The Stock Price-Volume Linkage on the Toronto Stock Exchange: Before and After Automation

Cetin Ciner*
College of Business Administration
Northeastern University

Abstract: This paper investigates the information content of trading volume on the Toronto Stock Exchange before and after the move towards fully electronic trading. It is argued that if price discovery improves under electronic trading, the predictive power of volume should be less significant. The empirical analysis supports more accurate price discovery under electronic trading. Results from both the structural and vector autoregression models indicate that the predictive power of volume for price variability disappears after full automation.

August 1, 2002

* Assistant Professor of Finance, 413 Hayden Hall, Northeastern University, Boston, MA, 02115. Tel: 617- 373 4775, E-mail: c.ciner@neu.edu. The author is grateful to the late Craig Hiemstra for sharing his software to calculate the nonlinear causality tests and an anonymous referee for useful comments. The paper has also benefited from comments by Harlan Platt, Emery Trahan and Andy Saporoschenko.

The Stock Price-Volume Linkage on the Toronto Stock Exchange: Before and After Automation

1. Introduction

The stock price-volume relation has been the subject of many studies. Early theoretical models, such as the *mixture of distributions* model (MDM) of Clark (1973) and the *sequential information flow* model (SIF) of Copeland (1976), suggest that volume and price are jointly determined.[1] Relying on the motivation of these models, numerous papers test and consistently find evidence for a positive contemporaneous correlation between volume and price variability on equity markets. Karpoff (1987) provides an extensive review of this literature.

Among the more recent theoretical studies on the role of trading volume in asset markets, Blume, Easley and O'Hara (1994) and Suominen (2001) investigate the information content of volume on financial markets. Both of these papers suggest that stock prices are noisy and cannot convey all available information to market participants and that volume could be used as an informative statistic. Blume, Easley and O'Hara (1994) show that traders learn from volume and use it in their decision-making because volume conveys information about the precision of the informative signal that reaches the market. In Suominen (2001), volume is informative because it helps to resolve information asymmetries. He shows that traders estimate the availability of private

---

[1] The *sequential arrival of information* model of Copeland (1976) postulates that new information that reaches the market is not disseminated to all participants simultaneously, but to one investor at a time. Final information equilibrium is reached only after a sequence of transitional equilibria. Hence, due to the sequential information flow, lagged trading volume may have predictive power for current absolute stock returns and lagged absolute stock returns could have predictive power for current trading volume. The *mixture of distributions* model of Clark (1973) argues that returns and trading volume are positively correlated because the variance of returns is conditional upon the volume of that transaction. In Clark's model, trading volume is a proxy for the speed of information flow, which is regarded as a latent common factor that affects prices and volume synchronously. No causal relation from trading volume to returns is predicted in this model.

information using past volume and adjust their strategies. A common conclusion of these studies is that volume conveys information to the market that cannot be obtained from price alone and significant linkages are suggested between past volume and future price variability.

This paper builds on the motivation of these models and investigates the information content of trading volume on the Toronto Stock Exchange (TSE). In prior work, articles by Gallant, Rossi and Tauchen (1992) and Hiemstra and Jones (1994) examine the linkages between volume and returns on the U.S. equity markets. While the former study concludes that volume does not forecast returns, Hiemstra and Jones (1994) detect causality from volume to returns using nonlinear tests. Recently, Lee and Rui (2002) test for the information content of volume on the stock exchanges of U.S., U.K and Japan. They find that volume forecasts the magnitude of price changes (return volatility), however, no evidence exists for causality from volume to returns in any of the markets.

An investigation of the TSE contributes to this literature because it permits to examine the impact of electronic trading on the dynamic stock price-volume linkages, since the TSE has moved all trading from floor to an electronic platform during the sample period of the study. The TSE is fundamentally structured as an order-driven market in which specialists have market-making responsibilities, similar to the New York Stock Exchange (NYSE).[2] An electronic trading system was gradually implemented as an

---

[2] The markets in Canada and the US share similar structures and regulatory environments. In fact, Longin and Solnik (1995) argue that these two are the most correlated international markets. However, most studies of the Canadian and US markets report evidence against market integration, suggesting different information flows in the two markets. For instance, Booth and Johnston (1984), Foerster and Karolyi (1993), and Doukas and Switzer (2000) examine dual listed Canadian stocks and argue that the two markets are segmented, causing expected returns for similarly risky assets to be different. Furthermore, Karolyi

addition to floor trading on the TSE since the 1970s. On April 23, 1997, the TSE stopped using the floor and all trading has been fully electronic since then. This paper investigates the information content of volume on the TSE between 1990 and 2002 and focuses on whether the stock price-volume dynamics is influenced by the move towards fully electronic trading.

The linkage between automation and the information content of volume depends on whether automation increases price efficiency. Volume contains predictive power for price variability in the models of Blume, Easley and O'Hara (1994) and Suominen (2001) only because prices are noisy and cannot convey all available information to the market. Hence, it can be argued that if price discovery is more accurate under electronic trading, volume will be less informative about future price movements.

