Review of Quantitative Finance and Accounting, 17: 327–350, 2001

Empirical Identification of Non-Informational Trades Using Trading Volume Data

BONG-SOO LEE
Department of Finance, Bauer College of Business, University of Houston, Houston, Texas 77204-6282
E-mail: blee@uh.edu

OLIVER M. RUI
Department of Accountancy, Hong Kong Polytechnic University, Hung Hom, KowLoon, Hong Kong
E-mail: acmrui@inet.polyu.edu.hk

Abstract. This paper empirically identifies non-informational and informational trades using stock returns and trading volume data of the U.S., Japanese, and U.K. stock markets and five individual firms. We achieve the identification by imposing a restriction from theoretical considerations. Our results show that trading volume is mainly driven by non-informational trades, while stock price movements are primarily driven by informational trades. We also find that, around the 1987 stock market crash, trading volumes due to non-informational trades increased dramatically, while the decline in stock market prices was due mainly to informational trades. Increases in volatilities both in returns and in trading volumes during and after the crash are mainly due to non-informational trades. Regarding the trading volume-serial correlation in the stock returns relationship, we find evidence that is consistent with theoretical predictions that non-informational components can account for high trading volume accompanied by a low serial correlation of stock returns.

Key words: informational trade, non-informational trade, trading volume

JEL Classification: C32, G12, G14

1. Introduction

Day-to-day movements in stock market prices and expected stock returns may occur for two reasons. One is due to public information that causes all investors to change their valuation of the stock market because of new information about fundamental shocks affecting it. The other is non-informational factors, such as interactions among different groups of investors with heterogeneous information. We call the former *informational* trade and the latter *non-informational* trade in this paper.[1] Some researchers introduce non-informational trades into models by exogenously shifting misperceptions of future stock payoffs, by irrational noise trading (e.g., DeLong, et al., 1989, 1990), or by over-confident investors who overestimate the precision of their private signal about security values (Daniel, Hirschleifer and Subrahmanyam, 1998). Others introduce them by heterogeneous information and investment opportunities or by shifts in the risk aversion of some traders (e.g., Admati and Pfleiderer, 1988, 1989; Campbell, Grossman and Wang, 1993; Wang, 1994).

It is very difficult to distinguish between these two different views of stock market movements using data on stock returns alone. One of the major differences between the two types of trades is their implications for stock market trading volume. If public information that affects all investors arrives, stock market trading volume may not be significantly affected. However, selling pressure by non-informational traders must have a substantial effect on trading volume. Therefore, the two types of trades can be distinguished by looking at trading volume. In addition, Campbell, Grossman and Wang (1993) (henceforth CGW) find that trades due to heterogeneous investors (i.e., non-informational trades) that are accompanied by high trading volume are expected to be associated with a low serial correlation in stock returns because market makers buying stocks would require higher expected returns to compensate for their bearing additional risk.

While CGW suggest the use of data on stock market trading volume as a means of distinguishing between the two types of trades and provide a model whose implications are consistent with this distinction, previous studies have not provided an *empirical identification* of the two types of trades using stock market trading volume data as well as stock return data. Our paper attempts to fill this void in the literature. Gallant, Rossi and Tauchen (1992) point out that previous empirical work on stock price-volume relations tends to be heavily data-based and not guided by rigorous equilibrium models with optimizing agents. In contrast, our empirical identification is based on the theoretical work of CGW (see also Daniel, Hirschleifer and Subrahmanyam, 1998).[2]

The purpose of this paper is to *empirically identify* the components of stock returns and trading volume due to non-informational and informational trades and examine whether the components due to non-informational trades can account for the empirical relationship between trading volume and the serial correlation of stock returns: high trading volume accompanied by a low serial correlation of stock returns. We employ a bivariate moving average (BMAR) model of trading volume and stock returns with two types of shocks—non-informational and informational shocks—as the basis of our analyses. But the model is in general under-identified to distinguish between the two shocks. We achieve the identification by exploiting a difference between the two types of shocks: the non-informational shock should affect trading volume, while the informational shock may not.[3]

Our results show that trading volume is mainly driven by non-informational trades, while stock price movements are primarily driven by informational trades. We also find that, around the 1987 stock market crash, trading volumes due to non-informational trades increased dramatically, while the decline in stock market prices was mainly due to informational trades. Increases in volatilities both in returns and in trading volumes after the crash are mainly due to non-informational trades. Regarding the trading volume-serial correlation in the stock returns relationship, we find evidence that is consistent with theoretical predictions that non-informational components can account for high trading volume accompanied by a low serial correlation of stock returns.

The paper is organized as follows. Section 2 briefly reviews related literature. Section 3 introduces an empirical framework that is used for the identification of two components of trading volume and stock returns. Section 4 describes data and reexamines the empirical relationship between trading volume and serial correlation in stock returns using the daily data of the U.S., Japanese and U.K. stock markets and five individual firms.

Section 5 presents the results of decomposition (i.e., identification). Section 6 concludes the paper.

2. Related literature

Campbell and Kyle (1993), DeLong, et al. (1989, 1990), Grossman and Miller (1988) and Shiller (1984) have developed models in which expected stock returns vary over time as some investors accommodate the shifting stock demands of other investors. But these studies do not pursue the implications of their models for trading volume. Wang (1994) develops a model of competitive stock trading in which investors are heterogeneous in their information and private investment opportunities. As a consequence, he can examine the link between the nature of heterogeneity among investors and the behavior of trading volume and its relation to price dynamics. But he does not explicitly examine the relation between trading volume and serial correlation of stock returns.

Campbell, Grossman and Wang (1993) present a model whose implications are consistent with the observed relation between trading volume and serial correlation of stock returns. Their model contains two types of investors with different risk aversions: liquidity investors and risk-averse expected utility maximizers, who effectively act as 'market makers.' If a large subset of investors becomes more risk-averse than the rest of the market, then the expected return from holding the stocks must rise to compensate marginal investors for bearing the risk. This leads to the reallocation of risk from those who become more risk-averse to the rest of the market, which will be reflected in a rise in trading volume. At the same time, the rise in expected future returns is achieved by a fall in current stock prices, which causes a negative current return. As a consequence, this model implies that price changes accompanied by high volume will tend to be reversed, and this will be less true of price changes on days with low volume.[4] Specifically, they introduce an informational signal (S_t) that all investors receive at time t about future dividend shocks and a non-informational variable (Z_t) as a measure of risk aversion of marginal investors in the market.

Conrad, Hameed and Niden (1994) test whether trading information is important in predicting price movements of securities using weekly individual securities. They find that the autocovariances of high- and low-transaction securities differ in sign and magnitude: High-transaction stocks experience price reversals or negative autocovariances in returns, but low-transaction stocks experience positive autocovariances in returns. Their finding is strongly consistent with the CGW (1993) model.[5]

In a related study, Kelly (1997) *empirically identifies* noise traders and smart money in a model of Campbell and Kyle (1993) by assuming that an individual's probability of being smart money (a noise trader) is increasing (falling) in wealth. Kelly (1997, p. 361) notes in his conclusion that it would be useful to know if noise variables could be the source of variation in trading volume in the model of Campbell, Grossman and Wang (1993).

3. An empirical framework

3.1. Identification based on a bivariate time-series representation

We consider a 2-by-l vector x_t consisting of two stationary variables: detrended trading volume v_t and stock return r_t series, i.e., $x_t = [v_t, r_t]'$. By the Wold theorem, x_t has the following bivariate moving average representation (BMAR)[6]:

$$x_t \equiv [v_t, r_t]' = B(L)\varepsilon_t, \tag{1}$$

or

$$\begin{bmatrix} v_t \\ r_t \end{bmatrix} = \begin{bmatrix} B_{11}(L) & B_{12}(L) \\ B_{21}(L) & B_{22}(L) \end{bmatrix} \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix} \tag{2}$$

where v_t = detrended trading volume; r_t = stock returns; ε_t is a 2×1 vector consisting of ε_t^1 and ε_t^2; ε_t^1 = shock due to non-informational trade, ε_t^2 = shock due to informational trade; L is the lag operator (i.e., $L^n k_t = x_{t-n}$); $B_{ij}(L)$ for $i, j = 1, 2$ is a polynomial in the lag operator L (i.e., $B_{ij}(L) = \sum_k b_{ij}(k)L^k$ with $\sum_k \equiv \sum_{k=0}^{\infty}$); and the innovations are orthonormalized such that $\mathrm{var}(\varepsilon_t) = I$.[7]

The representation implies that we interpret trading volume and stock returns to be driven by two types of shocks (i.e., disturbances or innovations): non-informational and informational shocks. We distinguish between the two types of shocks by the following identifying restriction: Non-informational disturbance ε_t^1 has a contemporaneous effect on trading volume, while informational disturbance ε_t^2 has no contemporaneous effect on trading volume. This restriction is motivated by theoretical considerations. For example, Campbell, Grossman and Wang (1993) derive in their model that trading volume is affected only by a non-informational variable such as the risk aversion variable of marginal investors in the market. Note that this restriction is not very stringent in that it still allows the informational disturbance to have some effects on trading volume after the initial period, and we let the data speak for themselves in this regard.

