overnight, on average, for stocks in quartiles 1–3. The results in Table 4 indicate that this information is incorporated by the close of trading. Taken as a whole, the results in Tables 3 and 4 suggest that when opportunities to benefit from technical analysis exist at the individual security level, they are short-lived.

These tables also show that order flow is an economically and statistically significant source of value-relevant information for securities in all quartiles. Between 17% and 28% of the variance in forecast revisions is attributable to current shocks to order flow, with similar percentages for daytime and opening order flow. This suggests that a large portion of permanent changes in prices occurs through trading on private information both at the open and during the day.

Although not reported in tables, we checked to see how sensitive the proportions are to estimating components using truncated sums of moving-average coefficients. The mean differences are small for five- and ten-lag truncations. Therefore, it probably makes little difference whether truncated or exact sums are used to compute cross-sectional averages of variance components. However, exact sums must be used for hypothesis testing because the statistics described here are not appropriate for assessing the significance of truncated sums. Exact sums also should be used if one is interested in explaining cross-sectional variation in the estimates. In some individual cases, the magnitudes of the errors due to truncation were large.

4.2.2. Order flow

The results so far indicate that opening prices do not fully reflect information from the previous night, but closing prices fully reflect information from the

Table 4

Variance decomposition of 48 h revisions in forecasts of price changes occurring after the closing

Parameters of the VAR model are estimated for each security using two-stage least squares from daily data over the period 1986–1989. Variance decompositions are computed for each security individually. The numbers reported in the table are cross-sectional means and medians of the proportional variance components; and cross-sectional means of z -statistics and rejection rates ($\alpha = 1\%$) relating to security-by-security tests of the hypothesis that a variance component is zero. Δ_t and Δ_c are revisions of forecasts of price changes occurring after the opening and closing, respectively. u_1 and u_2 are unexpected daytime and overnight returns; u_3 and u_4 are unexpected order flow during the day (exclusive of the opening) and at the opening, respectively. Each quartile category contains the top fifty stocks in each quartile of dollar trading volume for NYSE stocks during the 1986–1989 period. Columns in boldface type relate to the most current information (u_{1t} and u_{3t}), and information in the distant past (those labeled u_{2t-1} and u_{4t-1})

	Std. Dev. of $\Delta_{t-1}(ct - 2, ct)$	Percent of $\text{Var}[\Delta_{t-1}(ct - 2, ct)]$ attributable to																	
		u_{1t-1}	u_{2t-1}	u_{3t-1}	u_{4t-1}	u_{1t}	u_{2t}	u_{3t}	u_{4t}										
Quartile 1																			
Means	0.043	0.936	1.669	0.983	0.882	40.459	29.175	12.561	13.334										
Medians	0.029	0.334	0.220	0.242	0.146	42.946	24.350	7.904	8.040										
Rejection rates	n.a.	80%	84%	80%	80%	94%	96%	94%	94%										
Average z -value	n.a.	11.992	13.824	11.289	7.492	60.779	39.498	33.097	30.619										
Quartile 2																			
Means	0.041	2.040	0.942	1.368	1.128	38.425	34.427	9.217	12.452										
Medians	0.028	0.546	0.187	0.064	0.111	39.195	28.951	4.045	6.589										
Rejection rates	n.a.	80%	70%	74%	74%	96%	84%	84%	86%										
Average z -value	n.a.	16.088	8.446	8.460	7.145	69.936	22.813	24.701	20.933										
Quartile 3																			
Means	0.044	0.826	1.381	0.714	1.316	50.466	26.939	9.456	8.901										
Medians	0.030	0.227	0.277	0.081	0.144	53.367	18.664	5.254	3.851										
Rejection rates	n.a.	85%	77%	81%	81%	98%	94%	94%	92%										
Average z -value	n.a.	12.404	10.083	9.819	9.331	77.419	34.989	23.039	17.675										
Quartile 4																			
Means	0.202	1.521	1.382	1.417	1.537	44.834	21.965	13.814	13.530										
Medians	0.030	0.518	0.183	0.162	0.230	44.358	15.667	7.566	6.788										
Rejection rates	n.a.	74%	64%	64%	58%	86%	80%	78%	72%										
Average z -value	n.a.	11.499	5.166	7.889	6.836	54.111	19.051	25.837	14.078										

previous day's trading; and that this characterization is true to a greater degree for more active stocks than less active stocks. If trading does indeed facilitate the price discovery process, then trading that occurs during the day should help prices impound information in the trading history that is not fully reflected in opening prices. This means that some of the variation in daytime order flow should be related to shocks to the overnight return and opening order flow; and that this relation should be stronger for stocks in the quartiles of greater trading activity. Also, opening order flow should not be strongly related to information in the trading history because closing prices appear to reflect fully this information (except, perhaps, stocks in quartile 4).

The impact of past market statistics on trading patterns is examined in Table 5, which reports variance decompositions of the daytime and overnight order flow variables. Also reported are cross-sectional averages of z -statistics and rejection rates associated with z tests on individual securities.