Several papers, such as Domowitz (1990, 1993), Massimb and Phelps (1994) and Naidu and Rozeff (1994), discuss the effects of automation on price efficiency. Advocates of automation suggest that execution of trades is faster and less costly under computerized trading systems. Traders have access to broader information, including bid and ask prices, trades sizes and volume, at lower costs, due to the existence of a limit order book, than under systems that restrict access to information about standing orders above and below the market. That would attract more investors and improve volume and liquidity and generate better price discovery. However, critics of automation argue that electronic trading could lead to less efficient prices since judgmental aspects of trade

---

(1995) and Racine and Ackert (2000) find declining dependence in Canadian stock market returns to shocks in the US market.

execution is lost with automation, which could be particularly important in times of fast market movements.[3]

In a related empirical study, Freund and Pagano (2000) examine price efficiency, using the rescaled range analysis, before and after automation on the NYSE and the TSE. Although they find that automation is associated with an improvement in market efficiency on the TSE relative to the NYSE, they do not detect any change in the nonrandom patterns in returns before and after automation, which leads them to conclude that automation has not changed price efficiency on the TSE. However, Freund and Pagano (2000) also point out that their results should be interpreted with caution since they rely on a relatively short sample. Specifically, their data cover the period between 1986 and 1997 and they specify between 1992 and 1997 as the post-automation period. Since the floor trading co-existed with electronic trading on the TSE until April 23, 1997, they examine a brief period under full electronic trading. In contrast, the present paper uses data until May 5, 2002, which should enable a more complete analysis of the impact of electronic trading.

In other empirical studies, researchers generally find that electronic trading improves price discovery. For example, Taylor et al. (2000) and Anderson and Vahid (2001) investigate the impact of electronic trading on price efficiency on the London and Australian stock exchanges, using smooth transition error-correction models. These studies focus on arbitrage between spot and futures markets of stock indices and report a significant decrease in transaction costs faced by arbitrageurs and conclude that the

[3] Furthermore, it can be argued that price efficiency remains unchanged after automation. According to this viewpoint, liquidity and efficiency on a stock market depends on rules on handling and execution of trades. If these rules do not change, then liquidity and efficiency is not expected to change.

markets have become more efficient under electronic trading. Similarly, Naidu and Rozeff (1994) investigate autocorrelation of returns on the Singapore Stock Exchange after automation and find reduced autocorrelations, leading them to conclude that market efficiency improves after automation.

The study examines the predictive power of volume both for the magnitude and direction of price changes, i.e. for absolute value of returns and returns per se. The empirical approach follows that of Lee and Rui (2002) and investigates causal and contemporaneous relations separately. Linear causality relations are examined using conventional vector autoregressions (VAR). Also, as Gallant, Rossi and Tauchen (1992) and Hiemstra and Jones (1994) show, volume and returns can have nonlinear linkages undetected by linear tests. The modified Baek and Brock (1994) tests are used to test for nonlinear causality. The contemporaneous relations, on the other hand, are examined within the context of a structural model, which accounts for the simultaneity bias and estimated using an instrumental variables (IV)-based generalized method of moments (GMM) approach.

The empirical findings support the argument that the move to automation does indeed coincide with an increase in price discovery on the TSE. Both the GMM-estimation results and the VAR analysis show that, although trading volume contains predictive power for price variability before automation, the predictability completely disappears after automation. The study also shows, in a preliminary analysis, that the first-order autocorrelation on the TSE 300 index is significantly reduced after automation, consistent an increase in price efficiency after automation. Predictability, however, is restricted to price variability since no evidence is detected to suggest that volume forecasts returns neither before nor after automation, consistent with the results of Gallant, Rossi and Tauchen (1992) and Lee and Rui (2002). Moreover, it is found that a positive contemporaneous relation exists between volume and absolute returns, as suggested by the MDH and prior empirical studies.

The next section discusses the econometric approach and hypotheses. Section 3 presents the data set, while the empirical findings are discussed in Sections 4 and 5. The final section offers the concluding remarks of the study.

2. Econometric Approach and Hypotheses

2.1 Contemporaneous Relations

The empirical analysis begins by examining the contemporaneous linkages between volume and returns (and absolute returns, henceforth). Several theoretical and empirical studies, see Karpoff (1987) for a survey of this literature, suggest that there is a contemporaneous relation between price and volume, which makes it crucial to control for the simultaneity in estimation. The approach in this study, adopted from Foster (1995) and also used by Ciner (2002) and Lee and Rui (2002), uses an instrumental variable (IV) estimator as a GMM estimator and constructs the following structural model:

$$V_t = a_0 + a_1 R_t + a_2 V_{t-1} + a_3 V_{t-2} + u_{1t}$$

$$R_t = b_0 + b_1 V_t + b_2 V_{t-1} + b_3 R_{t-1} + u_{2t} \qquad (1)$$

in which V_t is (log) trading volume and R_t denotes returns, calculated as (log) price differences.

The model treats V_t and R_t as endogenous variables; hence, OLS estimates will be inconsistent.[4] To estimate equation (1), lagged values of V_t and R_t are used as instrumental variables and the system is estimated by the GMM. The IV approach controls for the simultaneity bias and the GMM estimation controls for possible heteroskedasticity in error terms. Within the context of this system, significance of a_1 and b_1 would indicate a contemporaneous relation between volume and returns and significance of b_2 would suggest that lagged volume contains information about returns, which is further examined using vector autoregression models as discussed below.