The time path of the dynamic effects of the two types of disturbances on trading volume and stock returns is represented by the coefficients of the polynomial $B_{ij}(L)$. Since $b_{12}(k)$ measures the effect of the second type of shock (ε_t^2) on the first variable (v_t) after k periods, the restriction that the informational shocks (ε_t^2) have no contemporaneous effect on trading volume is represented by the restriction:

$$b_{12}(k)|_{k=0} = b_{12}(0) = 0 \tag{3}$$

The non-informational shock ε_t^1 has a contemporaneous effect on trading volume in the absence of such a restriction.

3.2. A restricted BVAR model

The above BMAR on which an identifying restriction is imposed is, in practice, derived by inverting a bivariate vector autoregression (BVAR) model of x_t with non-orthonormalized innovations, and the restriction is imposed on this BVAR model. Hence, we discuss how to impose restrictions on the BVAR. Suppose that we estimate the following BVAR of $x_t = [v_t, r_t]'$ with m lags:

$$x_t \equiv \begin{bmatrix} v_t \\ r_t \end{bmatrix} = A(L)x_{t-1} + u_t \equiv \begin{bmatrix} A_{11}(L) & A_{12}(L) \\ A_{21}(L) & A_{22}(L) \end{bmatrix} \begin{bmatrix} v_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \tag{4}$$

where $A(L) = [A_{ij}(L)] = [\sum_{k=1}^{m} a_{ij}(k)L^{k-1}]$ for $i, j = 1, 2$, $u_t = [u_{1t}, u_{2t}]' = x_t - E(x_t \mid x_{t-s}, s \geq 1)$, with $\text{var}(u_t) = \Omega$. Thus, we have estimates of $A(L)$ and Ω. While ε_t is an orthonormalized innovation in z_t with $\text{var}(\varepsilon_t) = I$, u_t is a non-orthonormalized innovation in x_t.

The relationship between the bivariate moving average model (1) (or (2)) and the bivariate vector autoregression (4) is described in the following proposition.

PROPOSITION 1. *The bivariate model x_t with the restriction (3) provides a restriction that identifies ε_t^1 and ε_t^2 as non-informational and informational trade shocks, respectively.*

Proof. [See Appendix] ■

After we identify components of trading volume and stock returns due to non-informational shocks and those due to informational shocks, we can examine whether the observed negative relationship between trading volume and serial correlations in stock returns is due mainly to the non-informational components.

4. Data and preliminary results

4.1. Data

The data set comprises *daily* market price index and trading volume series for the U.S., Japan and the U.K. Regarding the trading volume measure, Kugele, Van Ness and Van Ness (1999) evaluate two interpretations of the volume metric-number of shares traded and number of trades-in terms of which reflects information events first in the market. They employ a non-linear threshold autoregressive model to identify the lag and delay to news on a stock-trading day basis for five actively traded NYSE issues. They find that in 1989 there is no statistically significant difference between the two, but in 1997 information is incorporated into volume 1.479 minutes faster than into the number of trades. We employ number of shares traded as the trading volume.

For the U.S., we use the S&P 500 index. The data cover the period of January 2, 1973–August 28, 1997, and consist of 6,436 observations for each series. For Japan, we use the TOPIX index. The data cover the period of January 7, 1970–August 28, 1997, and consist of 7,220 observations. For the U.K., we use the *Financial Times* 100 Ordinary Share Index. The index covers the period of October 27, 1986–August 28, 1997, and consists of 3,093 observations for each variable. The above data are taken from the Datastream database.

In addition to the three markets' index returns and trading volume data, we collect the price and volume series for Coca-Cola, Kodak, IBM, Amoco and Alcoa from the CRSP database. The data cover the period of January 2, 1970–August 28, 1997, and consist of 6992 observations.[8] This is in part because non-synchronous trading can cause positive serial correlation in returns, which may obfuscate the relationship between the trading volume-serial correlation in stock returns. The bid-ask spread may make short-horizon correlations in returns more negative (e.g., Kaul and Nimalendran, 1990). The spread may not have an appreciable effect when index returns are used, but it would have an effect if individual share returns are used.[9]

4.2. Trend and unit root tests

The bivariate framework in Section 3 requires that the variables are stationary. As such, we test for the stationarity of stock returns and trading volume data. There are two ways to achieve stationarity. Some series need to be detrended (called the *trend-stationary* process), and others need to be differenced (called the *difference-stationary* or *integrated of order one,* I (1), process, or *unit root* process).

Previous work reports strong evidence of both linear and nonlinear time trends in trading volume series (e.g., Gallant, Rossi and Tauchen, 1992). As such, trend stationarity in trading volume is tested by regressing the series on a deterministic function of time. To allow for a nonlinear time trend as well as a linear trend, we include a quadratic time trend term:

$$v_t = \alpha + \beta_1 t + \beta_2 t^2 + \varepsilon_t \tag{5}$$

where v_t is raw trading volume in each stock market.

To test for a unit root (or the difference stationary process), we employ both the augmented Dickey-Fuller (D-F) test (1979) and the Phillips-Perron (P-P) test (1988). The difference between the two unit root tests lies in their treatment of any 'nuisance' serial correlation. The P-P test tends to be more robust to a wide range of serial correlations and time-dependent heteroskedasticity. In these tests, the null hypothesis is that a series is nonstationary (i.e., difference stationary): $\rho = 0$ and $\alpha = 1$ (see Table 1).

Test results are reported in Table 1. Panel A of Table 1 shows that the coefficients (with t-ratios in parentheses) of regressing trading volumes on a linear time trend alone are 5.590 (144.001), 4.162 (17.771), and 8.509 (47.272) in the U.S., Japanese, and U.K. stock markets, respectively, and they are all significant. The coefficients are 3.238 (86.776), 2.117 (61.791), 5.021 (98.718), 1.320 (103.224), and 1.043 (79.393) for Coca-Cola, Kodak, IBM, Amoco, and Alcoa, respectively. When we add a quadratic time trend term, the coefficients are very significant for all three stock markets' volume as well as for the five individual stocks' volume. Therefore, we use trading volume adjusted for both linear and nonlinear trends for all volume.

Panel B of Table 1 shows that the null hypothesis that the stock return series and detrended trading volume series are nonstationary (i.e., have a unit root) is rejected in all three stock markets and five individual stocks whether we allow for three lags or five lags. This confirms that detrended trading volume and stock return series are both stationary, and we do not

Table 1. Test for stationarity of stock returns and trading volume

Panel A: Linear and nonlinear trend tests in trading volume

$$v_t = \alpha + \beta t + \chi t^2 + \varepsilon_t$$

where v_t is raw trading volumes

α	β	χ	R^2
A. The U.S.			
−6300(−45.551)*	5.590 (144.001)*		0.769
2400 (15.913)*	−2.801 (−25.537)*	0.0014 (79.001)*	0.884
B. Japan			
30200 (37.401)*	4.162 (17.771)*		0.05
2700 (2.469)**	3.173 (36.837)*	−0.0046 (−33.056)*	0.196
C. The U.K.			
11100 (38.645)*	8.509 (47.272)*		0.449
16500 (40.942)*	−3.553 (−5.225)*	0.0044 (18.325)*	0.509
D. Coca-Cola			
−5230 (−34.689)*	3.238 (86.776)*		0.518
2500 (13.239)*	−3.392 (−27.172)*	0.0094 (54.854)*	0.663
E. Kodak			
−1710 (−12.352)*	2.117 (61.791)*		0.353
−1210 (−5.843)*	1.691 (12.343)*	0.0006 (3.214)**	0.353
F. IBM			
−6030 (−29.379)*	5.021 (98.718)*		0.582
210 (−0.706)	2.356 (2.122)**	0.0007 (26.618)*	0.621
G. Amoco			
−1100 (−21.365)*	1.320 (103.224)*		0.603
−3800 (−5.068)*	7.071 (13.971)*	0.0009 (12.514)*	0.612
H. Alcoa			
−1110 (−20.877)*	1.043 (79.393)*		0.474
900 (1.172)	1.540 (2.302)***	0.0014 (20.808)*	0.504

Note: 1. Numbers in parentheses are t-statistics. * , ** and *** denote significant at the 1%, 5% and 10%, respectively. 2. Sample periods are 01/02/73–29/08/97 for the U.S., 01/07/74–29/08/97 for Japan, 10/27/86–29/08/97 for the U.K., and 01/02/70–29/08/97 for Coca-Cola, Kodak, IBM, Amoco and Alcoa.