The results are mostly consistent with these conjectures, but economic significance is not strong. Between 3.6% and 7.0% of daytime order flow is related to shocks to returns and order flow from the previous night; and these proportions are greater, on average, for stocks in quartile 1 than for stocks in quartiles 2 and 3. The average for quartile 4 is the largest, despite the earlier finding that opening prices of those stocks reflect information in the trading history more fully than stocks in the other quartiles. In all but one instance, rejection rates are 74% and above for tests of the hypothesis that the variance component is zero. Between 0.7% and 2.9% of opening order flow is related to shocks to returns and order flow from the previous daytime, on average, for stocks in quartiles 1–3. As expected, the estimates for quartile 4 are larger; on average, 7.5% of the variance of opening order flow is attributable to unexpected returns and order flow from the previous day. Statistical significance is weak, which could reflect the fact that our estimate of opening order flow is imprecise because orders are aggregated at the opening.

5. Conclusion

This study estimates the sensitivity of forecasts of permanent security price changes to current and past market statistics. Forecasts of price changes after the open of trading are related to price changes and order flow in the recent, but not distant, past; the relation is stronger for more actively traded stocks. By contrast, forecasts of price changes after the close are not significantly related to past market statistics; except for the least-actively traded stocks in the sample. These findings support the predictions of models of technical analysis such as Brown and Jennings (1989), Grundy and McNichols (1989), and Blume et al. (1994), which hypothesize that current prices do not always fully reflect value-relevant information in the trading history.

Table 5
Decomposition of variance of order flow variables

Parameters of the VAR model are estimated for each security using two-stage least squares from daily data over the period 1986–1989. Variance decompositions are computed for each security individually. The numbers reported in the table are cross-sectional means and medians of the proportional variance components; and cross-sectional means of z -statistics and rejection rates ($\alpha = 1\%$) relating to security-by-security tests of the hypothesis that a variance component is zero. v_{dt} and v_{nt} are net order flow in hundreds of shares during the day (exclusive of the opening) and net order flow at the opening, respectively. u_1 and u_2 are unexpected daytime and overnight returns; u_3 and u_4 unexpected order flow during the day (exclusive of the opening) and at the opening, respectively. Each quartile category contains the top fifty stocks in each quartile of dollar trading volume for NYSE stocks during the 1986–1989 period. Columns in boldface type relate to information available in the past

	Std. Dev. of v_{dt}				Std. Dev. of v_{nt}			
	u_1	u_2	u_3	u_4	u_1	u_2	u_3	u_4
Quartile 1								
Means	2331.36	3.009	73.539	3.403	299.96	1.332	28.667	1.262
Medians	2015.36	0.458	72.212	0.598	265.88	0.186	32.155	0.133
Rejection rates	n.a.	86%	96%	90%	n.a.	78%	88%	67%
Average z -value	n.a.	11.270	95.695	13.656	n.a.	12.152	10.462	6.823
Quartile 2								
Means	1358.56	1.893	72.816	1.686	84.24	1.659	30.722	1.255
Medians	1273.82	0.345	70.762	0.297	76.26	0.100	32.731	0.089
Rejection rates	n.a.	74%	88%	80%	n.a.	66%	70%	50%
Average z -value	n.a.	10.266	147.754	10.713	n.a.	7.417	6.582	4.403
Quartile 3								
Means	726.97	1.929	74.512	1.921	41.98	0.482	28.265	0.257
Medians	313.20	0.268	78.607	0.249	30.52	0.215	33.122	0.072
Rejection rates	n.a.	85%	96%	85%	n.a.	73%	85%	50%
Average z -value	n.a.	9.130	175.570	10.056	n.a.	8.107	8.181	2.990
Quartile 4								
Means	384.66	3.575	68.725	3.443	27.28	4.165	29.203	3.369
Medians	137.84	0.454	66.135	0.621	15.17	0.255	32.708	0.096
Rejection rates	n.a.	76%	84%	66%	n.a.	66%	82%	48%
Average z -value	n.a.	7.519	224.545	8.918	n.a.	9.882	6.512	4.236

The manner in which the results differ between open and close and across stocks grouped by trading activity suggests that trading facilitates the incorporation of past information into prices. A similar conclusion is reached by other researchers who examine return volatility, and bid-ask spreads throughout the day such as Amihud and Mendelson (1991), Hasbrouck (1991, 1993), Chan et al. (1994), and Madhavan et al. (1997). The results in this paper complement theirs by focusing specifically on information in the trading history, and the rate at which it becomes incorporated into prices through trading.

The theories of technical analysis are set in single-risky-asset economies, and the empirical work in this paper treats individual securities as independent because joint estimation for multiple securities is not practical. However, cross-security or market-wide effects might also be important to forecasts of long-run price changes. For example, cross-serial correlation exists in raw returns (see Lo and MacKinlay, 1990; Conrad et al., 1991), and returns within size deciles exhibit forecastable systematic components (see Conrad and Kaul, 1988). These effects, which are ignored here by focusing exclusively on security-specific information in the trading history, might make this paper's assessment of the models of technical analysis conservative.

