Dynamic Volume-Volatility Relation

Hanfeng Wang[♠]

EFMA Classification Code: 310,330

October 2004

Tel.: 852-28591032 Mobile: 852-61033726
Email: hfwang@hkusua.hku.hk

[♠] Contact author and presenting author, School of Economics and Finance,The University of Hong Kong, Pokfulam Road, Hong Kong, Tel.: 852-28591032, Mobile: 852-61033726, Email: hfwang@hkusua.hku.hk.

Dynamic Volume-Volatility Relation

ABSTRACT

We find that trading volume not only contributes positively to the contemporaneous volatility, as indicated in previous literature, but also contributes negatively to the subsequent volatility. And this pattern between trading volume and volatility is consistently held among individual stocks, volume-based portfolios, size-based portfolios, and market index, and among daily data and weekly data. These empirical findings tend to support that the Information-Driven-Trade (IDT) hypothesis is more pervasive and powerful in explaining trading activities in the stock market than the Liquidity-Driven-Trade (LDT) hypothesis. Our additional tests obtain three interesting findings, 1) liquidity and the degree of information asymmetry influence the relation between volume and subsequent volatility, 2) the effect of volume on subsequent volatility and volume size have a non-linear relationship, indicating that at least empirically there exists a most information-intensive volume for each stock, which is consistent with Barclay and Warner (1993, JFE)'s finding, 3) the effect of volume on subsequent volatility is asymmetry when the stock price moves up and when the stock price moves down, and we attribute this asymmetry to the short-selling constraints.

A. Introduction

Financial researchers have devoted considerable efforts in understanding the relationship between trading volume and other financial proxies such as stock price and stock return. In this paper, we examine the relationship between trading volume and stock volatility. We find further evidence for the contemporaneous positive correlation between trading volume and volatility, and at the same time we document new patterns in the dynamics between stock volatility and trading volume. Specifically, we find that trading volume contributes positively to the contemporaneous stock volatility but it contributes negatively to the subsequent stock volatility. Using both daily data and weekly data, we find this dynamics in individual stocks, capital-based portfolios, volume-based portfolios and stock indices, and among the U.S., Japan, and China stock markets.

We are interested in the relation between trading volume and volatility first because volume and volatility are two important variables in financial economics and their relation is not comprehensively examined in the literature.

There are a number of papers dealing with the role of trading volume in financial markets. The existing finance literature generally agrees upon the positive contemporaneous correlation between price changes and trading volume. Karpoff (1987) offers a comprehensive survey on the relation between price changes and trading volume. He points out that empirically volume is positively related to the magnitude of the price change. Epps and Epps (1976) and Rogalski (1978) both support the above positive interrelation between price changes and volume. In his seminal paper, Wang (1994) develops a model which captures the link between the nature of heterogeneity among investors and the behavior of trading volume and its relation to price dynamics. His model also predicts volume is positively correlated with absolute price changes. The equilibrium model in Blume, Easley and O'Hara (1994) examines the informational role of volume and its applicability for technical analysis and the model supports the positive correlation between price changes and trading volume. Epps (1975) theoretically predicts, and also finds empirical support for, that the ratio of transaction volume to price changes on upticks exceeds the absolute value of this ratio on downticks. All of these literature tend to support the contemporaneous positive relation between trading volume and stock volatility.

The joint dynamics of price changes (or returns) and trading volume are also examined in the literature. Gervais, Kaniel and Mingelgrin (2000) (GKM hereafter) document that stocks experiencing unusually high trading volume over a period of one day to a week tend to appreciate over the course of the following month. Llorente, Michaely, Saar and Wang (2002) (henceforth LMSW) examine the possible dynamic relations between return and volume of individual stocks. Stickel and Verrecchia (1994) (henceforth SV) find that large price changes on days with weak volume support tend to reverse and returns do not reverse following days of strong volume support. Gallant, Rossi and Tauchen (1992) use the daily stock market index data and examine the joint dynamics of price changes and aggregate volume. Campell, Grossman and Wang (1993) (henceforth CGW) investigate the relationship between aggregate stock market trading volume and the serial correlation of daily stock returns. Chen, Firth and Rui (2001) examine nine national markets and indicate that trading volume contributes some information to the returns process.

However, there are only a few papers directly dealing with the relation between trading volume and stock volatility, and these papers generally build merely upon a statistical notion rather than from economic sense, and moreover they only address the contemporaneous relation between trading volume and stock volatility. Morgan (1976) provides evidence that the variance of returns on common stocks is not constant through time but is related to the volume of shares traded. Lamoureux and Lastrapes (1990) find that daily trading volume, used as proxy for information arrival time, has significant explanatory power regarding the variance of daily return. Gallant, Rossi, and Tauchen (1992) documents a positive correlation between conditional volatility and volume. Brock and LeBaron (1995) claims that their model is able to qualitatively reproduce the feature of positive correlation between squared returns and current volumes, but as they themselves point out, their model has many serious problems.

Since the overall price change is the sum of the individual price changes in a specific interval, the variance of return over this interval should be positively related to the number of transactions if we assume that the inter-transaction price changes happen independently. It is then not difficult to understand that the contemporaneous relation between trading volume and volatility is positive if we take the trading volume as a measure of the number of transactions. Morgan (1976) and Lamoureux and Lastrapes (1990) both base on this intuition and find support for this positive contemporaneous relation between trading volume and stock volatility. However, is this all the relation between trading volume and volatility? Since trading volume is the reflection of the process through which information is incorporated into stock prices, and price movements are inherently driven by trades between various investors and moreover, volatility is very sensitive to new information, by intuition we know trade volume should not only have effects on the contemporaneous volatility, but also may influence the subsequent volatility.

We are interested in the dynamic relation between trading volume and stock volatility also because the existing literature provides conflicting implications about their dynamic relation. As far as we know, no existing literature directly addresses the possible effect of trading volume on subsequent volatility. However, we do find some papers in the existing literature which have some implications about this effect. In order to better understand both the motivation and intuition for our study, we limit our discussion to the several papers that are most relevant to our work.

CGW investigates the relationship between aggregate stock market trading volume and the serial correlation of daily stock returns. They find that for both stock indexes and individual large stocks, the first-order daily return autocorrelation tends to decline with volume. They explain this phenomenon by a case in which some investors, "liquidity" or more generally "non-informational" traders, desire to sell stock for exogenous reasons, other investors who are risk-averse utility maximizers, are willing to accommodate the selling pressure thus resulting in a certain trading volume, but they demand a reward in the form of a lower stock price and a higher expected stock return in the future. Thus the return dynamics may be a negative expected return when they buy the stocks and a positive expected return in subsequent period. They build their model based on this intuition and their model thus 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. Therefore, the CGW model implies that if there is a high volume, the subsequent return reversal will result in a high volatility, that is, the trading volume and subsequent volatility are positively related. And consistent with CGW's model, Conrad, Hameed, and Nidden (1994) find that the subsequent autocorrelation of weekly individual firm returns is decreasing in the number of firm transactions. All of these tend to support the positive relation between trading volume and subsequent volatility.

Wang (1994) is a comprehensive study of the role of trading volume in asset pricing. His model captures two types of heterogeneity across risk-averse investors: (1) heterogeneous investment opportunities outside the public stock market, and (2) heterogeneous expectation about public information, that is asymmetric information. And there are two groups of investors. "Informed investors" may trade due to changes of opportunities outside the public stock market, and they may also trade because of better information about individual publicly traded stocks. "Uninformed investors" generally extract information from public signals like realized dividend, prices, etc. And they may trade due to revising the positions held when they traded against informed traders' private information, and they may also trade to take on new positions as they perceive new needs of non-informational trading from the informed investors.