(*continued*)

Table 1. (*Continued*)

Panel B: Unit root tests for stock returns and trading volume

(a) Augmented Dickey-Fuller regression: $\Delta x_t = \rho_0 + \rho x_{t-1} + \sum_{i-1}^{n} \delta_i \Delta x_{t-i}$

(b) Phillips-Perron regression: $x_t = \alpha_0 + \alpha x_{t-1} + u_t$

	Lag 3		Lag 5	
Variable(x_t)	$t(\rho)$	$Z(t\alpha)$	$t(\rho)$	$Z(t\alpha)$
A. The U.S.				
r_t	−40.941*	−71.663*	−32.733*	−71.577*
v_t	−17.191*	−31.895*	−13.686*	−33.258*
r_t^2	−34.283*	−71.332*	−26.166*	−72.021*
B. Japan				
r_t	−36.955*	−68.508*	−31.765*	−68.514*
v_t	−13.351*	−22.586*	−11.684*	−23.298*
r_t^2	−30.567*	−58.953*	−25.988*	−59.787*
C. The U.K.				
r_t	−23.764*	−49.077*	−20.305*	−49.179*
v_t	−16.801*	−33.312*	−13.841*	−34.472*
r_t^2	−18.467*	−25.684*	−17.355*	−25.653*
D. Coca-Cola				
r_t	−44.101*	−79.989*	−36.091*	−79.982*
v_t	−23.428*	−42.843*	−19.043*	−44.449*
r_t^2	−32.279*	−49.399*	−27.853*	−49.714*
E. Kodak				
r_t	−46.831*	−85.752*	−37.056*	−86.021*
v_t	−26.874*	−45.821*	−22.314*	−46.992*
r_t^2	−35.714*	−52.983*	−30.434*	−52.927*
F. IBM				
r_t	−42.737*	−84.127*	−33.813*	−84.131*
v_t	−27.217*	−43.515*	−24.051*	−44.528*
r_t^2	−36.816*	−70.455*	−28.874*	−70.901*
G. Amoco				
r_t	−44.281*	−76.331*	−36.307*	−76.215*
v_t	−29.241*	−54.976*	−24.003*	−56.397*
r_t^2	−35.737*	−74.151*	−28.952*	−74.739*
H. Alcoa				
r_t	−42.728*	−75.037*	−35.168*	−74.930*
v_t	−22.941*	−44.761*	−19.193*	−46.676*
r_t^2	−35.299*	−67.930*	−28.229*	−68.597*

Notes: 1. r_t = returns; v_t = linear and nonlinear detrended trading volumes; r_t^2 = variance of returns. 2. Critical values for t with observations are: −2.57, 10%, 2.86, 5%, and −3.43, 1% (Fuller, 1976, Table 8.5.2, p. 373). For the Phillips-Perron test, two lags are used to calculate variance and the Fuller (1976) critical values are used to measure significance. 3. * , ** and *** denote significant at the 1%, 5% and 10%, respectively.

have to consider the possible cointegration problem associated with these variables. For the estimation of the BVAR, we use five lags considering both the Akaike information criterion (AIC) and the Schwarz criterion.

4.3. The empirical relationship between trading volume and serial correlation in stock returns

CGW (1993) investigate the relationship between aggregate stock market trading volume and serial correlation of daily stock returns by estimating the following regression using U.S. data:

$$r_{t+1} = \alpha + \left(\sum_{i=1}^{5} \beta_i D_i + \gamma_1 v_t + \gamma_2 v_t^2 + \gamma_3 \left(1000 h_t^2\right) \right) r_t \qquad (6)$$

where r_t is stock return; D_i are day-of-the-week dummies; v_t is the detrended trading volume; and h_t is the stock return volatility.[10]

We reexamine the relationship using daily data of the U.S., Japan, U.K. markets and the five individual stocks. In order to estimate the above regression, we need a measure of stock return volatility. We take the conditional variance series estimated by the EGARCH model.[11]

Table 2 presents the results of the empirical relationship between trading volume and the first autocorrelation of stock returns. First, we regress the one-day-ahead stock return (r_{t+1}) on the current stock return (r_t) interacted not only with day-of-the-week dummies (D_i) but also with volume (v_t). The coefficients on the product of volumes and returns, γ_1, are significantly negative for the U.S. stock market, Coca-Cola, Kodak and Amoco. The coefficients on volume are negative and marginally significant for the Japanese stock market, IBM and Alcoa. The coefficient on volume, γ_1, is negative but not significant for the U.K. stock market.

Second, we allow interaction of the current return with dummies and with estimated conditional variance (h_t^2). When volume is excluded from the regression, the coefficients on the product of volatility and returns, γ_3, are significantly negative for the U.S. and Japanese stock markets and Coca-Cola stock. The coefficients γ_3 are significantly positive for the U.K. stock market and Alcoa stock. The coefficients for Kodak, IBM and Amoco are negative but not significant.

Finally, we allow the current return to interact with dummies, volume, volume squared, and conditional variance. The volume squared is included to capture any nonlinearity that may exist in the relationship between trading volume and autocorrelation. The coefficients γ_2 are positively significant for the U.K. stock market, Coca-Cola, Amoco and Alcoa. The coefficients γ_2 are negatively significant for the U.S. stock market and Kodak stock. When volume and volume squared appear in the regression, γ_1 are significantly negative for the three stock markets and the four individual stocks except IBM.

Overall, the findings in Table 2 for the three markets and five individual firms confirm those of Campbell, Grossman and Wang (1993) for the U.S. market that the first

Table 2. Trading volume, volatility, and the first autocorrelation in stock returns

$$r_{t+1} = \alpha + \left(\sum_{i=1}^{5} \beta_i D_i + \gamma_1 v_t + \gamma_2 v_t^2 + \gamma_3 (1000 h_t^2)\right) r_t$$

where r_t is stock return for the period t, D_i are the dummy variables of days of the week. v_t is the linear and nonlinear detrended trading volumes at time t, h_t is the estimated stock return volatility at time t using EGARCH model.

Specification	γ_1	γ_2	γ_3	R^2
A. The U.S.				
Volume	−0.884 (−7.667)*			0.009
Volatility			−0.089 (−3.233)*	0.002
Volume and Volatility	−1.259 (−3.206)*	−8.272 (−5.919)*	0.045 (1.427)	0.014
B. Japan				
Volume	−0.37 (−1.873)***			0.001
Volatility			−0.071 (−6.238)*	0.006
Volume and Volatility	−0.049 (−1.788)***	0.038 (0.671)	−0.072 (−6.207)*	0.006
C. The U.K.				
Volume	−0.191 (−1.530)			0.001
Volatility			0.326 (6.579)*	0.015
Volume and Volatility	−1.501 (−4.966)*	2.663 (3.315)*	0.492 (7.311)*	0.024
D. Coca-Cola				
Volume	−0.059 (−4.879)*			0.003
Volatility			−1.093 (−5.563)*	0.004
Volume and Volatility	−0.105 (−3.875)*	0.019 (3.651)*	−0.562 (−2.081)**	0.006
E. Kodak				
Volume	−0.016 (−3.113)*			0.001
Volatility			−0.024 (−1.441)	0.001
Volume and Volatility	−0.059 (−4.251)*	0.004 (3.383)*	0.0104 (0.564)	0.001
F. IBM				
Volume	−0.031 (−1.673)***			0.001
Volatility			−0.244 (−1.501)	0.001
Volume and Volatility	−0.005 (−1.456)	−0.0033 (−0.236)	−0.284 (−1.639)	0.002
G. Amoco				
Volume	−0.146 (−4.051)*			0.002
Volatility			−0.166 (−1.613)	0.002
Volume and Volatility	−0.267 (−3.972)*	0.138 (2.251)**	−0.077 (−0.736)	0.003
H. Alcoa				
Volume	−0.016 (−1.789)***			0.001
Volatility			0.488 (3.981)*	0.002
Volume and Volatility	−0.152 (−2.551)**	0.092 (2.354)**	0.556 (4.202)*	0.006

Note: * , ** and *** denote significant at the 1%, 5% and 10%, respectively.

daily autocorrelation of stock returns is lower on high-volume days than on low-volume days.

5. Two components of trading volume and stock returns

5.1. Decomposition of trading volume and stock returns

By imposing the identifying restriction and setting each type of shock equal to zero individually, we can decompose the historical values of trading volume and return series into two parts: one arising from the effects of non-informational disturbances (the non-informational component), and the other from the effects of informational disturbances (the informational component).[12]

We plot non-informational and informational components of returns and trading volume of the three stock markets and five individual firms in Figure 1. The following observations are made about the figures. First, the informational components are dominant components for stock return movements, while the non-informational components are dominant components for trading volumes for all three markets and five individual firms. The former observation implies that the majority of stock return movements reflects common public information about the overall health of the economy. The latter observation is consistent with the implications of the CGW model. When public information that affects all investors arrives, there is no reason to expect high volumes, whereas selling pressure by non-informational traders must reveal itself in unusually high volumes. These findings are also consistent with an observation by Karpoff (1987) that price changes are interpreted as the market evaluation of new information, while the corresponding volume is considered an indication of the extent to which investors disagree about the meaning of the information.