2.2 Causal Relations

The study proceeds to test for Granger causality relations between volume and returns. Granger causality testing investigates whether the past of one time series improves the forecast of the present and future of another time series. Testing for linear Granger causality can be conducted within the context of a vector autoregression (VAR) model. The benefit of VAR models is that they account for linear intertemporal dynamics between variables, without imposing a priori restrictions of a particular model. Hence, they are ideally suited to detect stylized facts in the data. A VAR model to test for the dynamic linkages between volume and returns can be expressed as:

$$R_t = a_r + \sum_{i=1}^{l} b_{r,i} R_{t-i} + \sum_{i=1}^{m} c_r V_{t-i} + \sum_{i=1}^{k} D_i + u_{r,t} \tag{2}$$

$$V_t = a_v + \sum_{i=1}^{n} b_{v,i} R_{t-i} + \sum_{i=1}^{o} c_v V_{t-i} + \sum_{i=1}^{k} D_i + u_{v,t} \tag{3}$$

in which R_t represents returns, V_t denotes volume, Di's are dummy variables to account for the day of the week and month of the year effects in stock returns, $u_{r,t}$, $u_{v,t}$ are error

[4] More specifically, R_t is correlated with error term u_{1t}; hence, cov(R_t, u_{1t}) is not equal to zero, as required

terms and l, m, n and o denote the autoregressive lag lengths. Within the context of this VAR model, linear Granger causality restrictions can be defined as follows: If the null hypothesis that c_r's jointly equal zero is rejected, it is argued that volume Granger causes returns. Similarly, if the null hypothesis that b_v's jointly equal zero is rejected, returns Granger cause volume. If both of the null hypotheses are rejected, a bidirectional Granger causality, or a feedback relation, is said to exist between variables.

Different test statistics have been proposed to test for linear Granger causality restrictions. This study relies on the conventional χ^2-test for joint exclusion restrictions. Evidence reported in the literature suggests that this simplest form of linear causality testing is the most powerful (see, for example, Geweke, Meese and Dent (1983)).

In addition to linear linkages, volume and returns could have nonlinear linkages. For example, the models by Campbell, Grossman and Wang (1993) and Llorente et al. (2002) predict a nonlinear relationship between returns and volume. LeBaron (1992) and Duffee (1992) provide empirical evidence of significant nonlinear interactions between stock returns and trading volume. Hiemstra and Jones (1994) and Fujihara and Mougoue (1997) show that bidirectional nonlinear Granger causality exists between trading volume and returns in the U.S. equity and futures markets, respectively, although linear Granger causality tests cannot capture it.[5]

This study uses the modified Baek and Brock test, fully developed in Hiemstra and Jones (1994), to examine nonlinear causality relations. Baek and Brock (1992) offer a nonparametric statistical method to detect nonlinear causal relations that, by construction, cannot be uncovered by linear causality tests. Hiemstra and Jones (1994)

---

by OLS. Similarly, V_t is correlated with u_{2t} in the second equation.

modify their test to allow the variables to which the test is applied to exhibit short-term temporal dependence, rather than the Baek and Brock (1992) assumption that the variables are mutually independent and identically distributed.

The Baek and Brock (1992) approach begins with a testable implication of the definition of strict Granger noncausality. Consider two strictly stationary and weakly dependent time series $\{X_t\}$ and $\{Y_t\}$, t = 1, 2,.... Denote the m-length lead vector of X_t by X_t^m and the Lx-Length and Ly length lag vectors of X_t and Y_t, respectively. For given values of m, Lx, and Ly ≥ 1 and for e > 0, Y does not strictly Granger cause X if:

$$\Pr(\| X_t^m - X_s^m \| < e \big| \ \| X_{t-Lx}^{Lx} - X_{s-Lx}^{Lx} \| < e, \| Y_{t-ly}^{Ly} - Y_{s-Ly}^{Ly} \| < e \\ = \Pr(\| X_t^m - X_x^m \| < e \big| \ \| X_{t-Lx}^{Lx} - X_{s-Lx}^{Lx} \| < e) \quad (4)$$

in which Pr( ) denotes probability and || || denotes the maximum norm. The probability on the left side of equation (4) is the conditional probability that two arbitrary m-length lead vectors of $\{X_t\}$ are within a distance, e, of each other, given that the corresponding Lx-length lag vectors of $\{X_t\}$ and Ly-length lag vectors of $\{Y_t\}$ are within, e, of each other.

The strict Granger non-causality condition in equation (4) can then be expressed as

$$\frac{C1(m+Lx, Ly, e)}{C2(Lx, Ly, e)} = \frac{C3(m+Lx, e)}{C4(Lx, e)} \quad (5)$$

for given values of m, Lx, and Ly ≥1 and e>0, where C1,…,C4 are the correlation-integral estimates of the joint probabilities. Hiemstra and Jones (1994) discuss how to derive the joint probabilities and their corresponding correlation-integral estimators.

---

[5] In an empirical application of the Baek and Brock approach, Ciner (2001) shows that there are significant linkages between oil price changes and US stock price movements, uncovered by linear causality tests.