In Wang's framework, the dynamic relation between volume and returns varies depending on the different motivations behind the trading by the informed investors. A reversal in consecutive returns is likely if the trading by informed traders is driven by changes of investment opportunities outside the stock market, we call this kind of trade the liquidity-driven trade. This is the case examined in CGW's study. Under such condition, we predict trading volume will contribute positively to the subsequent volatility. We called the view of positive correlation between trading volume and subsequent volatility the Liquidity-Driven Trade hypothesis (LDT).

Another possible situation addressed in Wang's framework is that momentum in consecutive returns is likely if the informed investors trade due to better private information. The intuition is that, when a subset of informed investors sells a stock when they have unfavorable private information, its price decreases, reflecting the negative private information about its payoff. Since this information is usually only partially incorporated into the price at the beginning, the negative return in the current period will be followed by another negative return in the next period. Thus this trading volume leads to lower subsequent volatility since these two period returns tend to be of the same sign, which means that high trading volume will be followed by a low volatility, that is trading volume and subsequent volatility are negatively related. Similarly we can analyze the case when the informed investors have favorable private information and we will get the same conclusion. We thus call the view of the negative correlation between trading volume and subsequent volatility the Information-Driven Trade hypothesis (IDT). Stickel and Verrecchia (1994)'s empirical evidences tend to support the IDT hypothesis. They find that that trading volume does sustains stock price changes, and that price increases on high volume days tend to be followed by another price increase the next day. A recent paper, Connolly and Stivers (2003) (hereafter CS), document substantial momentum (reversals) in consecutive weekly returns that during time of unexpectedly high (low) turnover. This can also be viewed as the evidence of IDT hypothesis.

LMSW's work is based on Wang (1994). And similarly they also argue that volume and return dynamics depend on the motivation behind the trade. Their model shows that "hedging trades" (which is essentially the same with the liquidity-driven trade discussed above) generate negatively auto-correlated returns and "speculative trades" (which is basically the same with the information-driven trade discussed above) generate positively auto-correlated returns. Their empirical evidences also support their model's predictions.

Thus we have two opposing views about the relation between trading volume and subsequent volatility. The IDT hypothesis supports a negative correlation between trading volume and subsequent volatility, but the LDT hypothesis supports a positive one. In this paper we examine the dynamics between trading volume and volatility, shedding light on which one of these two hypotheses dominates in the stock market data. Our empirical tests based on the data from international stock markets provide further evidence to support the positive relation between trading volume and contemporaneous volatility, and we document a negative correlation between trading volume and subsequent volatility, which tends to support that the IDT hypothesis is more powerful and pervasive in explaining stock market trading activities. And our findings are robust with daily data and weekly data, and among individual stocks, size-based portfolios, volume-based portfolios, and stock market index.

Our further tests obtain some interesting findings. Firstly, we find that two factors may influence the relation between volume and subsequent volatility, liquidity and information asymmetry. Our regression results indicate that for more liquid stocks, trading volume contribute more negatively to subsequent volatility, which is consistent with LMSW's conjecture, and for stocks with higher degree of informed trading, volume also contributes more negatively to subsequent volatility. Secondly, we find that trading volume and subsequent volatility could have a non-linear relationship. As Barclay and

Warner (1993) indicates, informed trades are concentrated in the medium volume size category. Since more informed trading will contribute more negatively to subsequent volatility, we thus expect medium size volume will have the largest negative effect on subsequent volatility. The consistently significant and positive coefficients of the square term of lagged volume in our estimation results do indicate that there is a volume size that maximizes the negative effect of trading volume on subsequent volatility. We name this volume size "the implied most information-intensive volume size". Thus our finding provides further evidence which is consistent with Barclay and Warner's findings. Thirdly, we also find an asymmetric effect of volume on subsequent volatility when stock price moves up and when the stock price moves down. Volume will contribute less negative to the following volatility when the stock price moves down. We conjecture that plausibly this asymmetry can be attributed to the short-selling constraints. We argue that more informed trading will contribute more negatively to the following volatility. However, if the informed traders are refrained from trading when they have "bad" information about a stock, their information can not be incorporated into stock price thus the negative effect of volume on subsequent volatility can not be realized either.

Our empirical exploration contributes to the existing study of stock market in the following dimensions. First, in establishing the empirical tests, we include trading volume and its lagged terms in the standard GARCH and ARCH models to explain the conditional volatility and find that trading volume and its lagged terms have rather satisfactory power of explaining the conditional volatility. Second, we believe this paper is the first to document the dynamics between trading volume and volatility. By discovering the negative contribution of trading volume to the subsequent volatility, our findings enrich the conventional understanding of trading volume and volatility. Third, by supporting the IDT hypothesis, our study sheds lights on how stock market trading activities happen and enhance our understanding about what is the general motivation behind trading activities. Fourth, our additional tests find further evidence for the hypothesis that informed trades are concentrated in the medium volume size category, and that empirically there exists a most information-intensive volume size for each stock.

This rest of this paper is organized as follows. In the next section, we introduce the data and the major methodologies. In Section C we will report the main empirical results and the results of robustness check. In Section D we do additional tests to examine further the relation between volume and subsequent stock volatility. Section E concludes.

B. Data Description and Methodology

B.1 Data description

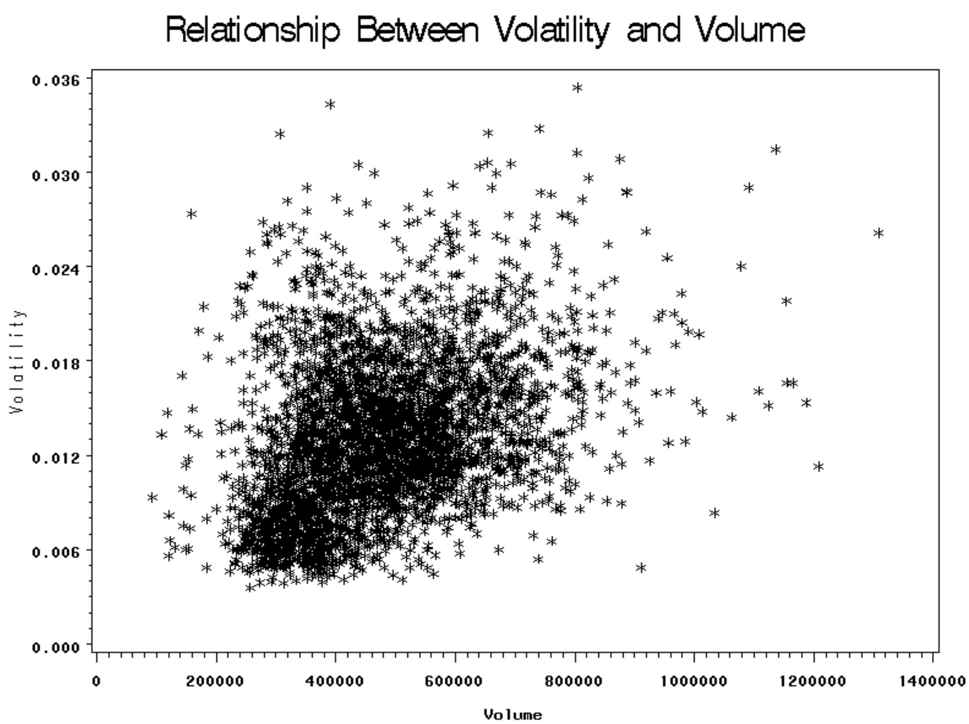
We gather our data from four primary sources. We use CRSP as our source for the U.S. equity data and PACAP (Pacific-Basin Capital Markets Research Center at the University of Rhode Island College of Business) as our source for international data. In our additional tests we also need to construct the proxies for the degree of informed trading, and we choose the average bid-ask spread and the average number of analysts following a stock as the proxy for the degree of asymmetric information, so from the I/B/E/S dataset we collect data about the monthly number of analysts who provide

I/B/E/S with end-of-fiscal-year earnings forecast for the current year, and from the TAQ database we get the daily open bid-ask spread data.