Second, around the 1987 market crash, net trading volumes due to non-informational trades increased dramatically as shown in Panels A3 (for the U.S.), B3 (for Japan), and C3 (for the U.K.), whereas trading volumes due to informational trades have been both up and down, resulting in minor net changes in the U.S., Japanese, and U.K. markets (see panels A4, B4, and C4). This finding suggests a wide-spread heterogeneous belief among investors during the market crash. A similar observation is made for the five individual firms, but the degree of the heterogeneous belief does not appear to be as strong as in the case of the three stock markets (see panels D3, E3, F3, G3, and H3).

Third, around the 1987 market crash, the decline in stock market returns was due mainly to informational trades. This is observed consistently for all three markets and five individual firms (see panels A2, B2, C2, D2, E2, F2, G2, and H2). Along with the second observation, this implies that, although trading volume due to informational trades was relatively small, its impact on stock market returns appears to be relatively large.

Fourth, increases in relative volatilities both in market returns and trading volumes in all three stock markets are primarily due to non-informational trades. Specifically, volatilities of returns and volumes due to informational trades are quite stable over time, but those due to non-informational trades have increased significantly since the 1987 market crash for all

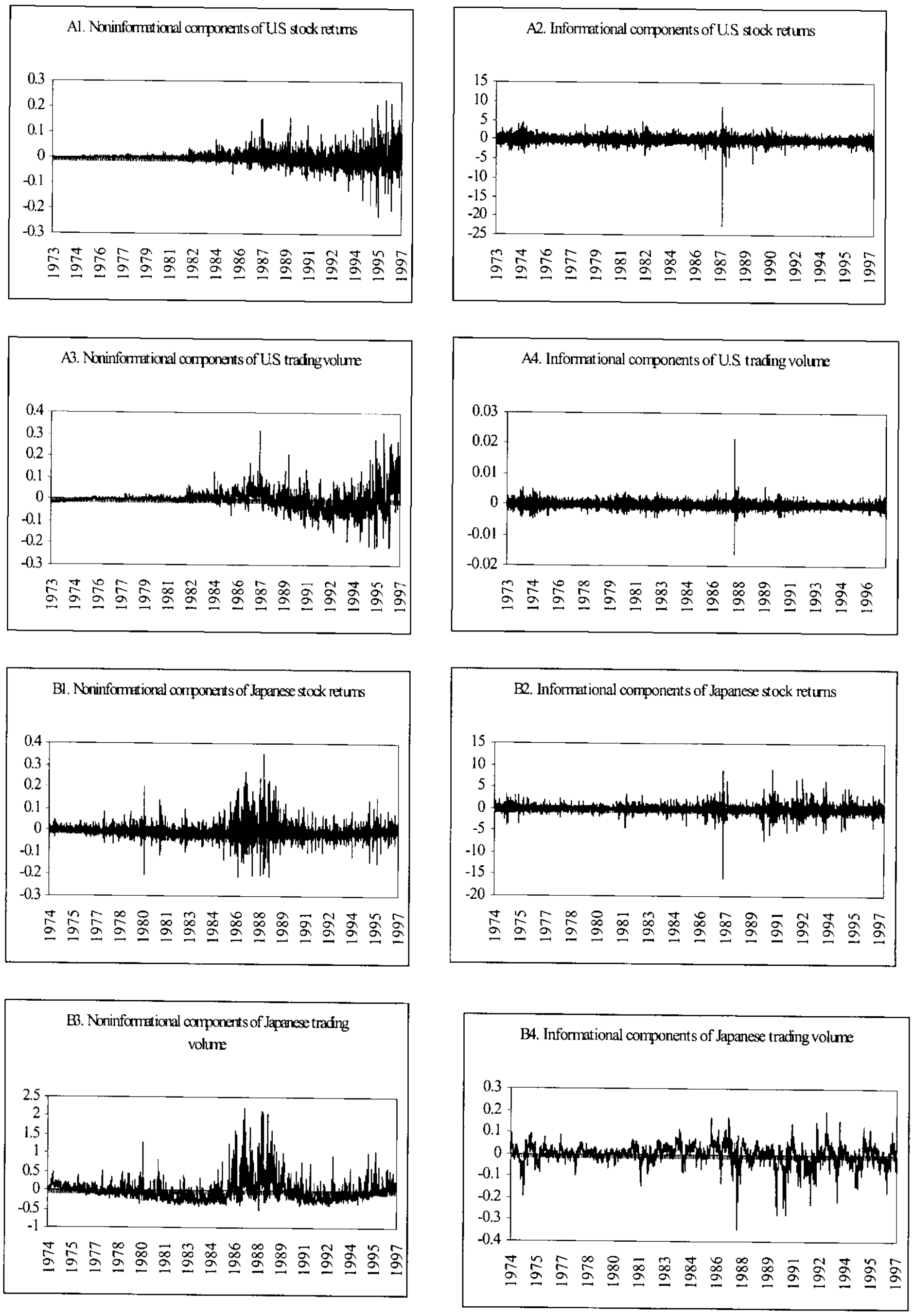

Figure 1. Two components of stock returns and trading volume.

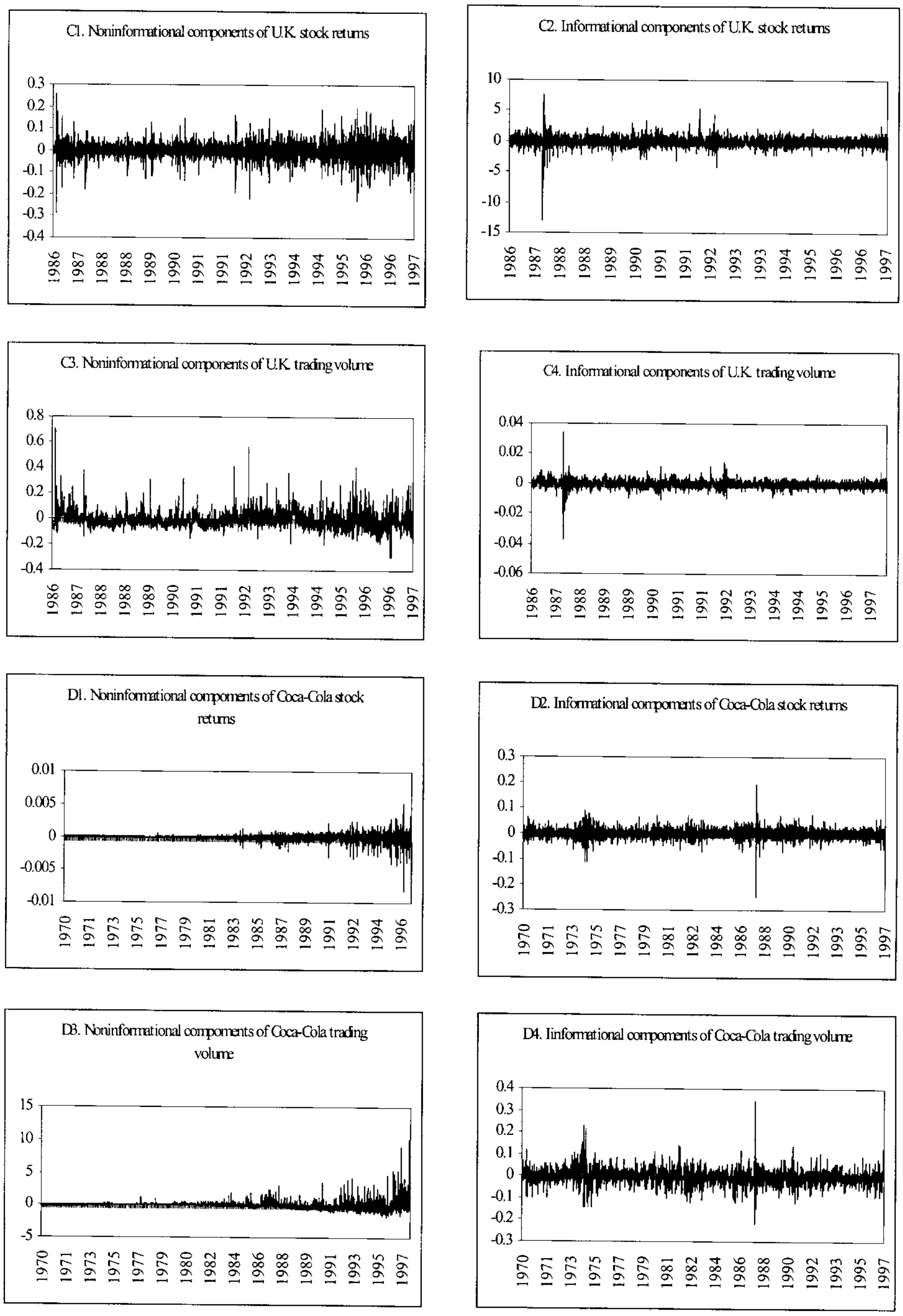

Figure 1. (*Continued*)