Assuming that X_t and Y_t are strictly stationary, weakly dependent, and satisfy the mixing conditions of Denker and Keller (1983), if Y_t does not Granger cause X_t, then,

$$\sqrt{n}\ \left(\frac{C1(m+Lx,Ly,e,n)}{C2(Lx,Ly,e,n)} = \frac{C3(m+Lx,e,n)}{C4(Lx,e,n)}\right) \sim N(0,\sigma^2\ (m,Lx,Ly,e)), \qquad (6)$$

Hiemstra and Jones (1994) show that a consistent estimator of the variance is, σ^2(m, Lx, Ly, e) = $\delta(n).\Sigma(n).\delta(n)'$.[6] To test for nonlinear causality between volume and returns, the test in equation (6) is applied to obtained residual series from the VAR models. Since the VAR model accounts for any linear dependencies, any remaining predictive power of one residual series for another can be considered nonlinear predictive power.

3. Data and Summary Statistics

The data consist of daily TSE 300 stock index closing values and trading volume, measured as (log) aggregate number of shares traded on the exchange, between January 2, 1990 and May 5, 2002. The TSE 300 is a value-weighted portfolio of 300 stocks from 14 industry groups, introduced in January 1977. There are a total of 3119 observations and the data are provided by the TSE. The period between January 2, 1990 to April 22, 1997 is specified as the pre-automation period and from April 23, 1997, when trading became fully electronic on the TSE, to May 5, 2002 is specified as the post-automation period.[7]

---

[6] A significantly positive value for the test statistic in (6) indicates that past values of Y help to forecast X, while a significantly negative value indicates that past values of Y confound the forecast of X. Therefore, Hiemstra and Jones (1994) argue that the test statistic should be evaluated with right-tailed critical values when testing for Granger causality.

[7] Freedman (2001) recently analyzes the performance of the Canadian economy during the 1990s. Canada has employed an inflation-targeting program since 1991 to maintain price stability. However, although inflation has fallen down to lower levels, Freedman suggests that Canada's economic performance in the 1990s, especially in the first half of the decade, was not entirely satisfactory, particularly when compared to the performance of the US economy over the same period. He argues that productivity growth in the US in the period has been largely in the production of high-technology machinery and electronics equipment, which are considerably less important sectors for Canada, as one of the possible reasons. That the prices of

The study first examines whether returns and volume contain a unit root, i.e. nonstationary. This is important because the VAR model requires that all variables are strictly stationary. The augmented Dickey Fuller (ADF) test is used to examine unit roots in variables and the results are reported in Table 1. The null hypothesis of the ADF test is nonstationarity and the lag lengths for the test are determined according to Akaike 's Information Criteria (AIC). The alternative hypothesis for volume is specified as stationary around a trend, since there is a growth in trading volume. Following Gallant, Rossi and Tauchen (1992) both a linear time trend, t, and a nonlinear, t^2, trend variables are included in the regression.[8] The alternative hypothesis for returns is specified as stationary with an intercept. The results indicate that the null hypothesis of unit root is safely rejected for all of the variables in both pre- and post- automation periods.

Table 1 also reports the sample statistics of the data set, which show that index returns have, on average, zero mean, negative skewness and excess kurtosis. Also, Naidu and Rozeff (1994) examine the impact of automation on price efficiency on the Singapore Stock Exchange by focusing on return autocorrelations before and after automation. Following their approach, the first-order autocorrelation of returns is calculated in the pre- and post-automation periods. It is found that the first-order autocorrelation of returns on the TSE 300 is .27 before automation, while it is reduced to .09 after automation. This is similar to the findings of Naidu and Rozeff (1994) and provides preliminary evidence suggesting that price efficiency on the TSE has improved after automation. To consider the economic significance of this reduction in dependency

---

raw materials have been weaker in the latter part of the 1990s, especially in the aftermath of the Asian crisis, could be another factor.

[8] The author thanks an anonymous referee for suggesting the inclusion of a nonlinear trend.

after automation, notice that the r-square of a regression of returns on a constant and its first lag is the square of the slope coefficient, which is simply the first-order autocorrelation. Hence, an autocorrelation of .27 implies that 7.29 percent of the variation on the TSE 300 index was predictable before automation, although only .81 percent is predictable after automation.

4. Contemporaneous Relations

The system of equations in (1) is estimated by the GMM and the results are reported in Table 2. An important point to determine is whether the system is exactly identified, i.e. a unique set of estimates for the coefficients in the model exists. If the system is overidentified, there will be multiple estimates for the coefficients. Hansen's (1982) test is used in this study to test for overidentification. The test statistics, also in Table 2, are very small in all of the cases, supporting a good fit of the model to the data.

The GMM-estimates suggest that there is a positive contemporaneous relation between volume and absolute returns both before and after automation. This finding is consistent with empirical results from the US equity markets as well as with the MDH of Clark (1973). According to the MDH, a contemporaneous relation between volume and absolute returns exists because a latent, exogenous variable, representing the rate of information arrival to the market, affects both volume and stock price variance, causing simultaneous movements. Prior research generally does not find a contemporaneous relation between volume and returns on equity markets, see Lee and Rui (2002) and Karpoff (1987). Consistent with this literature, no contemporaneous relation is detected between volume and returns after automation. However, a positive simultaneous relation between volume and return exists on the TSE in the pre-automation period.