We obtain daily firm returns, shares traded and shares outstanding for the January 1988 to December 2001 sample period from the CRSP database. We also obtain the market capitalization data for each firm from the database. Totally we get the daily data for 1789 firms' from the U.S. stock market. Table 1 is a summary statistics of the US raw sample. Here volume is the event-adjusted trading volume and we use the GARCH(1,1) model to forecast the daily conditional volatility for each stock. We get the three daily series of average by summing the variables cross-sectionally.

To get a preliminary impression of the relation between trading volume and volatility, in the Diagram 1, we plot the cross-sectional daily average of volatility over the cross-sectional daily average of raw trading volume for the U.S. market. From the diagram we know that the volatility tends to be high when there is a high trading volume, indicating a positive contemporaneous correlation between volatility and trading volume[1].

Diagram 1: Trading Volume and Volatility



[1] Taking turnover as the trading volume measure, we can also plot the similar diagram, but this positive contemporaneous relation between trading volume and volatility is less obvious by appearance.

Table 1: Summary statistics of the daily cross-sectional average of volume, return and volatility

The volume here is the cross-sectional average of daily event-adjusted volume (shares traded). We use this raw trading volume measure to give a direction description of trading volume. The volatility here is the cross-sectional average of the standard GARCH(1,1)-forecasted conditional daily volatility. Return is the cross-sectional average of daily returns.

Variables	Sample Period	Obs. No.	Sample Stocks	Mean	STD	Median	Min	Max
Volume	01/04/88-12/31/01	3533	1789	476693.8	148947.3	459550.1	93084.3(12/24/90)	1308434(09/21/01)
Return	01/04/88-12/31/01	3533	1789	0.00099	0.00617	0.00142	-0.0481(10/27/97)	0.0312(03/16/00)
Volatility	01/04/88-12/31/01	3533	1789	0.01314	0.00498	0.01261	0.00361(07/05/90)	0.0353(01/06/99)

			Serial Correlation at lag					
Variables	**Skewness**	**Kurtosis**	**1**	**2**	**3**	**4**	**5**	**20**
Volume	0.806	1.276	0.761	0.656	0.621	0.63	0.639	0.564
Return	-0.815	6.426	0.211	0.036	0.077	0.077	0.03	0.026
Volatility	0.6479	0.4677	0.847	0.774	0.738	0.721	0.701	0.596

We then divide the shares traded by corresponding shares outstanding and we get the turnover measure, which we use as the proxy of trading volume in the empirical tests. As Chan and Fong (2000) indicate, not only the number of trades but also the trade sizes are significant in the volatility and volume relation. Thus we choose turnover as our trading volume measure, which takes both the size of trade and the number of trades into consideration. See also Wang(1994) and Lo and Wang (2000) for a theoretical justification for using turnover in our setting, instead of other volume metrics.

We aggregate the daily return in the usual way to form weekly return using Wednesday as the week's end. Following from Lo and Wang (2000), our weekly turnover is the sum of five daily turnovers.

We then form five size-based portfolios by sorting firms on their market capitalization and we form five volume-based portfolios by sorting firms on their turnover. For the daily data the sorting is re-performed every day and for the weekly data the sorting is re-performed every week. The portfolio returns are a capitalization-weighted average of the component firm returns. A portfolio's turnover is the equally weighted average of the individual firm turnovers for the firms that comprise the portfolio.

We form size-based portfolio to check the effect of market capitalization on the relation between trading volume and volatility, because as pointed out by a large amount of literature, large stocks and small stocks tend be different in the degree of asymmetric information, and this difference may lead to different dynamics of volatility and trading volume. Indeed LMSW do find the return dynamics is related to firm size to some degree. And as GKM documented, stocks experiencing unusually high (low) trading volume over a period of one day to a week tend to appreciate (depreciate) over the course of the following month, we then assume that trading volume may also have some effect on the dynamics between trading volume and volatility. Thus we form volume-based portfolio to check our conclusions are sensitive to the magnitude of stock trading volume.

We don't detrend the turnover series as CS has done. First because our sample period is not so long, the time trend in our turnover series is not apparent. Second, as shown in Morgan (1976) and Lamoureux and Lastrapes (1990), trading volume has a good explanatory power of return variance, thus we use the GLS method to estimate our regression model specification and take the trading volume as the instrumental variable, and we believe the variation in trading volume series does not bring much inconvenience in our estimation, indeed we check the normality of some residual series of our regressions, it seems the heteroscedasticity is not a serious problem. Third, by using the raw turnover series, we do not lose any information that trading volume may contain, and this will facilitate our digging out the relation between trading volume and volatility. We will introduce our models in detail in the next subsection. Actually in unreported tests, we do detrend the turnover series as LMSW has done and we find that with the resulting series, our conclusion still holds.

We get the international market data from the PACAP database. For the Japan market we get the daily data from January 1989 to December 2001, and for the China market we get the daily data from April 1994 to December 2003. For the index testing we also collect the daily data of a group of individual firms that comprise a major large-cap market index, the DJIA stock market index.

We need to say a few more words about the data we use. Although CGW does make a conjecture that low-frequency dependencies in volume data may make the associations between volume and other variables difficult to find in long-horizon data, none of the theoretical papers mentioned above (BEO, CGW, or Wang (1994)) specifies a time interval over which their hypothesized relations hold, thus we haven't much existing theoretical guidance about the choice of horizon of the data. Generally speaking, when people have private information about the assets, they choose to trade, and according to the Efficient Market Hypothesis, the information should be instantly incorporated into asset prices, thus we should expect that the potential associations between volume and other financial variables can be more easily found in short-horizon data than in the long-horizon data. Table 2 reports a summary of the data used in the previous most relevant research. We do not claim that it is a complete set, rather we just want to provide some support for the data horizon we choose. From the table we can see that both daily and weekly data were used in previous studies, and also a number of papers use both the individual stock data and the index or stock portfolio data. Thus in our empirical tests, we use both daily and weekly data, individual and portfolio and index data. We use daily data to avoid that long-horizon data may disguise the relation between trading volume and other financial variables, and we use weekly data to control the effect of daily volume and return fluctuations that may have less direct economic relevance.

Table 2: Selected volume studies regarding the data horizon used

	Daily Data		**Weekly Data**	
	Individual	Index/Portfolio	Individual	Index/Portfolio
Chen, Firth and Rui (2001)	√			
Morgan (1976)			√*	
Gervais, Kaniel and Mingelrin (1998)	√	√	√	√
Lamoureux and Lastrapes (1990)	√			
Stickel and Verrecchia (1994)	√			
Epps (1976)	√			
Gallant, Rossi and Tauchen (1992)		√		
Connolly and Stivers (2003)			√	√
Conrad, Hameed, and Niden (1994)			√	

* Morgan (1976) used 4-day data, which we regard it as weekly data.