The tests in this paper for whether technical analysis has value are designed to conform to the spirit of models in which technical analysis is a rational phenomenon. Nevertheless, the tests cannot distinguish between rationality and irrationality as explanations for *why* technical analysis might have value. The results indicate that even rational (linear) forecasts of long-run price changes are sensitive to past events. This could occur because noise in the system prevents even a good Bayesian from fully understanding past events until more data is observed, or because market participants under- or over-react to observations whose interpretation should be unambiguous, or some combination of these effects. A careful laboratory experiment might be able to assess the relative importance of these effects on the behavior of individual market participants.

Acknowledgements

We thank Joyce Berg, Doug Foster, Yasushi Hamao, Joel Hasbrouck, Gautam Kaul, Ananth Madhavan, Tom Rietz, Paul Schultz, Avaniidhar Subrahmanyam (the editor), Ben Wilner, an anonymous referee and seminar participants at Hong Kong University of Science and Technology, Michigan State University, Northwestern University, the University of Iowa, the University of Maryland, and the University of Notre Dame for helpful comments.

Appendix A. Derivation of equations

Derivation of Eq. (4): Assuming that A^{-1} exists, Eq. (3) in the text can be written as

$$\mathbf{x}_t = \Theta \mathbf{x}_{t-1} + A^{-1} \mathbf{u}_t \quad \text{where} \quad \Theta = A^{-1}B,$$

or as

$$\mathbf{x}_t = \sum_{k=0}^{\infty} \Theta^k A^{-1} \mathbf{u}_{t-k} \quad (\text{A.1})$$

provided that the eigenvalues of Θ are inside the unit circle. If we define $\beta_{ih}(k)$ to be the (i, h) th element of $\Theta^k A^{-1}$, then the i th equation in Eq. (A.1) can be written as

$$x_{it} = \sum_{k=0}^{\infty} \beta_{i1}(k) u_{1t-k} + \sum_{k=0}^{\infty} \beta_{i2}(k) u_{2t-k} + \sum_{k=0}^{\infty} \beta_{i3}(k) u_{3t-k} + \sum_{k=0}^{\infty} \beta_{i4}(k) u_{4t-k}, \quad (\text{A.2})$$

the sum of four infinite moving averages. We can write $p_{ct+n} - p_{ot}$ as a sum of returns over the period from ot to $ct+n$, and $p_{ot+n} - p_{ct-1}$ as a sum of returns over the period from $ct-1$ to $ot+n$. Recalling that $x_{1t} \equiv r_{dt}$ and $x_{2t} \equiv r_{nt}$, these sums are given by

$$\begin{aligned} p_{ct+n} - p_{ot} &= x_{1t} + \sum_{j=1}^n (x_{1t+j} + x_{2t+j}), \\ p_{ot+n} - p_{ct-1} &= \sum_{j=0}^n (x_{1t+j} + x_{2t+j}). \end{aligned} \quad (\text{A.3})$$

Consider the case in which $\mathbf{u}_t = (u_{1t}, u_{2t})'$. Eq. (A.3) can be written as

$$p_{ct+n} - p_{ot} = \sum_{j=0}^n x_{1t+j} + \sum_{j=1}^n x_{2t+j},$$

implying that

$$\begin{aligned} &E_{ct}[p_{ct+n} - p_{ot}] - E_{ct-1}[p_{ct+n} - p_{ot}] \\ &= \left\{ E_{ct} \left[\sum_{j=0}^n x_{1t+j} \right] - E_{ct-1} \left[\sum_{j=0}^n x_{1t+j} \right] \right\} + \left\{ E_{ct} \left[\sum_{j=1}^n x_{2t+j} \right] \right. \\ &\quad \left. - E_{ct-1} \left[\sum_{j=1}^n x_{2t+j} \right] \right\}. \end{aligned}$$

This can be written more compactly as

$$\Delta E_{ct-1}^{ct} [p_{ct+n} - p_{ot}] = \sum_{j=0}^n \Delta E_{ct-1}^{ct} [x_{1t+j}] + \sum_{j=1}^n \Delta E_{ct-1}^{ct} [x_{2t+j}]$$

using the notation, $\Delta E_{ct-1}^{ct}[\cdot] \equiv E_{ct}[\cdot] - E_{ct-1}[\cdot]$. Since $\mathbf{u}_t = (u_{1t}, u_{2t})'$, the infinite MA representations of x_{2t} , x_{1t+j} and x_{2t+j} are

$$\begin{aligned} x_{2t} &= \sum_{k=0}^{\infty} \beta_{21}(k)u_{1t-k} + \sum_{k=0}^{\infty} \beta_{22}(k)u_{2t-k}, \\ x_{1t+j} &= \sum_{k=0}^{\infty} \beta_{11}(k)u_{1t+j-k} + \sum_{k=0}^{\infty} \beta_{12}(k)u_{2t+j-k}, \\ x_{2t+j} &= \sum_{k=0}^{\infty} \beta_{21}(k)u_{1t+j-k} + \sum_{k=0}^{\infty} \beta_{22}(k)u_{2t+j-k}. \end{aligned} \tag{A.4}$$