The other coefficient of interest in the structural model is b_2, which measures the predictive power of lagged volume for returns. Recall that Blume, Easley and O'Hara (1994) and Suominen (2001) predict that when prices are noisy and cannot convey the available information, volume will have predictive power for future price variability. The results indicate that lagged volume significantly predicts price variability in the pre-automation period (p-value is.04). However, the predictive power of volume disappears in the post- automation period (p-value is .43). Within the context of Blume, Easley and O'Hara (1994) and Suominen (2001), this indicates an improvement in price discovery on the TSE after automation, which is further investigated using the VAR analysis below. On the other hand, no evidence exits to suggest that lagged volume does not predict future returns throughout the sample of the study.

5. Causality Relations

5.1 Linear Causality

This section discusses the results of testing for linear Granger causality between volume and returns. The VAR models are estimated by the OLS, including dummy variables to account for day of the week and the January effects and White's (1980) heteroskedasticity consistent standard errors are used to calculate the test statistics. Also, volume series is regressed over linear and nonlinear trend variables and the residuals from this regression are used in the analysis, to remove deterministic trends. The optimal lag lengths in the VAR models are determined by the AIC, with a maximum of 40 for univariate and 20 for bivariate lags. Thornton and Batten (1985) and Jones (1989) show that the VAR anlaysis is sensitive to lag length and suggest that the lags for dependent and independent variables should be determined differently.

The results of the χ^2-tests, reported in Table 3, show that no causality exists between volume and returns in either period, indicating that volume does not contain predictive power for the direction of price changes. This is consistent with results in prior studies, such as Gallant, Rossi and Tauchen (1992), Hiemstra and Jones (1994) and Lee and Rui (2001) as well as results from the GMM-analysis above. Residual diagnostics, also in Table 3, suggest that the VAR models successfully account for linear dependencies. However, the residuals exhibit nonlinear dependencies evinced by the significant values of the Ljung-Box Q-test applied to squared residuals.

Perhaps more important for the main motivation of the study are causality tests between volume and absolute returns. It is observed that volume contains predictive power for absolute returns in the pre-automation period, indicated by the significant value of the χ^2-test (p-value is .01). However, the predictability completely disappears in the post-automation period (p-value is .76). Hence, the findings from both the VAR approach and the GMM-analysis point to the same conclusion that the predictive power of volume becomes insignificant under fully electronic trading.

5.2 Nonlinear Causality

As mentioned before, Heimstra and Jones (1994) and Gallant, Rossi and Tauchen (1992) show that there are nonlinear linkages between volume and returns on the NYSE, uncovered by linear tests. Fujiahara and Mougoue (1997) reach the same conclusion in an examination of the U.S. oil futures markets. On the theoretical front, Campbell, Grossman and Wang (1993) and Llorente et al. (2001) argue that the relation between volume and returns could be nonlinear. Also, the residual diagnostics of the VAR models suggest that nonlinear dependencies remain in the error terms. Therefore, the study proceeds to test whether uncovered nonlinear causality remains between the variables.

The results of the modified Baek and Brock test statistics, applied to residuals from the VAR model, are reported in Table 4. To implement the modified Baek and Brock test, lead and lag truncation lengths (m, Lx and Ly) and the length scale parameter, e, have to be selected. Unlike in linear causality analysis, there are no established criteria to determine the optimal values for these parameters. Hence, this study relies on the Monte Carlo evidence in Hiemstra and Jones (1993), who find that for samples sizes of 500 or more observations, a lead length of m=1, lag lengths of Lx=Ly=1,2,…5 and length scale of e=1.0 provide good finite-sample size and power properties. However, the nonlinear causality test statistics are insignificant in all of the cases, suggesting no undetected causality between volume and returns. Hence, the conclusions of the linear analysis remain unchanged.

6. Concluding Remarks

Consistent with the common use of volume as an important statistic by practitioners, recent theoretical models by Blume, Easley and O'Hara (1994) and Suominen (2001), argue that volume contains useful information to forecast future price variability. According to their analysis, volume emerges as a useful statistic because prices are noisy and cannot convey all relevant information. Their approach is novel because they show that volume contains information independent from price that could affect the strategies of traders. Motivated by these theoretical models, this study investigates the information content of volume for subsequent price movements on the TSE. An investigation of the TSE is of interest since the TSE has moved trading from floor to an electronic platform in sample period examined. Hence, the TSE provides an opportunity to analyze whether electronic trading impacts the stock price volume dynamics.

The empirical findings indicate that the information content of volume is indeed not significant after automation. Although the evidence from both the VAR models and the GMM-analysis suggests that volume forecasts future price variability before automation, the predictability disappears after automation. Within the context of Blume, Easley and O'Hara (1994) and Suominen (2001), this indicates that price discovery improved after automation on the TSE. This conclusion is also supported by the observation that the first-order return autocorrelation on the main index of the TSE is significantly reduced after automation.

These findings are consistent with prior studies such as Naidu and Rozeff (1994), Taylor et al. (2000) and Andersen and Vahid (2001), who conclude that automation coincides with an improvement in price efficiency on the stock exchanges of Singapore, London and Australia. The findings, however, are not consistent with Freund and Pagano (2000), who find that there are no changes in nonrandom patterns on the TSE returns before and after automation, which leads them to conclude that automation on the TSE did not lead to a change in price efficiency. One explanation for different results, also mentioned by Freund and Pagano (2000), is that their data cover a relatively short time period under fully electronic trading and hence, could not capture the full impact of automation on the price formation process.