B.2 Methodology

In this subsection, we develop the empirical methodologies to detect the possible relation between trading volume and volatility. Basically we use two series of models. The first series of models is the regression model. In the regression models , we use the absolute value of return as our volatility measure. This measure is very common in the finance literature to be used as the volatility measure, such as Ahn, Bae and Chan (2001). We also use other kind of volatility measure such as absolute price changes in our robustness check section and the results are mainly the same. Specifically the models will have the following specification,

$$Abs\,\mathrm{Re}t_t = \alpha_0 + \sum_{i=1}^{m} \alpha_i Abs\,\mathrm{Re}t_{t-i} + \sum_{j=1}^{n} \beta_j TO_{t-j-1} + \varepsilon_t \quad (1),$$

Where $Abs\,\mathrm{Re}t_t$ is the absolute value of the individual stock return. It can also denote the absolute value of portfolio return and index return according to the data the model is describing. $Abs\,\mathrm{Re}t_{t-i}$ is the corresponding lagged term. TO_{t-j-1} is the turnover of t-j-1 period, and similarly it can be the turnover of individual stocks, portfolios and indexes. All the above variables can be weekly or daily, and m and n are the total number of lagged terms and they may vary according to the data we use. α's and β's are the estimated parameters.

The second series of models is the modified GARCH and ARCH models. Since the GARCH(1,1) and ARCH model series are widely used in the literature to capture the characteristics of volatility, naturally we could try to use these models to capture the relation between volume and volatility. Although theoretically there are different forms of GARCH model series to capture different aspects of volatility, as Bollerslev, Chou and Kroner (1992) point out, the GARCH(1,1) model often appears adequate in practice. So we construct our model primarily based on GARCH(1,1) model. We also modify the standard ARCH model to check the effect of trading volume on volatility.

Specifically the modified GARCH(1,1) model will have the following representation,

$$\begin{aligned} ret_t &= \mu + \varepsilon_t \\ \varepsilon_t &\mid (TO_t, TO_{t-1}, \ldots, TO_{t-j}, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) \sim Distr(0, h_t) \quad (2), \\ h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 h_{t-1} + \sum_{j=1}^{N} \beta_j TO_{t-j-1} \end{aligned}$$

where ret_t is the return of period t, μ is the mean return, and ε_t is the return residual of period t, h_t is the conditional volatility of period t, and TO_{t-j-1} has the same definition as in the first series of models. Similarly these variables can be daily or weekly, according to the data which the models are describing. α's and β's are the estimated parameters. And the modified ARCH model is as follows,

$$\begin{aligned} ret_t &= \mu + \varepsilon_t \\ \varepsilon_t &\mid (TO_t, TO_{t-1}, \ldots, TO_{t-j}, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) \sim Distr(0, h_t) \quad (3) \\ h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \sum_{j=1}^{N} \beta_j TO_{t-j-1} \end{aligned}$$

Our GARCH model here is very alike to the one in Lamoureux and Lastrapes (1990). The major difference is that we use not only the contemporaneous term of volume but also its lagged terms to explain the conditional volatility. And they only test it on daily return of a small specific sample of individual stocks. We extend the test to daily return, weekly return of a large sample of individual stocks, portfolios and stock indexes.

To further examine the relation between the trading volume and subsequent volatility, we also use some other models to do the empirical tests. We will introduce these models in Section D.

In the next section we report our estimation results with these model specifications.

C. Main Empirical Results

In this section we do empirical test with the models and data introduced in the last section. We estimate the model series (1), (2) and (3) and we are interested in the coefficients of the trading volume measure, that is, the β's in these models. As predicted in the previous literature, trading volume contributes positively to the contemporaneous volatility, thus we expect the estimated β_1's in these models, i.e., the coefficient of the contemporaneous trading volume term, are significant and positive. As for the estimated coefficients of the lagged trading volume term, β_2's in these models, IDT hypothesis and LDT hypothesis provide different predictions. IDT hypothesis tends to assume trading happens mainly because of arrivals of new information, and it predicts that trading volume will contribute negatively to the subsequent volatility. LDT hypothesis holds that trading happens because of investors' non-informational needs, among which liquidity is a major one. LDT hypothesis predicts that trading volume and subsequent volatility are negatively correlated. To explore which of the two hypotheses are more pervasive and powerful in explaining trading activities, we move on to test these predictions with the stock market data.

Model series (1) are estimated by GLS, and we take the contemporaneous volume term as the instrumental variable. We choose GLS method instead of more complicated estimation procedures because this method provides consistent and unbiased parameter estimates under certain conditions and it is easy to perform and simple to understand. As indicated by Morgan (1976) and Lamoureux and Lastrapes (1990), trading volume does well in explaining the return variance, thus we believe our situation does satisfy the conditions under which GLS works. Indeed we check the normality of several residual series from our regressions, and the problem of heteroscedasticity is not serious. Model series (2) and (3) are estimated by FIML. FIML method provides approximate t statistics and this may cause problem when the sample is not large enough. However, we don't think this is a serious problem in our case since our sample is relatively large, and for the daily data we even have 3533 observations.

In the next several subsections, we first report our test results with the individual stock data, then we report our test results with different portfolios and market index. In each subsection, we report the results from both the daily data and weekly data. To save space, we put some of the results in the appendix.

C.1 Empirical results with individual stocks

We run regressions with model series (1) for each stock in our sample. Due to the large number of regressions, we don't report the result of each regression. Instead we merely report the summary of the regression results.

Table 3, Group 1, 2 and 3, report summary results from estimating model series (1) for the daily individual stock sample. In group 1, we take m=3 and n=3. We take m=3 because when running regressions with more lagged terms of volatility, we find in most regressions the estimated coefficients of the fourth and larger lagged terms are statistically insignificant. To control the effect of different model specifications on our results, we also run another two groups of regressions. In group 2, we take m=3 and n=2, and in group 3, m=2 and n=3.

Table 3: Volatility and trading volume regression results summary

The table reports the results from estimating three variations of the following model,

$$Abs\,\mathrm{Re}\,t_t = \alpha_0 + \sum_{i=1}^{m} \alpha_i Abs\,\mathrm{Re}\,t_{t-i} + \sum_{j=1}^{n} \beta_j TO_{t-j-1} + \varepsilon_t .$$

Where $Abs\,\mathrm{Re}\,t_t$ is the daily return of individual stocks from the US sample, and TO_{t-j} is the daily turnover of corresponding individual stocks. We run three groups of regressions, in the first group we choose m=3 and n=3. We choose m=3 because when running the regressions with more lagged terms of volatility, we find in most regressions the estimated coefficients of the fourth lagged term of volatility are statistically insignificant. To control the effect of different model specification on our regression results, we also run another two groups of regressions. We take m=3 and n=2 in group 2, and m=2 and n=3 in group 3. The coefficients are estimated by GLS and we take the TO_t term as the instrumental variable. We run regressions for all the individual sample stocks. We report the summary results in Group1, 2, 3. In each group, Column 1 reports the total number of regressions. The first row in Column 2 in each group reports the number of regressions in which the estimated β_1 is positive and the estimated β_2 is negative, the second row in Column 2 in each group reports the number of regressions in which the estimated β_1 is positive and the estimated β_2 is also positive. Column 3 reports the ratios of these numbers to the total number of regressions. Column 4 reports the similar number as in Column 2 but we take the statistical significance into consideration. Column 5 reports the corresponding ratios.