Note that expectations change only with realizations of shocks. Between $ct - 1$ and ct , the shocks u_{1t} and u_{2t} are realized. Consequently,

$$\Delta E_{ct-1}^{ct}[u_{ht-s}] = \begin{cases} u_{ht} & \text{if } s = 0 \\ 0 & \text{otherwise} \end{cases} \quad \text{for } h \in \{1, 2\}.$$

Therefore,

$$\begin{aligned} \Delta E_{ct-1}^{ct}[x_{2t}] &= \beta_{21}(0)u_{1t} + \beta_{22}(0)u_{2t}, \\ \Delta E_{ct-1}^{ct}[x_{1t+j}] &= \beta_{11}(j)u_{1t} + \beta_{12}(j)u_{2t}, \\ \Delta E_{ct-1}^{ct}[x_{2t+j}] &= \beta_{21}(j)u_{1t} + \beta_{22}(j)u_{2t}. \end{aligned}$$

Substituting these expressions into the definition of $\Delta_{ot}(ct - 1, ct)$ yields

$$\begin{aligned} \Delta_{ot}(ct - 1, ct) &\equiv \lim_{n \rightarrow \infty} \{ \Delta E_{ct-1}^{ct}[p_{ct+n} - p_{ot}] \} \\ &= \lim_{n \rightarrow \infty} \sum_{j=0}^n \{ \beta_{11}(j)u_{1t} + \beta_{12}(j)u_{2t} \} + \lim_{n \rightarrow \infty} \sum_{j=1}^n \{ \beta_{21}(j)u_{1t} \\ &\quad + \beta_{22}(j)u_{2t} \} \\ &= \left\{ \sum_{j=0}^{\infty} \beta_{11}(j)u_{1t} + \sum_{j=0}^{\infty} \beta_{12}(j)u_{2t} \right\} + \left\{ \sum_{j=1}^{\infty} \beta_{21}(j)u_{1t} \right. \\ &\quad \left. + \sum_{j=1}^{\infty} \beta_{22}(j)u_{2t} \right\} \end{aligned} \tag{A.5}$$

which is the first statement in Lemma 1 for the special case in which $\mathbf{u}_t = (u_{1t}, u_{2t})'$. By the way \mathbf{x}_t is defined in the general case, the shocks u_{1t} and u_{3t} are contemporaneous, as are the shocks u_{2t} and u_{4t} . Consequently, the order-flow shocks u_{3t} and u_{4t} enter the functional forms of Δ_{ot} in the same manner as their contemporaneous return-shock counterparts. Generalizing Eq. (A.5) to reflect shocks to order flow yields Eq. (4) in the text.

Derivation of Eq. (7): Define the matrix

$$\psi_{\beta}(k) \equiv \begin{bmatrix} \beta_{11}(k) & \cdots & \beta_{14}(k) \\ \vdots & \ddots & \vdots \\ \beta_{41}(k) & \cdots & \beta_{44}(k) \end{bmatrix}.$$

Using this definition, $\beta_{ih}(k) = e'_i \psi_{\beta}(k) e_h$, and

$$\sum_{k=a}^{\infty} \beta_{ih}(k) = e'_i \left(\sum_{k=a}^{\infty} \psi_{\beta}(k) \right) e_h.$$

By definition, $\beta_{ih}(k)$ is the (i, h) th element of $\Theta^k A^{-1}$, so $\psi_{\beta}(k) = \Theta^k A^{-1}$. Therefore, if it exists,

$$\Psi_{\beta} \equiv \sum_{k=0}^{\infty} \psi_{\beta}(k) = \sum_{k=0}^{\infty} \Theta^k A^{-1} = (I - \Theta)^{-1} A^{-1}$$

and

$$\Psi_{\beta} - \psi_{\beta}(0) = (I - \Theta)^{-1} A^{-1} - A^{-1}.$$

Using the definition $\Theta = A^{-1}B$, yields

$$\sum_{k=a}^{\infty} \beta_{ih}(k) = \begin{cases} e'_i (I - A^{-1}B)^{-1} A^{-1} e_h & \text{if } a = 0, \\ e'_i ((I - A^{-1}B)^{-1} A^{-1} - A^{-1}) e_h & \text{if } a = 1, \end{cases}$$

which is equal to Eq. (7) in the text.

Appendix B. Description of test statistics

B.1. Permanent components of returns

Let Γ denote the vector of time-series parameters arranged in a convenient manner (e.g., $(\text{Vec}(A)', \text{Vec}(B)')$), and define $\sigma \equiv (\sigma_1, \dots, \sigma_4)'$ to be the list of diagonal elements of the diagonal covariance matrix of \mathbf{u}_t . Finally, define

$$f_a^{ih}(\Gamma) = \sum_{k=a}^{\infty} \beta_{ih}(k),$$

the formula for which is given in Eq. (8). A 'generic' raw component of the variance of permanent returns (e.g., Eq. (6) or Eq. (7)) can be written as