Predictability, however, is restricted to price variability. There is no evidence to suggest that returns can be predicted by volume, consistent with Gallant, Rossi and

Tauchen (1992) and Lee and Rui (2002). This is, of course, consistent with the efficient markets hypothesis, which argues that returns should not be forecast by publicly available information, like trading volume.

References

Anderson, H. M. and F. Vahid, 2001, Market architecture and nonlinear dynamics of Australian stock and futures indices, *Australian Economic Papers* 40, 541-566.

Baek, E. and W. Brock, 1992, A general test for nonlinear Granger causality: Bivariate model, *Working Paper*, Iowa State University and University of Wisconsin, Madison.

Blume, L., D. Easley and M. O'Hara, 1994, Market statistics and technical analysis: The role of volume. Journal of Finance 49, 153-181.

Booth, L. and D. Johnston, 1984, The ex-dividend day behavior of Canadian stock prices: Tax changes and clientele effects, *Journal of Finance* 39, 457-476.

Campbell, J., S. Grossman and J. Wang, 1993, Trading volume and serial correlation in stock returns, *Quarterly Journal of Economics* 108, 905-939.

Ciner, C., 2002, Information Content of Volume: An Investigation of Tokyo Commodity Futures Markets, *Pacific-Basin Finance Journal* 10, 201-215.

Ciner, C., 2001, Energy Shocks and Financial Markets: Nonlinear Linkages, *Studies in Nonlinear Dynamics and Econometrics* 5, 203-212.

Clark, P., 1973, A subordinated stochastic process model with finite variances for speculative prices, *Econometrica* 41, 135-155.

Copeland, T., 1976, A model of asset trading under the assumption of sequential information arrival, *Journal of Finance* 31,135-155.

Conrad, J., A. Hameed and C. Niden, 1994, Volume and autocovariances in short-horizon individual security returns. Journal of Finance 49, 1305-1329.

Denker, M. and G. Keller, 1983, On u-statistics and Von Mises statistics for weakly dependent processes, *Zetschrift fur Wahrscheninlichkeistheorie und Verwandte Gebiete* 64, 505-522.

Domowitz, I., 1990, The mechanics of automated trade execution systems, *Journal of International Money and Finance* 12, 607-631.

Domowitz, I., 1993, Automating the price discovery process: Some international comparisons and regulatory implications, *Journal Financial Services Research* 6, 305-326.

Doukas, J. and L. N. Switzer, 2000, Common stock returns and international listing announcements: Conditional tests of the mild segmentation hypothesis, *Journal of Banking and Finance* 47, 471-502.

Duffee, G, 1992, Trading volume and return reversals, *Working Paper*, Federal Reserve Board.

Foerster, S. R. and G. A. Karolyi, 1993, International listings of stocks: The case of Canada and the US, *Journal of International Business Studies* 24, 763-784.

Foster, A. J., 1995, Volume-volatility relationships for crude oil futures markets. Journal of Futures Markets 8, 929-951.

Fujihara, R. A. and M. Mougoue, 1997, An examination of linear and nonlinear causal relationships between price variability and volume in petroleum futures markets. Journal of Futures Markets 17, 385-416.

Freedman, C., 2001, Inflation targeting and the economy: Lessons from Canada's first decade, *Contemporary Economic Policy* 19, 2-19.

Freund, W. C. and M. S. Pagano, 2000, Market efficiency in specialist markets before and after automation, *Financial Review* 35, 79-

Gallant, R., P. Rossi and G. Tauchen, 1992, Stock prices and volume, *Review of Financial Studies* 5, 199-242.

Geweke, J., R. Meese and W. Dent, 1983, Comparing alternative tests of causality in temporal systems: Analytic results and experimental evidence, *Journal of Econometrics* 21, 161-194.

Hansen, L. P., 1982, Large sample properties of generalized method of moments estimators. Econometrica 50, 1029-1054.

Hiemstra, C. and J. D. Jones, 1993, Monte Carlo results for a modified version of the Baek and Brock nonlinear Granger causality test, *Working Paper*, University of Strathclyde and Securities and Exchange Commission.

Hiemstra, C. and J. D. Jones, 1994, Testing for linear and nonlinear Granger causality in the stock price-volume relation, *Journal of Finance* 49, 1639-1664.

Jones, J., 1989, A comparison of lag-length selection techniques in tests of Granger causality between money growth and inflation: Evidence for the U.S., 1959-86, *Applied Economics* 21, 809-822.

Karolyi, G. A., 1995, A multivariate GARCH model of international transmissions of stock returns and volatility: The cases of the US and Canada, *Journal of Business and Economic Statistics* 13, 11-25.

Karpoff, J., 1987, The relation between price changes and trading volume: A survey, *Journal of Financial and Quantitative Analysis* 22, 109-126.

Lee, B.-S. and O. M. Rui, 2002, The dynamic relationship between stock returns and trading volume: Domestic and cross-country evidence, *Journal of Banking and Finance* 26, 51-78.

LeBaron, B, 1992, Persistence of the Dow Jones index on rising volume. *Working Paper*, University of Wisconsin, Madison.

Llorente G., R. Michaely, G. Saar and J. Wang, 2002, Dynamic Volume-Return Relation of Individual Stocks, forthcoming, *Review of Financial Studies*.

Lo, A. and A. C. McKinlay, 1988, Stock market prices do not follow random walks: Evidence from a simple-specification test, *Review of Financial Studies* 1, 41-66.