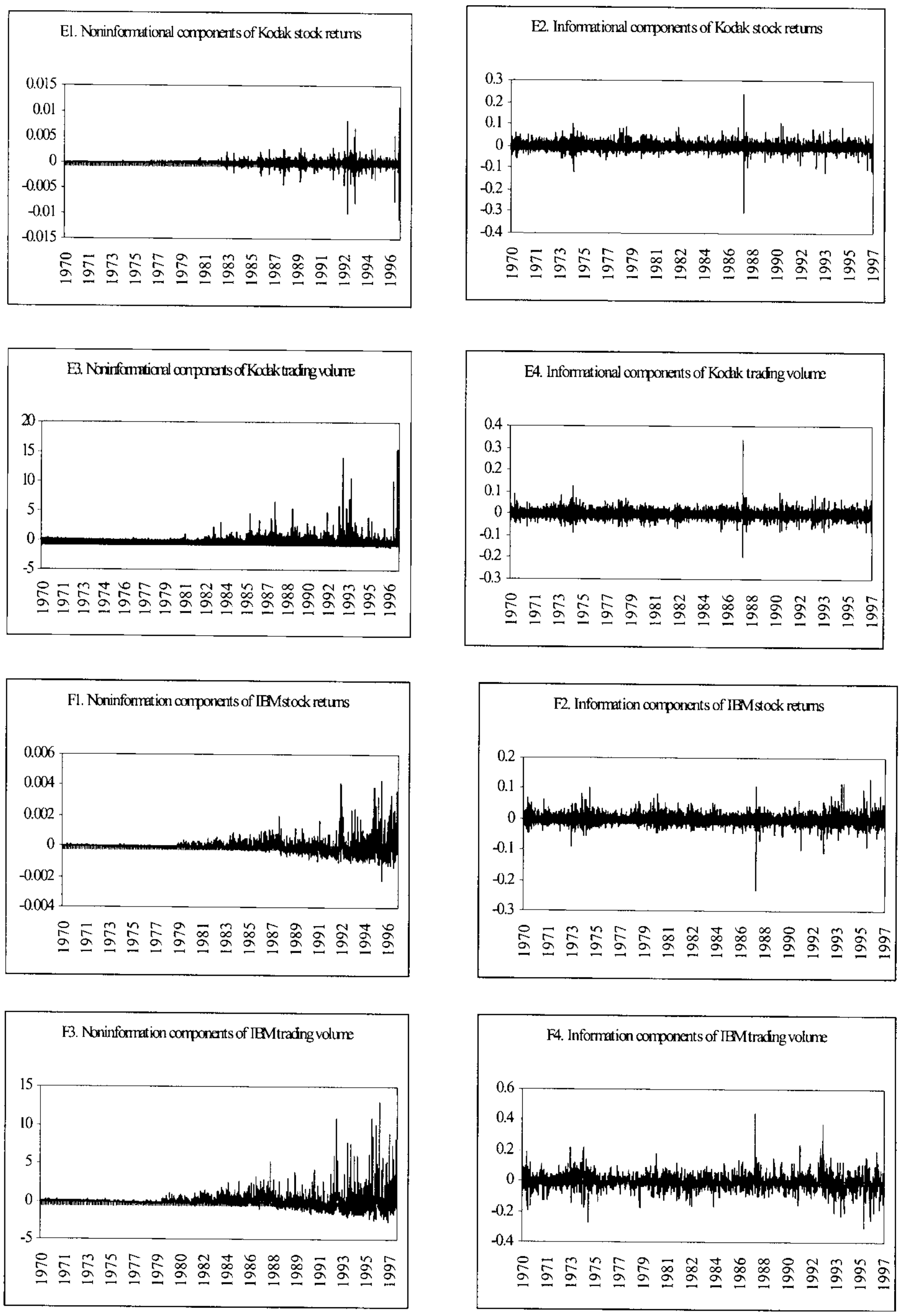

Figure 1. (*Continued*)

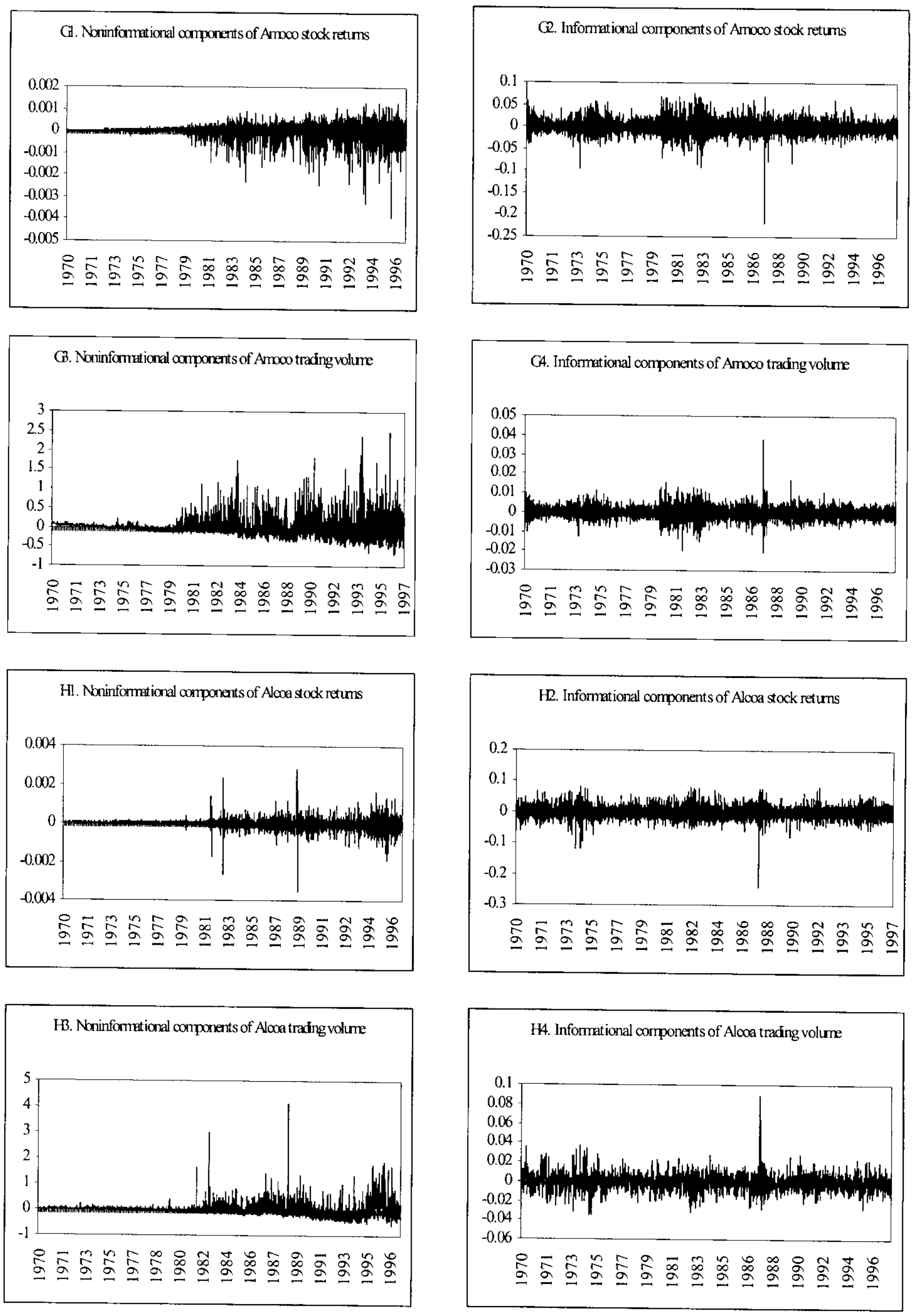

Figure 1. (*Continued*)

Table 3. Identification of noninformational and informational components of stock returns and trading volumes using trading volume data

$$\begin{bmatrix} v_t \\ r_t \end{bmatrix} = \begin{bmatrix} B_{11}(L) & B_{12}(L) \\ B_{21}(L) & B_{22}(L) \end{bmatrix} \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix}$$

with restriction $b_{12}(k)|_{k=0} = b_{12}(0) = 0$, where r_t = stock return for the period t; v_t = the linear and nonlinear detrended trading volumes at time t,; ε_t^1 = shock due to noninformational trading, ε_t^2 = shock due to informational trading.

Variables	Mean of Trading Volume due to ε_t^1	Mean of Trading Volume due to ε_t^2	First Autocorrelation of ε_t^1	First Autocorrelation of ε_t^2
The U.S.	7.24	$-1.10\ e^{-4}$	0.093	0.121
Japan	2.94	-4.20	0.091	0.115
The U.K.	1.59	$1.53\ e^{-2}$	-0.426	0.067
Coca-Cola	1.71	-0.451	0.034	0.344
Kodak	1.26	-0.510	-0.024	0.126
IBM	5.50	-4.50	0.376	0.704
Amoco	1.07	$3.95\ e^{-3}$	0.090	0.340
Alcoa	2.24	$-7.80\ e^{-4}$	-0.038	0.106

three markets.[13] However, this observation does not necessarily apply to the five individual firms.