	Column 1	Column 2	Column 3	Column 4	Column 5
Group 1	Regression with three volatility terms and three volume terms				
	1789	1394	77.9%	1136	63.5%
	1789	262	14.6%	110	6.1%
Group 2	Regression with three volatility terms and two volume terms				
	1789	1393	77.9%	1156	64.6%
	1789	262	14.6%	110	6.1%
Group 3	Regression with three volatility terms and three volume terms				
	1789	1397	78.1%	1153	64.4%
	1789	258	14.4%	109	6.1%

In Table 3 the first row in Column 4 in each group reports the number of regressions in which the estimated β_1 is positive and significant and the estimated β_2 is negative and significant, and the next column in the same row reports the percentage ratio of this number to the total number of regressions. The second row in Column 4 in each group report the number of regressions in which both the estimated β_1 and β_2 are positive and significant, and the next column in the same row reports the percentage ratio of this number to the total number of regressions. We choose 5% as our confidence level. From Column 5 we note that in the three groups of regressions, these two ratios are a sharp contrast, 63.5% to 6.1% in Group 1, 64.6% to 6.1% in Group2 and 64.4% to 6.1% in Group 3. The sharp contrast in each group at least show that among individual stocks, under most cases trading volume contributes positively to the contemporaneous volatility and contributes negatively to the subsequent volatility, which tends to support that the IDT hypothesis is more pervasive and powerful in explaining trading activities in the stock market. The contrast is consistent among the three groups of regressions, which indicates that our results here are not sensitive to model specifications. For the weekly data, we also run the same regression as in group 2 and we summarize the regression result as we did above, the percentage ratio contrast is 60.4% to 3.4%. The conclusion is largely the same.

C.2 Volume-based portfolios tests

Different volume may have different effects on the dynamics of volume and volatility found in the last subsection. Indeed, Gervais, Kaniel and Mingelgrin (1998) examine the question of whether trading activity (measured by trading volume) contains information about future evolution of stock prices. They find that stocks experiencing unusually high trading volume over a period of one day to a week tend to appreciate over the course of the following month. So we group our US sample into five volume portfolios by sorting the stocks over their trading volume. For the daily portfolio, the sorting is re-performed once a day, and for the weekly portfolio, the rebalancing is re-performed every week. In case that some outliers may influence the largest-volume portfolio and the smallest-volume portfolio and bias our conclusion, we allocate more stocks to these two volume portfolios to control the effects of outliers.

Table 4 and Table 5 report the results of estimation model series (1) and model (2) with the weekly volume-based portfolio data. From Table 5 we can see that among the regression results for the five volume-based portfolios, the estimated β_1 's are consistently significant and positive, and the estimated β_2 's are significant and negative except that the estimated β_2 for portfolio 3 is not significant (with a p-value of 0.149). From Table 6 in the Appendix, we can also see that in a different model specification, the estimated coefficient of contemporaneous volume β_1 is still positive and significant among all the five volume-based portfolios, and the estimated coefficient of lagged volume term β_2 is negative and significant. All these results provide further evidence that the contemporaneous volume and volatility are positively correlated, and that volume and subsequent volatility are negatively correlated.

Table 4: Weekly volume-based portfolios regression results

This table reports the results from estimating the following model with the five volume-based portfolios data,

$$Abs\,\mathrm{Re}\,t_t = \alpha_0 + \alpha_1 Abs\,\mathrm{Re}\,t_{t-1} + \alpha_2 Abs\,\mathrm{Re}\,t_{t-2} + \beta_1 TO_t + \beta_2 TO_{t-1} + \varepsilon_t$$

where $Abs\,\mathrm{Re}\,t_t$ is the volatility measure of the portfolios in period t, TO_t is the portfolio trading volume measure in period t, which is the equal-weighted average of the weekly turnover of the component stocks in the portfolio. The coefficients are estimated by GLS, with t-statistics in the next column.

	Portfolio 1 (Largest vol.)		**Portfolio 2**		**Portfolio 3**		**Portfolio 4**		**Portfolio 5 (Smallest Vol.)**	
	Coefficients	t. Stat.	Coefficients	t. Stat.	Coefficients	t. Stat.	Coefficients	t. Stat.	Coefficients	t. Stat.
Inter.	0.01743	8.09	0.00138	0.54	-0.00416	-1.89	-0.00691	-3.8	0.00147	1.33
α_1	0.17389	4.77	0.08443	2.23	0.06176	1.62	0.14713	3.81	0.05168	1.36
α_2	0.10269	2.81	0.09815	2.69	0.11929	3.32	0.12551	3.55	0.10157	2.74
β_1	0.000171	10.43	0.00098	10.97	0.00161	12.19	0.00307	15.1	0.00468	12.08
β_2	-0.00019	-11.22	-0.00044	-3.8	-0.00026	-1.46	-0.00059	-2.08	-0.0017	-3.33

Table 5: Modified GARCH model estimation results with weekly volume based portfolio data

The table reports the results from estimating the modified GARCH model with five weekly volume-based portfolios. The modified GARCH model has the following specification,

$$ret_t = \mu + \varepsilon_t$$

$$\varepsilon_t \mid (TO_t, TO_{t-1}, \ldots, TO_{t-j}, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) \sim Distr(0, h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 h_{t-1} + \beta_1 TO_t + \beta_2 TO_{t-1}$$

where μ is the capitalization-weighted portfolio return, TO_t is the turnover of period t, h_t is the conditional volatility. The coefficients are estimated by FIML, with t-statistics in the next column. These approximate t-statistics under FIML do make sense since our sample is large enough (722 observations used in the estimation).

	Portfolio 1 (Largest)		**Portfolio 2**		**Portfolio 3**		**Portfolio 4**		**Portfolio 5 (Smallest vol.)**	
	Coeff. Est.	T stats.	Coefficients Est.	T stats.	Coefficients Est.	T stats.	Coefficients Est.	T stats.	Coefficients Est.	T stats.
α_0	0.0000170	1.89	0.0000007	0.06	-0.0000300	-2.84	-0.0000100	-3.35	-0.0000040	-3.38
α_1	0.0553080	4	0.0413120	3.79	0.0328340	2.7	0.0372160	3.63	0.0346830	3.71
α_2	0.9291090	48.29	0.9448490	66.21	0.9411860	47.43	0.9427590	63.28	0.9557770	74.62
β_1	0.0000033	5.42	0.0000220	4.63	0.0000310	4.91	0.0000450	6.44	0.0000730	6.35
β_2	-0.0000034	-5.65	-0.0000200	-4.33	-0.0000300	-4.4	-0.0000400	-6.06	-0.0000700	-6.19
μ	0.0060100	7.39	0.0036210	5.44	0.0024770	4.24	0.0014830	2.97	0.0006820	1.74

We also estimate model (1), (2) and (3) with the daily volume-based portfolio data. And we also estimate model (3) with the weekly volume-based portfolio data. The results in these estimations consistently support the contemporaneous positive relation between volume and volatility, and negative interrelation between volume and subsequent volatility, thus indicating the IDT hypothesis can explain trading activities more pervasively than the LDT hypothesis. To save space the estimation results of daily volume based portfolios are put in the Appendix. Table 7, Table 8 and Table 9 in the Appendix report the model (1), (2) and (3) estimation results with daily volume-based portfolio data, and Table 6 reports the model (3) estimation result with the weekly volume-based portfolio data.

C.3 Size-based portfolios tests

In this subsection, we group the US sample stocks according to their market capitalization. There are a large amount of papers which aim at capturing the different aspects between large-size firms and small-size firms. As for the aspect of the influence of firm size on the dynamics of trading volume and other financial proxies, LMSW finds that stocks of smaller firms show a tendency for return continuation following high-volume days, and larger firms show almost on continuation and mostly return reversal following high-volume days. This indicates that firm size may have influence on the relation between trading volume and subsequent return dynamics. Thus we assume that firm size may have some effect on the relation between trading volume and volatility.