$$F(\Gamma; \sigma) = \left[\sum_i f_a^{ih}(\Gamma) \right]^2 \sigma_h$$

for choices of i, h and a that correspond the component of interest. Since F is continuously differentiable, then using an estimator $\hat{\Gamma}$ that is asymptotically

normal implies that

$$\sqrt{T}(F(\hat{\Gamma}; \sigma) - F(\Gamma; \sigma)) \rightarrow^d N(0, F_r(\Gamma; \sigma)\Omega_r F_r(\Gamma; \sigma)'),$$

where T is the sample size, F_r is the row-vector of partial derivatives of F with respect to the elements of Γ and $\Omega_r = E[(\hat{\Gamma} - \Gamma)(\hat{\Gamma} - \Gamma)']$ is the asymptotic covariance matrix of $\hat{\Gamma}$ [see, for example, Hamilton (1994) (Proposition 7.4)]. Therefore, a statistic can be constructed that is asymptotically standard normal to test whether $F(\Gamma; \sigma) = 0$. The form of the z -statistic for such a null is

$$z_F \equiv \frac{F(\hat{\Gamma}; \sigma)}{\sqrt{(1/T)\{F_r(\hat{\Gamma}; \sigma)\hat{\Omega}_r F_r(\hat{\Gamma}; \sigma)'\}},$$

where $\hat{\Omega}_r$ is a consistent estimate of Ω_r .¹¹

The mathematical structure of the exact sums in Eq. (8) is such that the elements of F_r take a very simple form involving products of the infinite sums of moving-average coefficients.

Lemma 1. Let γ be any element of Γ . Then

$$\frac{\partial f_o^{ih}}{\partial \gamma} = \begin{cases} -f_o^{il}(\Gamma)f_o^{jh}(\Gamma) & \text{if } \gamma = a_{lj}, \\ f_o^{il}(\Gamma)f_o^{jh}(\Gamma) & \text{if } \gamma = b_{lj} \end{cases}$$

and

$$\frac{\partial f_1^{ih}}{\partial \gamma} = \begin{cases} -f_o^{il}(\Gamma)f_o^{jh}(\Gamma) + a_{il}^{(-1)}a_{jh}^{(-1)} & \text{if } \gamma = a_{lj}, \\ f_o^{il}(\Gamma)f_o^{jh}(\Gamma) & \text{if } \gamma = b_{lj}, \end{cases}$$

where a_{lj} is the (l, j) th element of A , b_{lj} is the (l, j) th element of B , and $a_{ij}^{(-1)}$ is the (i, j) th element of A^{-1} .

Proof. To compute an element of F_r , we have to compute derivatives of the f_a^{ih} s with respect to each time series parameter. The computations are all about the same, and rely on Corollary 41 in Dhrymes (1984) (p. 125) that says if M is a non-singular matrix whose elements depend on the scalar parameter α , then $\partial M^{-1}/\partial \alpha = -M^{-1}(\partial M/\partial \alpha)M^{-1}$. The computation when $a = 0$ is

$$f_o^{ih}(\theta) = e_i' M^{-1} e_h \quad \text{where } M = (A - B),$$

¹¹Note that an estimate has not been substituted for the true value of σ . This is unnecessary because the z -ratio is independent of σ ; F and F_r depend linearly on the same element of σ , so it cancels from numerator and denominator. This makes sense because the question being addressed is whether the moving average coefficients are large or small, not whether the variance of the disturbance is large or small.

so

$$\frac{\partial f_o^{ih}}{\partial a_{lj}} = e'_i \frac{\partial M^{-1}}{\partial a_{lj}} e_h,$$

where a_{lj} is the (l, j) th element of A , and

$$\frac{\partial M^{-1}}{\partial a_{lj}} = -M^{-1} \frac{\partial M}{\partial a_{lj}} M^{-1}.$$

Note that $\partial M / \partial a_{lj}$ is the matrix filled with zeros except the (l, j) th entry which is unity, so

$$\frac{\partial M^{-1}}{\partial a_{lj}} = -(A - B)^{-1} e_l e'_j (A - B)^{-1};$$

thus,

$$\begin{aligned} \frac{\partial f_o^{ih}}{\partial a_{lj}} &= -e'_i (A - B)^{-1} e_l e'_j (A - B)^{-1} e_h \\ &= -(e'_i M^{-1} e_l)(e'_j M^{-1} e_h) = -f_o^{il} f_o^{jh}. \end{aligned}$$

The symmetry of the problem implies that if b_{lj} is the (l, j) th element of B then, for $a = 0$,

$$\frac{\partial f_o^{ih}}{\partial b_{lj}} = f_o^{il} f_o^{jh}.$$

When $a = 1$, the same form (product of f 's) obtains, but derivatives with respect to elements of A have an additional term:

$$\begin{aligned} \frac{\partial f_1^{ih}}{\partial a_{lj}} &= \frac{\partial f_o^{ih}}{\partial a_{lj}} - e'_i \frac{\partial A^{-1}}{\partial a_{lj}} e_h \\ &= -f_o^{il} f_o^{jh} + (e'_i A^{-1} e_l)(e'_j A^{-1} e_h) \\ &= -f_o^{il} f_o^{jh} + a_{il}^{(-1)} a_{jh}^{(-1)}, \end{aligned}$$

where $a_{il}^{(-1)}$ is the (i, l) th element of A^{-1} . Derivatives of f_1^{ih} with respect to the elements of B are identical to those of f_o^{ih} .||