Now we turn to the trade volume-serial correlation in the returns relationship using the decomposed components. Table 3 presents mean values of trading volumes and serial correlations in stock returns in both non-informational and informational components. The mean trading volume due to non-informational disturbance is substantially greater than that due to informational disturbance in all three stock markets and five individual firms, which confirms the observations about Figure 1. The first-order autocorrelation of stock returns due to non-informational shocks is lower than that due to informational shocks in all three stock markets and five individual firms. That is, in all cases we consider, high trading volume is accompanied by a lower autocorrelation of stock returns, and this is due to non-informational trades as predicted by the CGW model.

5.2. Dynamic effects of informational and non-informational shocks

By plotting the orthonormalized moving average coefficient, $b_{ij}(k)$, after imposing the identifying restriction, we obtain Figure 2 depicting the response of trading volume and stock returns to non-informational and informational disturbances of one standard deviation magnitude. A non-informational shock initially has a significantly positive effect on trading volume, and its effect declines gradually over time in all three markets and five individual firms. It has, however, negligible effects on stock returns in all three markets and five individual firms. In contrast, an informational shock initially has a significantly positive effect on stock returns but negligible effects on trading volume in all three markets

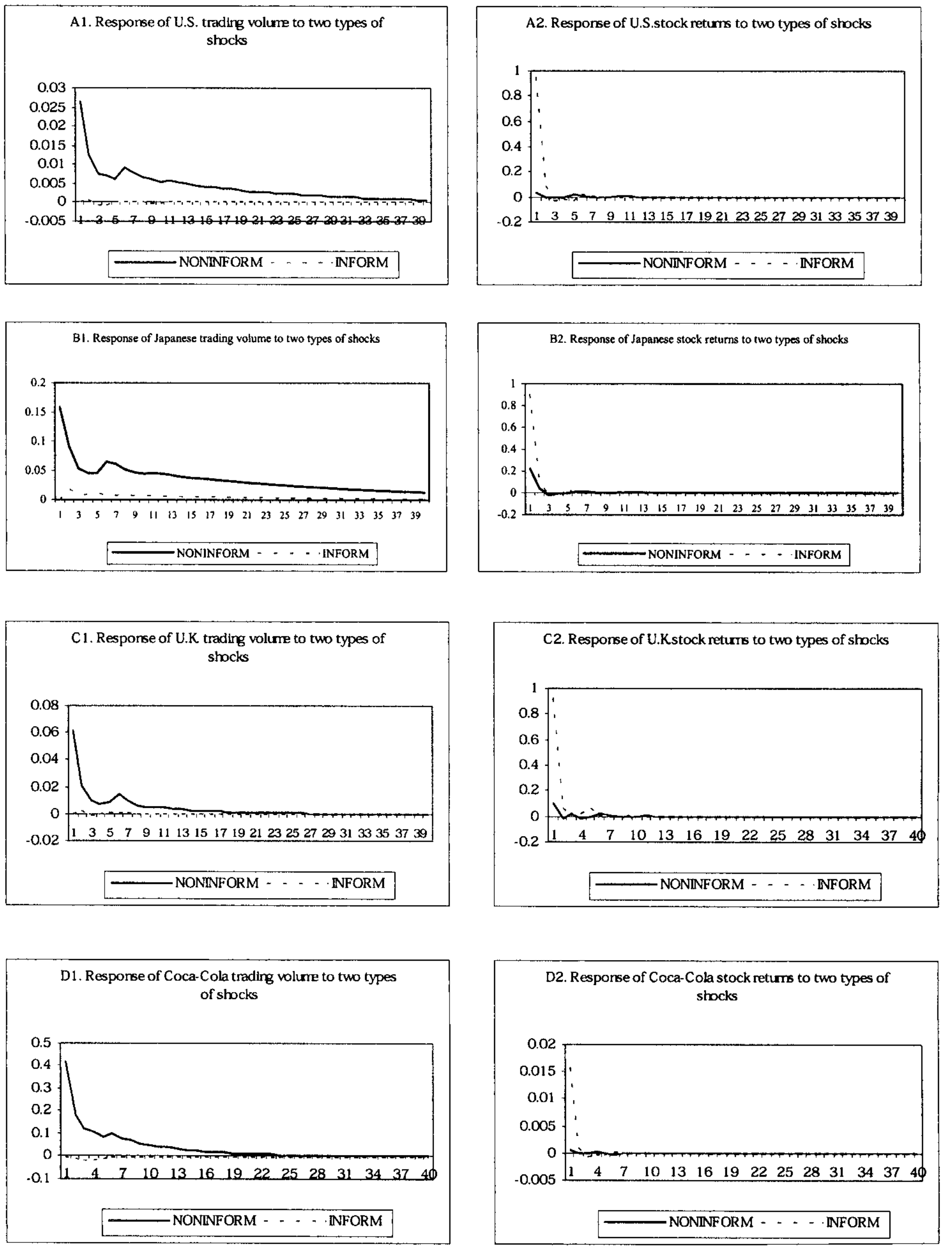

Figure 2. Dynamic response of trading volume and stock returns to two types of shocks.

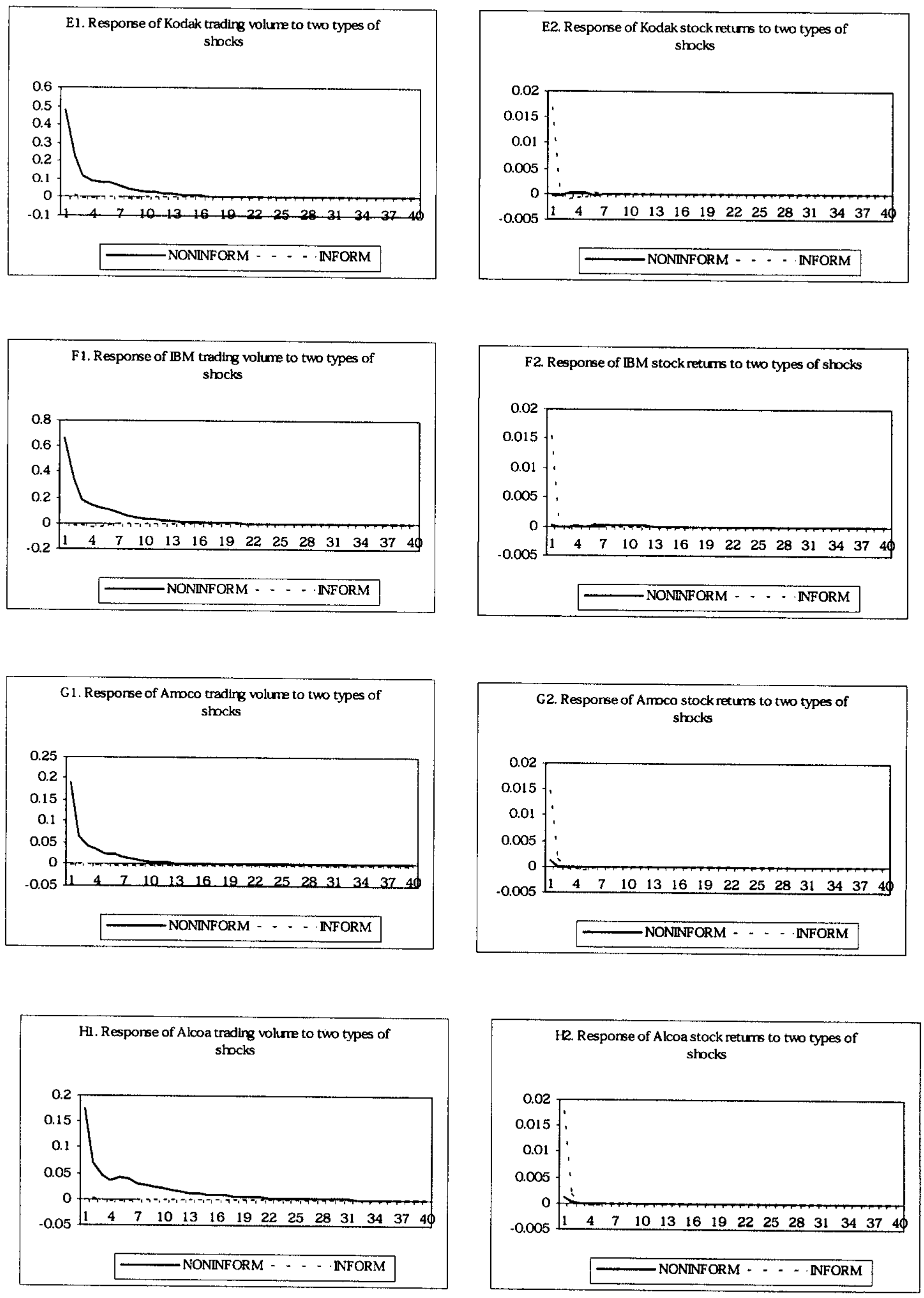

Figure 2. (*Continued*)

and five individual firms. These findings imply that trading volume is primarily driven by non-informational trades, but stock return movements are caused mainly by informational shocks.[14] It is also observed that stock prices adjust immediately (usually within two days), in particular, to informational shocks, whereas the effects of non-informational shock on trading volume tend to persist for days.[15]

6. Concluding remarks

We have attempted to empirically identify the non-informational and informational components of stock returns and trading volumes using stock market data of the U.S., the U.K. and Japan and five individual firms. We achieve the identification by imposing a restriction from theoretical considerations. Our results are encouraging in that they are consistent with the predictions of theoretical models: Trading volume is driven mainly by non-informational trades, while stock price movements are driven primarily by informational trades. Our decomposition also provides several interesting observations about the behavior of stock market returns and trading volume around the 1987 market crash and about the volatilities of the variables. We find that, around the 1987 stock market crash, trading volumes due to non-informational trades increased dramatically, while the decline in stock market prices was due mainly to informational trades. Increases in volatilities both in returns and in trading volumes during and after the crash are due mainly to non-informational trades. Regarding the trading volume-serial correlation in the stock returns relationship, we find evidence from the three stock markets and five individual firms that is consistent with the theoretical predictions that non-informational components can account for high trading volume accompanied by a low serial correlation of stock returns.