We sort the sample stocks by their capitalization and equally allocate the sample stocks into five size-based portfolios. The sorting is re-performed every day for the daily data and every week for the weekly data.

Table 10 and Table 11 report the estimation results of model (1) and model (2) with weekly size-based portfolio data. From Table 11, we can see that, among the five portfolio regressions the estimated β_1's are consistently positive and significant and the estimated β_2's are consistently negative and significant. In table 11, the estimated β_1's for the five size-based portfolios are all positive and significant, however, not all the five portfolios' estimated β_2's are negative and significant. But we can see that in portfolio 2, the estimated β_2 has a negative sign, although it is statistically insignificant. Overall, we can still conclude that the estimation results with the weekly size-based portfolio data support the dynamics between trading volume and volatility found in the former sections. And the effect of firm size on this dynamics is not apparent.

Table 10: Weekly size-based portfolios regression results

This table reports the results from estimating the following model with the five weekly size-based portfolios data,

$$Abs\,\mathrm{Re}\,t_t = \alpha_0 + \alpha_1 Abs\,\mathrm{Re}\,t_{t-1} + \alpha_2 Abs\,\mathrm{Re}\,t_{t-2} + \beta_1 TO_t + \beta_2 TO_{t-1} + \varepsilon_t$$

Where $Abs\,\mathrm{Re}\,t_t$ is the volatility measure of the portfolios in period t, TO_t is the portfolio trading volume measure in period t, which is the equal-weighted average of the weekly turnover of the component stocks in the portfolio. The coefficients are estimated by GLS, with t-statistics in the next column.

	Portfolio 1 (largest size)		**Portfolio 2**		**Portfolio 3**		**Portfolio 4**		**Portfolio 5(Smallest size)**	
	Coeff. Est.	T stat.	Coeff. Est.	T stat.	Coeff. Est.	T stat.	Coeff. Est.	T stat.	Coeff. Est.	T stat.
α_0	0.011960	7.41	0.008060	6.07	0.009260	7.22	0.001410	0.92	0.006620	4.39
α_1	0.160690	4.29	0.234680	6.18	0.189470	5.13	0.266030	7.11	0.174980	4.53
α_2	0.080740	2.21	0.076870	2.04	0.166550	4.47	0.175170	4.73	0.203190	5.32
β_1	0.000259	10.22	0.000257	13.29	0.000392	10.7	0.000765	14.95	0.000317	4.66
β_2	-0.000264	-9.88	-0.000227	-9.84	-0.000404	-10.46	-0.000559	-9.54	-0.000240	-3.01

Table 11: Modified GARCH model estimation results with weekly size-based portfolio data

The table reports the results from estimating the modified GARCH model with five weekly size-based portfolios. The modified GARCH model has the following specification,

$$ret_t = \mu + \varepsilon_t$$

$$\varepsilon_t \mid (TO_t, TO_{t-1}, \ldots, TO_{t-j}, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) \sim Distr(0, h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 h_{t-1} + \beta_1 TO_t + \beta_2 TO_{t-1}$$

where μ is the capitalization-weighted portfolio return, TO_t is the turnover of period t, h_t is the conditional volatility. The coefficients are estimated by FIML, with t-statistics in the next column. These approximate t-statistics under FIML do make sense since our sample is large enough (722 observations used in the estimation).

	Portfolio 1 (Largest size)		**Portfolio 2**		**Portfolio 3**		**Portfolio 4**		**Portfolio 5 (Smallest size)**	
	Coeff. Est.	T stats.	Coefficients Est.	T stats.	Coefficients Est.	T stats.	Coefficients Est.	T stats.	Coefficients Est.	T stats.
α_0	0.000010	1.76	0.000006	1.73	0.000042	3.68	0.000023	0.76	0.000005	0.46
α_1	0.036548	3.11	0.030686	4.19	0.161555	3.43	0.313453	2.65	0.207491	4.41
α_2	0.947176	46.19	0.957433	82.11	0.699853	13.23	0.578798	5.45	0.707056	13.65
β_1	0.000005	5.93	0.000000115	0.31	0.000005	4.1	0.000000521	0.64	0.000008	4.06
β_2	-0.000005	-6.15	-0.000000150	-0.38	-0.000005	-4.97	0.000000282	0.13	-0.000007	-4.15
μ	0.004166	6.4	0.003781	5.87	0.003867	6.28	0.005033	6.95	0.004219	7.35

We also estimate model (1), (2) and (3) with the daily size-based portfolio data. Table 12, Table 13 and Table 14 report the estimation results of model (1), (2) and (3) with the daily size-based portfolio data respectively. As we can see from these tables, the estimation results with these models largely support our findings about the dynamics between trading volume and volatility. And this effect of firm size on this dynamics is not apparent. We do not report the model (3) estimation results with the weekly size based portfolio data because estimations of two portfolios parameters among the five do not converge, but from the estimation results of the other three portfolios, we can still get the conclusion which is consistent with our findings in previous sections.

C.4 Total sample and market index tests

We find the same volume and volatility dynamics in the total US sample portfolio. Table 15 reports the estimation results of model (1), (2) and (3) with the weekly total sample portfolio. From the results we can get the same conclusion as we do in the previous sections. Thus we can extend our conclusion to market index. Indeed we test our finding using the DJIA index data, and we get the similar conclusion.

C.5 International evidence

We also check whether our primary results extend to non-US market. Specifically we estimated model (1) for individual stocks from other markets' sample and we summarize the results as we do in Table 4. For the Japan stock market sample, the ratio contrast in the three groups of regressions are 93.7% to 1.8%, 93.7% to 2.2%, and 93.9% to 1.8% respectively, and for the China market, 65.7% to 12.9%, 64.9% to 14.1%, and 69.0% and 9.7% respectively.

We further check our primary results with model (2) using the individual stock data from the Japan market. Among the estimations which converge, the percentage of the cases in which the estimated β_1 is positive and significant and β_2 is negative and significant is 93.2%, compared with 4.1%, the percentage of the cases in which both the estimated β_1 and β_2 are positive and significant. This again supports the contemporaneous positive correlation between volume and volatility, and the negative correlation between volume and subsequent volatility. This international evidence also gives support to the conclusion that the IDT hypothesis are more pervasive and powerful in explaining stock market trading activities than the LDT hypothesis.

C.6 Out-of-sample tests and alternative volatility measure

We also collect the sample from 1962 to 1980 from the US market to check whether our primary finding is only a characteristic of the US market during the specific sample period. Regression results with model series (1) for individual stocks are not materially different from what we report in Table 3. And we also choose the absolute daily price change as the volatility measure, the results are largely the same. Thus we can conclude that our conclusion about the relation between trading volume and volatility are not sensitive to the sample and volatility measure we choose.