B.2. Order flow

As before, let Γ denote the vector of time-series parameters and σ the list of elements of the diagonal covariance matrix of u_t . Also define

$$g_a^{ih}(\Gamma) = \sum_{k=a}^{\infty} \beta_{ih}(k)^2$$

as given in Eq. (11). A generic raw component of the variance of order flow can be written as

$$G(\Gamma; \sigma) = [g_a^{ih}(\Gamma)]\sigma_h$$

for choices of i, h and a that correspond the \mathcal{S}_h component of interest. Since G is continuously differentiable, using an estimator $\hat{\Gamma}$ that is asymptotically normal makes it possible to test whether $G(\Gamma; \sigma) = 0$ using a z -statistic:

$$z_G \equiv \frac{G(\hat{\Gamma}; \sigma)}{\sqrt{(1/T)\{G_{\Gamma}(\hat{\Gamma}; \sigma)\hat{\Omega}_{\Gamma}G_{\Gamma}(\hat{\Gamma}; \sigma)'\}}$$

where G_{Γ} is the row-vector of partial derivatives of G with respect to the elements of Γ .¹² It is clear from Eq. (11) that the vector G_{Γ} can be constructed from the derivatives of the diagonal terms of \mathcal{S}_h and $(\mathcal{S}_h - \Lambda_h)$ with respect to the elements of Γ . The next result provides formulas for these derivatives.

Lemma 2. Let γ be any element of Γ . Then $\partial g_o^{ih}/\partial\gamma$ is the (i, i) th element of the matrix $\partial\mathcal{S}_h/\partial\gamma$, and $\partial g_1^{ih}/\partial\gamma$ is the (i, i) th element of the matrix $(\partial\mathcal{S}_h/\partial\gamma - \partial\Lambda_h/\partial\gamma)$, where

$$Vec\left[\frac{\partial\mathcal{S}_h}{\partial\gamma}\right] = (I - \Theta \otimes \Theta)^{-1} Vec\left[\frac{\partial\Lambda_h}{\partial\gamma} + \frac{\partial\Theta}{\partial\gamma}\mathcal{S}_h\Theta' + \Theta\mathcal{S}_h\frac{\partial\Theta'}{\partial\gamma}\right]$$

and

$$\frac{\partial\Lambda_h}{\partial\gamma} = -\left\{A^{-1}\frac{\partial A}{\partial\gamma}A_h + \Lambda_h\frac{\partial A'}{\partial\gamma}(A^{-1})'\right\}$$

$$\frac{\partial\Theta}{\partial\gamma} = A^{-1}\left\{-\frac{\partial A}{\partial\gamma}\Theta + \frac{\partial B}{\partial\gamma}\right\}.$$

If $\gamma = a_{ij}$ then $\partial A/\partial\gamma$ is the matrix having the same dimension as A with unity as its (i, j) th entry and zeros elsewhere, and $\partial B/\partial\gamma$ is the matrix of zeros with the same dimension as B . If $\gamma = b_{ij}$ then $\partial B/\partial\gamma$ is the matrix having the same dimension as B with unity as its (i, j) th entry and zeros elsewhere, and $\partial A/\partial\gamma$ is the matrix of zeros with the same dimension as A .

Proof. From Hamilton (1994) (p. 265) we have

$$\mathcal{S}_h - \Theta\mathcal{S}_h\Theta' = \Lambda_h.$$

¹² As before, σ cancels from the ratio.

Differentiating with respect to γ_o yields

$$\begin{aligned} \frac{\partial \mathcal{S}_h}{\partial \gamma_o} - \left\{ \frac{\partial \Theta}{\partial \gamma_o} (\mathcal{S}_h \Theta') + \Theta \left(\frac{\partial \mathcal{S}_h \Theta'}{\partial \gamma_o} \right) \right\} &= \frac{\partial A_h}{\partial \gamma_o}, \\ \frac{\partial \mathcal{S}_h}{\partial \gamma_o} - \left\{ \frac{\partial \Theta}{\partial \gamma_o} (\mathcal{S}_h \Theta') + \Theta \left(\frac{\partial \mathcal{S}_h}{\partial \gamma_o} \Theta' + \mathcal{S}_h \frac{\partial \Theta'}{\partial \gamma_o} \right) \right\} &= \frac{\partial A_h}{\partial \gamma_o}, \\ \frac{\partial \mathcal{S}_h}{\partial \gamma_o} - \Theta \frac{\partial \mathcal{S}_h}{\partial \gamma_o} \Theta' &= \frac{\partial A_h}{\partial \gamma_o} + \frac{\partial \Theta}{\partial \gamma_o} \mathcal{S}_h \Theta' + \Theta \mathcal{S}_h \frac{\partial \Theta'}{\partial \gamma_o}. \end{aligned}$$