There may be alternative ways of identifying components of stock returns and trading volume due to non-informational disturbances. We have tried a few alternative bivariate models, but none of those performs better than the one reported in this paper. We may need to expand the model by introducing more than two types of disturbances into the model with more stringent identifying restrictions.

Appendix: Proof of Proposition 1

By inverting the BVAR of x_t in (4), we obtain a BMAR of x_t:

$$x_t = [I - A(L)L]^{-1}u_t, \tag{A-1}$$

where I is the identity matrix of rank 2.

By comparing the BMAR in (1) with that in (4), estimates of $B(L)$ can be obtained as follows. First, a typical innovation in the BMAR of (1) and that in the BVAR of (4) should be identical:

$$B_0\varepsilon_t = u_t, \tag{A-2}$$

where $B_0 = [b_{ij}(k)|_{k=0}] = [b_{ij}(0)]$. Second, x_t in (1) and (A-1) should be identical:

$$x_t = B(L)\varepsilon_t = [I - A(L)L]^{-1}u_t = [I - A(L)L]^{-1}B_0\varepsilon_t \tag{A-3}$$

where the last equality is obtained by using (A-2). Then, (A-3) implies that

$$B(L) = [I - A(L)L]^{-1}B_0. \tag{A-4}$$

Equation (A-4) implies that, to determine $B(L)$, we need an estimate of B_0 given an estimate of $A(L)$. This can be obtained by taking the variance of each side of (A-2):

$$B_0 B_0' = \Omega \tag{A-5}$$

In the bivariate model of x_t, (A-5) is given as

$$\begin{bmatrix} b_{11}(0) & b_{12}(0) \\ b_{21}(0) & b_{22}(0) \end{bmatrix} \begin{bmatrix} b_{11}(0) & b_{21}(0) \\ b_{12}(0) & b_{22}(0) \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \cdot \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \tag{A-6}$$

Here, we obtain three restrictions for the four elements of B_0. Hence, we need an additional restriction to just-identify the four elements of B_0. Our bivariate model of x_t in Section 3.1 provides this restriction: $b_{12}(0) = 0$. With estimates of B_0 along with $A(L)$, all the estimates of $B(L)$ can be obtained from (A-4).

Acknowledgment

We thank John Campbell, Ramon Rabinovitch, Jia He, Ron Singer, Raul Susmel, Joaquin Diaz-Saiz, and participants of the FMA meeting in Chicago for their valuable suggestions and comments. Any errors are entirely our own.

Notes

1. Different researchers use somewhat different terminology regarding informational and non-informational trades, and the distinction can be unclear because all trades are to some extent information-generated. Our distinction is very close to that of Grossman (1990).

   Grossman (1990) states, "... it is useful to partition the sources of stock price movement. First, the price of equity can change for purely informational reasons related to the market receiving news that changes its perception of the payoff stream to holding equity (e.g., the expected dividends or volatility of the dividend may change). Second, the price of equity may change because of changes in the marginal investor's tolerance for risk. Obviously, equity risk tolerance can change because of the arrival of information about perspective consumption changes which may be correlated with changes in equity payoffs, so that the above distinction can be cloudy."

   Wang (1994) develops a model with heterogeneous investors: informed and uninformed. Informed investors trade when they receive private information about the stock's future cash flow, which is referred to as their *informational* trading. They also trade to optimally rebalance their portfolios when their private investment opportunity changes, which is referred to as their *non-informational* trading. Uninformed investors trade only for non-informational reasons. When a trade from the informed investors is perceived to be non-informational, the uninformed investors will take the other side. Wang (1994) himself acknowledges a potential abuse of the terminology (see p.129, footnote 2).

2. If informed traders are overconfident about their abilities to forecast asset values given their information, then they will trade aggressively. Such aggressive trading may generate over-reaction in stock prices, and non-informational shock may cause high trading volume and low serial correlation in stock returns (see, e.g., Daniel, Hirshleifer and Subrahmanyam, 1998).

   Karpoff (1987) provides an extensive survey on the empirical relation between volume and contemporaneous price changes. He points out that the price-volume relation can indicate the importance of private versus public information in determining investors' demands (p. 110) (see also Pfleiderer, 1984).

   For the patterns of volume and price variability in intra-day transaction data, see Wood, McInish and Ord (1985) and Admati and Pfleiderer (1988, 1989).
3. In this paper, we achieve the just-identification of the bivariate model by imposing a restriction. This implies that, in the absence of over-identifying restrictions, the primary reason for our empirical analyses is not to formally (i.e., statistically) test the hypothesis but to assess the order of magnitude involved. Our intention is to show that our results derived from the just-identified model are consistent with implications of the informed and non-informed trades that we specify in the paper.
4. To justify the assumption in more detail, we quote CGW: "Suppose that one observes a fall in stock prices. This could be due to public information that has caused all investors to reduce their valuation of the stock market, or it could be due to exogenous selling pressure by noninformational traders. In the former case, there is no reason why the expected return on the stock market should have changed. In the latter case, market makers buying stock will require a higher expected return, so there will tend to be price increases on subsequent days. The two cases can be distinguished by looking at trading volume. If public information has arrived, there is no reason to expect a high volume of trade, whereas selling pressure by noninformational traders must reveal itself in unusual volume. Thus the model with heterogenous investors suggests that price changes accompanied by high volume will tend to be reversed; this will be less true of price changes on days with low volume." (p. 906)
5. Conrad, Hameed and Niden (1994) also find that, in low-transaction stocks, trading can reliably predict the next period's returns and that these relations are stronger in small firms. This finding is consistent with the predictions of a model in Blume, Easley and O'Hara (1994). Conrad and Niden (1992) find that significant increases in trading volume begin three trading days before the first acquisition announcement. However, they do not rule out the possibility that the increase in trading volume may be due to non-informational trades in our framework.
6. Bernanke (1986) provides an alternative method of identification for VAR models. For examples of bivariate identification applied to a present value relation and dividend policy, see Lee (1995, 1996).
7. For the relationship between non-orthonormalized and orthonormalized innovations, see Section 3.2.
8. Anderson (1996) uses these individual stocks in his study of stock return volatility and trading volume.
9. The relationship we test would be stronger if we were to use intra-day returns. However, given the difficulty of obtaining the intra-day trading volume data, our empirical analyses are conducted using daily data.
10. Campbell, Grossman and Wang (1993) use turnover data for trading volume, which is the ratio of the number of shares traded to the number of shares outstanding. This is sometimes called relative volume. To achieve stationarity of the trading volume, they detrend it by subtracting a one-year backward moving average of log turnover.

    Gallant, Rossi and Tauchen (1992) observe a time trend in trading volume series for the sample period of 1928 to 1985. To achieve stationarity, they experiment with transforming the volume series into a turnover series by dividing the volume by measures of the number of outstanding shares. However, they find a U-shaped pattern in the turnover series, and so they detrend the volume series by including a quadratic time trend term.

    Conrad, Hameed and Niden (1994) use the number of transactions as a measure of trading. However, they point out that their results are not sensitive to using trading volume data. They use weekly data for individual securities for the period of 1983–1990.
11. CGW take the conditional variance series estimated by Campbell and Hentschel (1992), who use a quadratic generalized autoregressive heteroskedasticity (QGARCH) model. They use a QGARCH model with one

autoregressive term, two moving average terms, and a mean return assumed to change in proportion to volatility. We use the EGARCH model for two reasons. First, Nelson (1991) finds that an EGARCH model adequately captures volatility persistence in *daily* aggregate stock returns for the 1962 to 1987 period. Second, Nelson demonstrates that EGARCH models have lognormal conditional variances in continuous time. An implication of this is that, as the sampling interval becomes shorter in discrete time, the distribution of innovation is a normal-lognormal mixture of distributions.