Longin, F. and B. Solnik, 1995, Is the correlation in international equity returns constant: 1960-1990?, *Journal of International Money and Finance* 14, 3-26.

Massimb, M. N. and B. D. Phelps, 1994, Electronic trading, market structure and liquidity, *Financial Analysts Journal* 50, 39-50.

Naidu, G. N. and M. S. Rozeff, 1994, Volume, volatility, liquidity and efficiency on the Singapore Stock Exchange before and after automation, *Pacific-Basin Finance Journal* 2, 23-42.

Racine, M.D. and L. F. Ackert, 2000, Time-varying volatility in Canadian and U.S. stock index and index futures markets: A multivariate analysis, *Journal of Financial Research* 23, 129-143.

Suominen, M., 2001, Trading volume and information revelation in stock markets, *Journal of Financial and Quantitative Analysis* 36, 546-565.

Taylor, N., D. Van Dijk, P. H. Franses and A. Lucas, 2000, SETS, artibrage activity and stock price dynamics, *Journal of Banking and Finance* 24, 1289-1306.

Thornton, D. and D. Batten, 1985, Lag-length selection and tests of Granger causality between money and income, *Journal of Money, Credit and Banking* 17, 164-178.

White, H., 1980, A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity, *Econometrica* 48, 817-838.

Table 1. Sample Statistics

Panel A: Pre-Automation

N = 1841	R_t	V_t	$\|R_t\|$
Mean	.0002	17.022	.004
Std. Deviation	.005	.556	.003
Skewness	-.457	-.226	1.890
Kurtosis	2.264	-.588	5.705
ADF	-19.11	-4.25	-8.76

Panel B: Post-Automation

N = 1278	R_t	V_t	$\|R_t\|$
Mean	.0002	18.051	.008
Std. Deviation	.012	.344	.008
Skewness	-.687	-.433	2.592
Kurtosis	4.575	1.209	11.909
ADF	-9.19	-4.17	-5.33

Note- This table provides descriptive statistics for daily returns (and absolute returns) and trading volume on the Toronto stock exchange. The pre-automation period is January 2, 1990-April 22, 1997 and the post-automation period is April 23, 1997-May 5, 2002. The ADF test for unit roots is calculated with an intercept for R_t and $|R_t|$ and with an intercept and linear and nonlinear trends for V_t. 3, 13 and 37 augmentation lags are used in the ADF tests for R_t, $|R_t|$ and V_t, respectively, in the pre-automation period. 12, 33 and 34 lags are used for the variables in the post-automation period. The critical values of the tests are –2.86 and –3.41.

Table 2. Contemporaneous Relations

Panel A: Pre-Automation

Volume-Absolute Returns	Estimate	P-value	Volume-Returns	Estimate	P-value
a_o	1.909	(.00)	a_o	2.045	(.00)
a_1	14.822	(.27)	a_1	19.332	(.00)
a_2	.614	(.00)	a_2	.608	(.00)
a_3	.269	(.00)	a_3	.270	(.00)
b_o	-.006	(.11)	b_o	-.009	(.09)
b_1	.002	(.02)	b_1	.001	(.31)
b_2	-.001	(.04)	b_2	-.001	(.31)
b_3	.121	(.02)	b_3	.259	(.00)
Hansen	.00	(.99)	Hansen	.00	(.99)

Panel B: Post-Automation

Volume-Absolute Returns	Estimate	P-value	Volume-Returns	Estimate	P-value
a_o	6.579	(.00)	a_o	5.772	(.00)
a_1	22.517	(.03)	a_1	-6.051	(.40)
a_2	.492	(.00)	a_2	.515	(.00)
a_3	.133	(.00)	a_3	.166	(.00)
b_o	-.068	(.07)	b_o	-.072	(.23)
b_1	.006	(.19)	b_1	.011	(.15)
b_2	-.002	(.43)	b_2	-.007	(.13)
b_3	.092	(.00)	b_3	.094	(.00)
Hansen	.00	(.99)	Hansen	.00	(.99)

Note- This table provides the estimates of the following model:

$$V_t = a_0 + a_1 R_t + a_2 V_{t-1} + a_3 V_{t-2} + u_{1t}$$

$$R_t = b_0 + b_1 V_t + b_2 V_{t-1} + b_3 R_{t-1} + u_{2t}$$

in which R_t denotes returns (and absolute returns) and V_t denotes (log) trading volume. The model is estimated by the GMM and p-values for statistical significance are in parentheses. The row labeled Hansen refers to Hansen's (1982) goodness of fit test. The null hypothesis of this test is no overidentification restrictions.