This expression can be vectorized and solved for

$$\text{Vec} \left[\frac{\partial \mathcal{S}_h}{\partial \gamma_o} \right] = (I - \Theta \otimes \Theta)^{-1} \text{Vec} \left[\frac{\partial A_h}{\partial \gamma_o} + \frac{\partial \Theta}{\partial \gamma_o} \mathcal{S}_h \Theta' + \Theta \mathcal{S}_h \frac{\partial \Theta'}{\partial \gamma_o} \right].$$

We also have

$$\begin{aligned} \frac{\partial A_h}{\partial \gamma_o} &= \frac{\partial}{\partial \gamma_o} (A^{-1} (e_h e_h') (A^{-1})') \\ &= \frac{\partial A^{-1}}{\partial \gamma_o} e_h e_h' (A^{-1})' + A^{-1} (e_h e_h') \left(\frac{\partial A^{-1}}{\partial \gamma_o} \right)' \\ &= -A^{-1} \frac{\partial A}{\partial \gamma_o} A^{-1} e_h e_h' (A^{-1})' + A^{-1} (e_h e_h') \left\{ -A^{-1} \frac{\partial A}{\partial \gamma_o} A^{-1} \right\}' \\ &= - \left\{ A^{-1} \frac{\partial A}{\partial \gamma_o} A_h + A_h \frac{\partial A'}{\partial \gamma_o} (A^{-1})' \right\}, \end{aligned}$$

and, finally,

$$\begin{aligned} \frac{\partial \Theta}{\partial \gamma_o} &= \frac{\partial}{\partial \gamma_o} (A^{-1} B) = -A^{-1} \frac{\partial A}{\partial \gamma_o} A^{-1} B + A^{-1} \frac{\partial B}{\partial \gamma_o} \\ &= -A^{-1} \frac{\partial A}{\partial \gamma_o} \Theta + A^{-1} \frac{\partial B}{\partial \gamma_o} \\ &= A^{-1} \left\{ -\frac{\partial A}{\partial \gamma_o} \Theta + \frac{\partial B}{\partial \gamma_o} \right\}. \end{aligned}$$

Appendix C. Description of algorithm to sign trades and compute returns

If the first transaction/quotation record of the day is a trade rather than a quote, we record the volume of the opening trade and sign it as positive (negative) if the opening transaction price is greater (less) than the previous day's closing price – the midquote if the stock closes with the posting of quotes, or the

last transaction price if it closes with a trade. If the price is the same, the trade is regarded to be a cross (buy volume roughly offsets sell volume) and a zero is recorded for opening order flow. The midpoint of the opening quotes is then used as the terminal price in calculating the overnight return, and as the initial price for calculating the subsequent daytime return. Opening order flow is not used in calculating daytime order flow. If the first record of the day is a quote rather than a trade, opening order flow is zero.

We sign trades that occur after the opening according to their nearness to the prevailing quotes. If such quotes are not available, the trade is signed according to whether its transaction price is higher or lower than the price of the preceding transaction. We follow the procedure suggested by Lee and Ready (1991) and select the freshest quotes that are at least five seconds old, relative to the timing of the trade. (The exception to this is that immediately after the first transaction of the day, the oldest quote may not be five seconds old, but we use it anyway.) We call these quotes a_t and b_t ; and denote the transaction price and quantity as p_t and x_t , respectively.

If $|a_t - p_t| < |p_t - b_t|$ then $n_t = x_t$,

Else if $|a_t - p_t| > |p_t - b_t|$ then $n_t = -x_t$,

Else if $|a_t - p_t| = |p_t - b_t|$, or there is a transaction following the opening transaction without an intervening quote revision, then

If $p_t > p_{t-1}$ then $n_t = x_t$,

Else if $p_t < p_{t-1}$ then $n_t = -x_t$,

Else if $p_t = p_{t-1}$ then $n_t = 0$.

We define daily net volume for day i to be $N_i = \sum_{t \text{ in day } i} n_t$.

References

- Admati, A., 1985. A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica* 53, 629–657.
- Allen, F., Karjalainen, R., 1995. Using genetic algorithms to find technical trading rules. Working paper, University of Pennsylvania.
- Amihud, Y., Mendelson, H., 1987. Trading mechanisms and stock returns: an empirical investigation. *Journal of Finance* 42, 533–553.
- Amihud, Y., Mendelson, H., 1991. Volatility and trading: evidence from the Japanese stock market. *Journal of Finance* 46, 1765–1789.
- Barclay, M., Litzenberger, R., Warner, J., 1990. Private information, trading volume and stock return variances. *The Review of Financial Studies* 3, 233–253.
- Beveridge, S., Nelson, C., 1981. A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary Economics* 7, 151–174.
- Blume, L., Easley, D., O'Hara, M., 1994. Market statistics and technical analysis: the role of volume. *Journal of Finance* 49, 153–181.