Specifically, we employ the following EGARCH(1,1) model to estimate stock return volatility:

$$
\begin{aligned}
& r_t = b_0 + \varepsilon_t \\
& \varepsilon_t \mid (\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) \sim N(0, h_t), \\
& \ln h_t = \alpha_0 + \alpha_1 \left(\frac{|\varepsilon_{t-1}| + \gamma \varepsilon_{t-1}}{h_{t-1}^{1/2}} \right) + \alpha_2 \ln h_{t-1}
\end{aligned}
$$

The exponential GARCH (EGARCH) model of Pagan and Schwert (1990) and Nelson (1991) was developed for two reasons. First, by using the exponential formulation, the restrictions of positive constraints on the estimated coefficients in ARCH and GARCH are no longer necessary. Second, a weakness of the GARCH model is that the conditional variance is dependent upon the magnitude of the disturbance term but not its sign. GARCH fails to capture the negative asymmetry apparent in many financial time series. The EGARCH model ameliorates this problem by allowing the standardized residual as an MA regressor in the variance equation while preserving the estimation of the magnitude effect. In addition, the ARCH/GARCH approach to modeling changing volatility precludes the testing of Black's (1976) leverage effect. The tendency for negative shocks to increase volatility more than positive shocks do is not captured in the ARCH/GARCH class of models.

12. Specifically, using notations in (2), the non-informational components of volume and stock returns are $B_{11}(L)\varepsilon_t^1$ and $B_{21}(L)\varepsilon_t^1$, respectively. The informational components of volume and stock returns are $B_{12}(L)\varepsilon_t^2$ and $B_{22}(L)\varepsilon_t^2$, respectively.
13. A dramatic increase in stock return volatility during and after the October 1987 stock market crash is well documented in Schwert (1990).
14. The mean values of return volatility for each component are follows:

	Non-Information Component	Information Component
U.S.	5.27	909.3
Japan	1.17	875.8
U.K.	2.02	856.1
Coca-Cola	1.88	247.4
Kodak	3.25	293.6
IBM	1.29	233.9
Amoco	9.18	272.5
Alcoa	4.45	326.3

15. As an alternative way of identification of the two components in trading volume and stock returns, one may consider the restriction $b_{11}(0) = 0$ on $x_t = [v_t, r_t]' = B(L)\varepsilon_t$ in (1). This requires that non-informational trades have no contemporaneous effect on trading volume, which contradicts our definition of non-informational trades. In addition, we have tried various identifying restrictions. We estimate a model consisting of trading volume squared and stock returns, $x_t = [v_t^2, r_t]'$, with an identifying restriction, $b_{12}(0) = 0$, which requires that non-informational disturbances should not affect trading volume volatility contemporaneously. The outcomes are inconsistent with the predictions of theoretical models. We find that the non-informational components have low trading volumes and higher serial correlations of stock returns compared with the informational components in all three stock markets, which is inconsistent with the predictions of the CGW model.

We also estimate a model consisting of the first-differenced trading volume and stock returns, $x_t = [\Delta v_t, r_t]'$, with an identifying restriction, $b_{12}(0) = 0$, which requires that non-informational disturbances should not affect changes in trading volume contemporaneously. Since detrended trading volume series is already stationary, the first-differenced trading volume prevents us from obtaining a decomposition after imposing the identifying restriction. We also tried identifying restrictions based on permanent and transitory components of trading volume and stock returns by imposing the restriction $B_{12}(L)|_{L=1} = B_{12}(1) = 0$. Besides the problem associated with theoretical justification of this restriction, the results from this model are not encouraging.

References

Admati, A. R. and P. Pfleiderer, "Selling and Trading on Information in Financial Markets." *American Economic Review* 78, 96–103, (1988a).

Admati, A. R. and P. Pfleiderer, "A Theory of Intraday Patterns: Volume and Price Variability." *Review of Financial Studies* 1, 3–40, (1988b).

Admati, A. R. and P. Pfleiderer, "Divide and Conquer: A Theory of Intraday and Day-of-The Mean Effects." *Review of Financial Studies* 2, 189–223, (1989).

Anderson, T ., "Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility." *Journal of Finance* 5, 169–204, (1996).

Bernanke, B., "Alternative Explanations of the Money-Income Correlations." *Carnegie-Rochester Conference Series on Public Policy* 25, 49–100, (1986).

Black, F ., "The Dividend Puzzle." *Journal of Portfolio Management* 2, 5–8, (1976).

Blume, L., D. Easley and M. O'Hara, "Market Statistics and Technical Analysis: The Role of Volume." *Journal of Finance* 49(1), 153–182, (1994).

Campbell, J. Y. and L. Hentschel, "No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns." *Journal of Financial Economics* 31, 281–318, (1992).

Campbell, J. Y., S. J. Grossman and J. Wang, "Trading Volume and Serial Correlation in Stock Returns." *The Quarterly Journal of Economics* 108, 905–939, (1993).

Campbell, J. Y. and A. S. Kyle, "Noise Trading and Stock Price Behavior." *Review of Economic Studies* 60, 1–34, (1993).

Conrad, J. and C. M. Niden, "Order Flow, Trading Costs and Corporate Acquisition Announcements." *Financial Management* Winter, 22–31, (1992).

Conrad, J., A. Hameed and C. Niden, "Volume and Autocovariances in Short-Horizon Individual Security Returns." *Journal of Finance* 49, 1305–1329, (1994).

Daniel, K., D. Hirshleifer and A. Subrahmanyam, "Investor Psychology and Security Market Under- and Overreactions." *Journal of Finance* 53, 1839–1885, (1998).

De Long, J. B., A. Shleifer, L. H. Summers and R. J. Waldmann, "The Size and Incidence of the Losses from Noise Trading." *Journal of Finance* 44, 681–196, (1989).

De Long, J. B., A. Shleifer, L. H. Summers and R. J. Waldmann, "Noise Trader Risk in Financial Markets." *Journal of Political Economy* 98(4), 703–738, (1990).

Dickey, D. A. and W. A. Fuller, "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of American Statistical Association* 74, 427–431, (1979).

Gallant, A. R., P. E. Rossi and G. Tauchen, "Stock Prices and Volume." *Review of Financial Studies* 5(2), 199–242, (1992).

Grossman, S. J., "Introduction to NBER Symposium on the October 1987 Crash." *Review of Financial Studies* 3(1), 1–3, (1990).

Grossman, S. J. and M. H. Miller, "Liquidity and Market Structure." *Journal of Finance* 43, 617–633, (1988).

Karpoff, J. M., "The Relation Between Price Changes and Trading Volume: A Survey." *Journal of Financial and Quantitative Analysis* 22, 109–126, (1987).

Kaul, G. and M. Nimalendran, “Price Reversals: Bid-Ask Errors or Market Overreaction?” *Journal of Financial Economics* 28, 67–93, (1990).

Kelly, M., “Do Noise Traders Influence Stock Prices?” *Journal of Money, Credit, and Banking* 29(3), 351–363, (1997).

Kugele, L. P., B. F. Van Ness and R. A. Van Ness, “Volume or Number of Trades: Which Reflects News First?” Working Paper, Kansas State University, (1999).

Lee, B. S., “The Response of Stock Prices to Permanent and Temporary Shocks to Dividends.” *Journal of Financial and Quantitative Analysis* 30, 1–22, (1995).

Lee, B. S., “Time Series Implications of Aggregate Dividend Behavior.” *Review of Financial Studies* 9, 585–614, (1996).

Nelson, D. B., “Conditional Heteroskedasticity in Asset Returns: A New Approach.” *Econometrica* 59, 347–370, (1991).

Pagan, A. R. and G. W. Schwert, “Alternative Models for Conditional Stock Volatility.” *Journal of Econometrics* 45, 267–290, (1990).

Pfleiderer, P., “The Volume of Trade and The Variability of Prices: A Framework for Analysis in Noisy Rational Expectation’s Equilibria.” Working Paper, Stanford University, (1984).

Phillips, P. C. B. and Pierre Perron, “Testing for a Unit Root in Time Series Regression.” *Biometrika* 75, 335–346, (1988).

Schwert, G. W., “Stock Volatility and the Crash of 1987.” *Review of Financial Studies* 3(1), 77–102, (1990).

Shalen, C. T., “Volume, Volatility, and the Dispersion of Beliefs.” *The Review of Financial Studies* 6, 405–434, (1993).

Shiller, R. J., “Stock Price and Social Dynamics.” *Brookings Papers on Economic Activity,* 457–498, (1984).

Tauchen, G. E. and M. Pitts, “The Price Variability-Volume Relationship on Speculative Markets.” *Econometrica,* 51, 485–505, (1983).

Wang, J., “A Model of Competitive Stock Trading Volume.” *Journal of Political Economy* 102, 127–168 (1994).

Wood, R. A., T. H. McInish and J. K. Ord, “An Investigation of Transaction Data for NYSE Stocks.” *Journal of Finance* 40, 723–741, (1985).