Table 3. Linear Causality Tests

Panel A: Pre-Automation

Volume-Absolute Returns	χ^2-value	p-value	Volume-Returns	χ^2-value	p-value
$\|R_t\| \rightarrow V_t$	19.36	(.00)	$R_t \rightarrow V_t$	53.75	(.00)
$V_t \rightarrow \|R_t\|$	33.89	(.01)	$V_t \rightarrow R_t$	10.51	(.39)
Residual Diagnostics			Residual Diagnostics		
$Q(\|R_t\|)$	.14	(.99)	$Q(R_t)$:	6.04	(.91)
$Q(V_t)$	2.54	(.99)	$Q(V_t)$	2.32	(.99)
$Q^2(\|R_t\|)$	42.27	(.00)	$Q^2(R_t)$	109.78	(.00)
$Q^2(V)$	21.79	(.03)	$Q^2(V)$	23.32	(.02)

Panel B: Post-Automation

Volume-Absolute Returns	χ^2-value	p-value	Volume-Returns	χ^2-value	p-value
$\|R_t\| \rightarrow V_t$	12.73	(.02)	$R_t \rightarrow V_t$	17.18	(.64)
$V_t \rightarrow \|R_t\|$	4.98	(.76)	$V_t \rightarrow R_t$	4.63	(.32)
Residual Diagnostics			Residual Diagnostics		
$Q(\|R_t\|)$	.30	(.99)	$Q(R_t)$:	.10	(.99)
$Q(V_t)$	1.97	(.99)	$Q(V_t)$	2.18	(.99)
$Q^2(\|R_t\|)$	42.04	(.00)	$Q^2(R_t)$	128.56	(.00)
$Q^2(V)$	48.61	(.03)	$Q^2(V)$	48.32	(.00)

Note- This table provides the results of testing for linear Granger causality within the context of the following VAR model:

$$R_t = a_r + \sum_{i=1}^{l} b_{r,i} R_{t-i} + \sum_{i=1}^{m} c_r V_{t-i} + \sum_{i=1}^{k} D_i + u_{r,t}$$

$$V_t = a_v + \sum_{i=1}^{n} b_{v,i} R_{t-i} + \sum_{i=1}^{o} c_v V_{t-i} + \sum_{i=1}^{k} D_i + u_{v,t}$$

in which R_t denotes returns (and absolute returns) and V_t denotes (log) trading volume. The arrows indicate the direction of causality. The VAR model for returns and volume is estimated using 3 and 37 univariate and 15 and 10 bivariate lags in the pre-automation period, and 12 and 34 univariate and 4 and 20 bivariate lags in the post-automation period for returns and volume, respectively. The model for absolute returns and volume is estimated using 37 and 13 univariate and 7 and 18 bivariate lags in the pre-automation and 34 and 33 univariate and 5 and 8 bivariate lags in the post-automation period for absolute returns and volume, respectively. The χ^2-tests for joint exclusion restrictions are calculated using White's (1980) heteroskedasticity consistent standard errors. Q(12) and $Q^2(12)$ are Ljung-Box test statistics applied to residuals and squared residuals, respectively, at 12 lags. The results of the Ljung-Box tests are, however, robust to other lag length specifications.

Table 4. Nonlinear Causality Tests

Panel A: Pre-Automation

Volume-Absolute Returns				Volume-Returns			
$\lvert R_t \rvert \rightarrow V_t$		$V_t \rightarrow \lvert R_t \rvert$		$R_t \rightarrow V_t$		$V_t \rightarrow R_t$	
CS	TVAL	CS	TVAL	CS	TVAL	CS	TVAL
-.006	-2.019	-.002	- .972	-.004	- 1.336	-.0004	-1.169
-.008	-1.768	-.007	-1.551	.0003	.064	.0005	.533
-.002	-.390	-.010	-1.521	.007	1.026	.001	1.029
-.009	-1.024	-.003	-.401	.005	.553	.004	1.199
.004	.384	-.006	-.536	.004	.357	.009	.656

Panel B: Post-Automation

Volume-Absolute Returns				Volume-Returns			
$\lvert R_t \rvert \rightarrow V_t$		$V_t \rightarrow \lvert R_t \rvert$		$R_t \rightarrow V_t$		$V_t \rightarrow R_t$	
CS	TVAL	CS	TVAL	CS	TVAL	CS	TVAL
.006	1.362	.002	.585	.004	1.017	.008	1.835
.004	.587	-.005	-.880	.010	1.588	.008	1.182
-.001	-.169	-.001	-.167	.010	1.168	.005	.520
-.002	- .170	-.004	-.428	.013	1.268	.003	.266
-.0007	.051	-.003	-.227	.012	.952	.008	.546

Note-This table presents the results of testing for nonlinear causality between daily returns (and absolute returns) and trading volume. The modified Baek and Brock test is applied to the obtained residuals from the VAR models. The tests are applied to unconditionally standardized series, the lead length, m, is set to 1 and the length scale, e, is set to 1.0. CS and TVAL are the difference between the two conditional probabilities in the following equation

$$\Pr(\| X_t^m - X_s^m \| < e \mid \| X_{t-Lx}^{Lx} - X_{s-Lx}^{Lx} \| < e, \| Y_{t-ly}^{Ly} - Y_{s-Ly}^{Ly} \| < e$$
$$= \Pr(\| X_t^m - X_x^m \| < e \mid \| X_{t-Lx}^{Lx} - X_{s-Lx}^{Lx} \| < e)$$

and the standardized test statistic in

$$\sqrt{n}\ (\frac{C1(m+Lx,Ly,e,n)}{C2(Lx,Ly,e,n)} = \frac{C3(m+Lx,e,n)}{C4(Lx,e,n)}) \sim N(0,\sigma^2\ (m,Lx,Ly,e)),$$

The null hypothesis of the test statistic is no nonlinear Granger causality and it is asymptotically distributed N(0,1). The critical value at 5% significance level is 1.64.