Table 15: Estimation with the whole weekly sample

This table reports the estimation results with the total weekly sample portfolio.
Panel A reports the results from estimating the following model with the whole weekly sample portfolio,

$$Abs\,\mathrm{Re}\,t_t = \alpha_0 + \alpha_1 Abs\,\mathrm{Re}\,t_{t-1} + \alpha_2 Abs\,\mathrm{Re}\,t_{t-2} + \beta_1 TO_t + \beta_2 TO_{t-1} + \varepsilon_t$$

Panel B reports the result from estimating the following model with the whole weekly sample portfolio. The coefficients are estimated by FIML, with t-statistics in the next row. These approximate t-statistics under FIML do make sense since our sample is large enough (722 observations used in the estimation).

$$ret_t = \mu + \varepsilon_t$$

$$\varepsilon_t \mid (TO_t, TO_{t-1}, ..., TO_{t-j}, \varepsilon_{t-1}, \varepsilon_{t-2}, ...) \sim Distr(0, h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 h_{t-1} + \beta_1 TO_t + \beta_2 TO_{t-1}$$

Panel C reports the result from estimating the following model with the whole weekly sample portfolio. The coefficients are estimated by FIML, with t-statistics in the next row. These approximate t-statistics under FIML do make sense since our sample is large enough (722 observations used in the estimation).

$$ret_t = \mu + \varepsilon_t$$

$$\varepsilon_t \mid (TO_t, TO_{t-1}, ..., TO_{t-j}, \varepsilon_{t-1}, \varepsilon_{t-2}, ...) \sim Distr(0, h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 TO_t + \beta_2 TO_{t-1}$$

Panel A: Total weekly sample regression result

α_0	α_1	α_2	β_1	β_2
0.0142	0.12154	0.08133	0.00036125	-0.00040939
8.2	3.34	2.27	6.51	-7.43

Panel B: Modified GARCH estimation result with total weekly sample portfolio

α_0	α_1	α_2	β_1	β_2	μ
0.00001300	0.04985900	0.92893600	0.00000789	-0.00000801	0.00398100
2.04	3.51	42.27	6.02	-6.28	6.32

Panel C: Modified ARCH estimation result with total weekly sample portfolio

α_0	α_1	β_1	β_2	μ
0.00044500	0.15693300	0.00000696	-0.00000996	0.00341200
8.77	3.51	6.87	-14.57	4.46

D. Further Analysis and Discussion of Results

In the above sections we find evidence that generally trading volume and subsequent volatility are negatively correlated. But questions such as what factors may have an influence on this negative relation still remain. In this section, we try to address this issue from several dimensions. In the first subsection we will check what factors may affect the relation between volume and subsequent volatility. The second subsection we will check whether the relation between volume and subsequent volatility is non-linear, and in the last subsection we investigate the possible asymmetric effect of volume on subsequent volatility.

D.1 Factors affecting the relation between trading volume and subsequent volatility

In this subsection, we try to look for the factors that may influence the relation between volume and subsequent volatility. In previous sections, we only show that volume and subsequent volatility are correlated, but we have no idea about what factors may influence this correlation. Here we will develop models to examine the possible factors. Basically our idea about constructing the econometric model here was enlightened by the one in LMSW. We choose the estimated β_2 for individual sample stocks, i.e., the coefficient of the lagged volume term in model (1) as the dependent variable and use other relevant variables to explain it. Specifically our model has the following representation,

$$\beta_2 = \gamma_0 + \gamma_1 F + \varepsilon \quad (4),$$

where β_2 is the estimated coefficient in model (1) for individual sample stocks, and F denotes possible factors that may influence the relation between volume and subsequent volatility. Here we choose four proxies, the average volume, the average market capitalization, the average number of analysts following, and the average relative bid-ask spread.

We choose the average volume (denoted by AvgTO) as one possible factor, because as LMSW conjectures, " for less liquid stocks, high volume is associated with a higher price impact and a larger subsequent return reversal than for more-liquid stocks". Thus if we choose the average volume as the proxy for liquidity, we can expect the estimated β_2's for stocks with larger average volume, that is, more liquid stocks, are more negative than those for stock with smaller average volume ,that is, less liquid stocks. Therefore when the F in model (4) denotes the average volume of a stock, we will expect the estimated γ_1 to be positive. The rest of the three proxies have something related to the information asymmetry. Some papers such as Lo and Mackinlay (1990) argues that larger firms have a lower degree of information asymmetry, so we choose the average market capitalization (denoted by AvgCAP) of stocks as a candidate proxy for the degree of information asymmetry. Since more information asymmetry will lead to more negative β_2, we expect γ_1 to be positive when we choose average market capitalization as the proxy for information asymmetry. The bid-ask spread is also regarded as a proxy for information asymmetry by many market microstructure papers, such as Lee, Mucklow, and Ready (1993). So we also choose the average relative open bid-ask spread (AvgBA) as a proxy for information asymmetry. And for robust check purpose we also choose the average numbers of analysts (AvgNUM) following a stock as a proxy for information asymmetry. Recent studies by Brennan and Subrahmanyam (1995) and Easley, O'Hara, and Paperman (1998) find that firms that are followed by a large number of analysts have a lower degree of information asymmetry. Similarly we will expect γ_1 to be correspondingly negative and positive when we choose average bid-ask spread and average number of analysts following as proxy for information asymmetry respectively.

We get the daily open bid-ask spread data from the TAQ database. Since the TAQ database records begin in 1993, we estimated model (1) for individual stocks with the sub-sample from 1993 to 2001 to keep consistency. We divide the difference between bid price and ask price by the mid-point of bid-ask prices to get the relative bid-ask spread, then we define AvgBA to be the average of the daily relative bid-ask spread. We summarize the variable AvgBA and we can see that it has a mean of 0.095, a median of 0.041.

We obtain the data of the monthly numbers of analysts following stocks from the I/B/E/S database for the same sample period as our sample from CRSP. We then define AvgNUM to be the average monthly number of analysts over the sample period. Summary statistics for AvgNUM indicate that it has an average value of 8.11, with the median 4.96, and minimum value 1, and maximum value 40.89.

As LMSW did in their paper, we also adopt an ordinal transformation of the variables, AvgTO, AvgCAP, AvgBA and AvgNUM. That is, we order the sample stocks in an ascending order according to the proxy and assign a rank of one to the first firm and a rank which equals to the total number of sample stocks to the last firm. We divide the rank by the corresponding total number of sample stocks to get an ordinal variable for each of the proxies. As LMSW argued, "this monotonic transformation preserves the intuition of the differences between low and high information asymmetry without reading too much into the specific differences in magnitude". We denote the resulting ordinal variables by OrdTO, OrdCAP, OrdBA and OrdNUM respectively.

Although our methodology here appears to be of little difference with that in LMSW, there is one major improvement in the estimation process. LMSW estimated the following model specification to get C_2,

$$R_t = C_0 + (C_1 + C_2 V_t) R_{t-1} + \varepsilon_t,$$

where R_t and R_{t-1} denote the return of period t and period t-1 respectively, and V_t is their period t volume measure. They then used the estimated C_2 as the dependent variable and the average market capitalization as independent variables to run regressions. Here we argue that there may be some inherent bias in their estimation. Since R_t and R_{t-1} are of nearly the same magnitude across large and small firms, we then expect their first order autocorrelations to be of nearly the same magnitude. For the convenience of illustration, let's assume it's a positive number. Thus for stocks which on average have a large volume, the estimated C_2's tend to be small, and for stocks which on average have a small volume, the estimated C_2's tend to be large. Unfortunately, according to our sample, the average volume and the average market capitalization are positively correlated, that is to say large stocks also tend to have large average volume. Indeed, even as LMSW indicated themselves, their average volume measure also increases with firm size. Thus without running regressions we can know that C_2 and the average market capitalization are negatively related and this negative relation has nothing to do with the information asymmetry, which is what LMSW tried to illustrate at the very beginning.

Thus before estimating β_2's, we divide the volume of individual stocks by their own average volume to avoid that the variation of magnitude of volume across stocks might bias our final conclusion, while keeping the variations of volume within stocks at the same time.