- Brock, W., Lakonishok, J., LeBaron, B., 1992. Simple technical trading rules and the stochastic properties of stock returns. *Journal of Finance* 47, 1731–1764.
- Brown, D., Jennings, R., 1989. On technical analysis. *The Review of Financial Studies* 4, 527–551.
- Campbell, J., Ammer, J., 1993. What moves the stock and bond markets? A variance decomposition for long-term asset returns. *Journal of Finance* 48, 3–37.
- Campbell, J., Grossman, S., Wang, J., 1993. Trading volume and serial correlation in stock returns. *Quarterly Journal of Economics* 108, 905–940.
- Chan, K.C., Christie, W., Schultz, P., 1994. Market structure, price discovery, and the intraday pattern of bid-ask spreads for NASDAQ securities. *Journal of Business* 68, 35–60.
- Conrad, J., Hameed, A., Niden, C., 1994. Volume and autocorrelations in short-horizon individual security returns. *Journal of Finance* 49, 1305–1330.
- Conrad, J., Kaul, G., 1988. Time-variation in expected returns. *Journal of Business* 61, 409–425.
- Conrad, J., Kaul, G., 1996. Anatomy of a trading strategy, Working Paper, University of North Carolina.
- Conrad, J., Kaul, G., Nimalendran, M., 1991. Components of short-horizon individual security returns. *Journal of Financial Economics* 29, 365–384.
- Danthine, J., 1978. Information, futures prices, and stabilizing speculation. *Journal of Economic Theory* 17, 79–98.
- Diamond, D., Verrecchia, R., 1981. Information aggregation in a noisy rational expectations economy. *Journal of Financial Economics* 9, 221–236.
- Forster, M., George, T., 1995. Trading hours, information flow and international cross-listing. *International Review of Financial Analysis* 4, 19–34.
- Forster, M., George, T., 1996. Pricing errors at the NYSE open and close: evidence from internationally cross-listed stocks. *Journal of Financial Intermediation* 5, 95–126.
- French, K., Roll, R., 1986. Stock return variances, the arrival of information and the reaction of traders. *Journal of Financial Economics* 17, 5–26.
- Gallant, A.R., Rossi, P., Tauchen, G., 1992. Stock prices and volume. *The Review of Financial Studies* 5, 199–242.
- George, T., Kaul, G., Nimalendran, M., 1991. Estimation of the bid-ask spread and its components: a new approach. *The Review of Financial Studies* 4, 623–656.
- Grossman, S., 1976. On the efficiency of competitive stock markets when traders have diverse information. *Journal of Finance* 31, 573–585.
- Grossman, S., 1978. Further results on the informational efficiency of competitive stock markets. *Journal of Economic Theory* 18, 81–101.
- Grossman, S., Stiglitz, J., 1980. On the impossibility of informationally efficient markets. *The American Economic Review* 70, 393–408.
- Grundy, B., McNichols, M., 1989. Trade and revelation of information through prices and direct disclosure. *The Review of Financial Studies* 2, 495–526.
- Hamilton, J., 1994. *Time Series Analysis*. Princeton University Press, Princeton, NJ.
- Hasbrouck, J., 1991. The summary informativeness of stock trades: an econometric analysis. *The Review of Financial Studies* 4, 571–595.
- Hasbrouck, J., 1993. Assessing the quality of a security market: a new approach to transaction cost measurement. *The Review of Financial Studies* 6, 191–212.
- Huang, R., Stoll, H., 1994. Market microstructure and stock return predictions. *The Review of Financial Studies* 7, 179–213.
- Jones, C., Kaul, G., Lipson, M., 1994. Information, trading and volatility. *Journal of Financial Economics* 36, 127–154.
- Judge, G. et al., 1984. *The Theory and Practice of Econometrics*. 2nd Ed., Wiley, New York.
- Kleidon, A., Werner, I., 1996. U.K. and U.S. trading of british cross-listed stocks: an intraday analysis of market integration. *The Review of Financial Studies* 9, 619–664.

- Leach, C., Madhavan, A., 1993. Price experimentation and security market structure. *The Review of Financial Studies* 6, 375–404.
- Lee, C., Ready, M., 1991. Inferring trade direction from intraday data. *Journal of Finance* 46, 733–746.
- Lo, A., MacKinlay, C., 1990. When are contrarian profits due to stock market overreaction?. *The Review of Financial Studies* 3, 175–205.
- Madhavan, A., 1992. Trading mechanisms in securities markets. *Journal of Finance* 47, 607–641.
- Madhavan, A., Richardson, M., Roomans, M., 1997. Why do security prices change? A transaction-level analysis of NYSE stocks. *The Review of Financial Studies* 10, 1035–1064.
- Neftci, S., 1991. Naive trading rules in financial markets and Wiener–Kolmogorov prediction theory: a study of ‘technical analysis’. *Journal of Business* 64, 549–571.
- Stoll, H., 1989. Inferring the components of the bid-ask spread: theory and empirical tests. *Journal of Finance* 44, 115–134.
- Stoll, H., Whaley, R., 1990. Stock market structure and volatility. *Review of Financial Studies* 3, 37–71.
- Verrecchia, R., 1982. Information acquisition in a noisy rational expectations economy. *Econometrica* 50, 1415–1430.