Table 16: Factors affecting the relation between volume and subsequent volatility

This table report the results of the regressions which aim at finding the factors that may influence the relation between volume and subsequent volatility. Basically we are estimating the following model,

$$\beta_2 = \gamma_0 + \gamma_1 F + \varepsilon$$

where F can denote OrdTO, OrdCAP, OrdNUM, and OrdBA. We estimate these models with OLS.

Panel A: Examining the liquidity effect on the relation between volume and subsequent volatility

$$\beta_2 = \gamma_0 + \gamma_1 OrdTO + \varepsilon$$

Variable	Estimate	t Value	Pr > |t|
γ_0	0.000146	2.19	0.029
γ_1	-0.00358	-30.98	<.0001

Panel B: Examining the effect of average market capitalization

$$\beta_2 = \gamma_0 + \gamma_1 OrdCAP + \varepsilon$$

Variable	Estimate	t Value	Pr > |t|
γ_0	-0.00114	-13.97	<.0001
γ_1	-0.00102	-7.23	<.0001

Panel C: Examining the effect of average number of analysts following

$$\beta_2 = \gamma_0 + \gamma_1 OrdNUM + \varepsilon$$

Variable	Estimate	t Value	Pr > |t|
γ_0	-0.00175	-28.7	<.0001
γ_1	0.17558	2.34	0.0192

Panel D: Examining the effect of average bid-ask spread

$$\beta_2 = \gamma_0 + \gamma_1 OrdBA + \varepsilon$$

Variable	Estimate	t Value	Pr > |t|
γ_0	-0.0013	-24.69	<.0001
γ_1	-0.00471	-11.38	<.0001

We report the regression results in Table 16. From Panel A, C, and D we can see that the estimated γ_1's are all significant under 5% confidence level and have the expected sign, indicating that the liquidity effect and information asymmetry do influence the dynamic relation between trading volume and subsequent volatility. However, as for the regression result in Panel B, the estimated γ_1 is negative and significant, which is not consistent with our prediction. Since with the other two proxies for information asymmetry we get the consistent results, here we argue that possibly the average market capitalization is merely a crude proxy for the degree of information asymmetry.

Our regression results largely support that trading volume contributes more negatively to subsequent volatility for more liquid stocks, which is consistent with the LMSW's conjecture, and higher degree of informed trading contribute more negatively to subsequent volatility. Liquidity and information asymmetry are two factors that may influence the relation between volume and subsequent volatility.

D.2 Non-linear relation between trading volume and subsequent volatility

In this subsection we examine the possible non-linear relation between volume and subsequent volatility. The consideration of a possible non-linear relation between volume and subsequent volatility stems from Barclay and Warner (1993), which examines the proportion of a stock's cumulative price change that occurs in each trade-size category and finds evidence for that informed traders are concentrated in the medium-size category, while our previous conclusions support that when there are more informed trading, trading volume will contribute more negatively to subsequent volatility. Since Barclay and Warner (1993) argues that informed trading is more related to medium-size trades, we could conjecture that medium-size trade volume contributes more negatively to subsequent volatility than smaller-size and larger-size trade volume do. Thus trading volume and subsequent volatility could have a non-linear relationship.

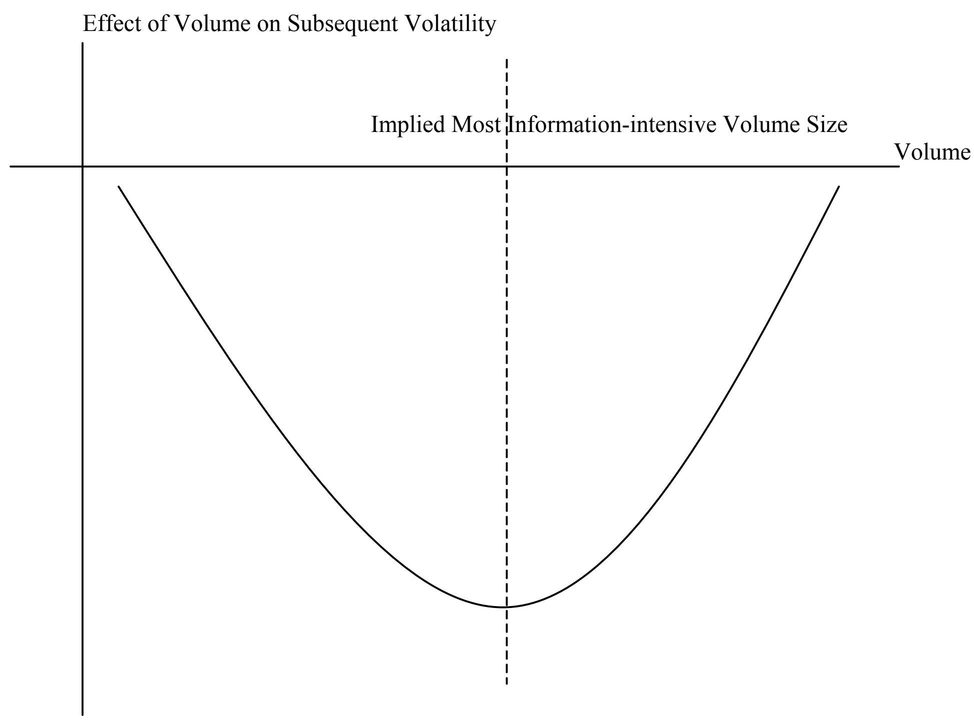
We examine this hypothesis with the following model specification,

$$\begin{aligned} &ret_t = \mu + \varepsilon_t \\ &\varepsilon_t \mid (TO_t, TO_{t-1}, ..., TO_{t-j}, \varepsilon_{t-1}, \varepsilon_{t-2}, ...) \sim Distr(0, h_t) \qquad (5) \\ &h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 h_{t-1} + \beta_1 TO_t + \beta_2 TO_{t-1} + \beta_3 TO_{t-1}^2 \end{aligned}$$

which is much alike to model (2), except an additional term of the square of the lagged volume. We also modify model (1) similarly to examine the possible non-linearity. We estimate these models with the daily and weekly data, both for individual stock and portfolios. To save space we only report the result of estimation with model (5) for the daily size-based portfolio data. Other estimation results will provided upon request.

Table 17 reports the estimation result of model (5) with the daily size-based portfolio data. We can see that under the model settings in (5), the coefficients of the contemporaneous volume and lagged volume are both consistent with the previous results. Our concern is β_3, the coefficient of the square of the lagged volume, of which the estimated results are set in boldface in Table 18. All the estimated β_3's are positive and statistically significant under the 5% confidence level, which indicates a nonlinear relation between volume and subsequent volatility. Our result indicates that indeed medium-sized volume may contain more information than smaller and larger volume, which is consistent with the findings in Barclay and Warner (1993). In Diagram 2 we illustrate the nonlinear relation between volume and its effect on subsequent volatility, which is indicated in our estimation results. We name the volume size that maximize the negative effect of volume on subsequent volatility "the most information-intensive volume size".

Diagram 2: Volume and Its Effect on Subsequent Volatility



After a further look at the result in Table 17, we can also find that from the largest size group to the smallest size one, the implied medium-size volume which maximize the possible information effect, i.e., the most information-intensive volume size in each portfolio decreases monotonically, with the exception of portfolio 5, the estimated result of which possibly is influenced by some outliers.

D.3 Asymmetric influence of trading volume on subsequent volatility

In this subsection we want to check whether the effect of trading volume on subsequent volatility are symmetric when return is positive and return is negative. A huge body of literature has documented the asymmetric effect of price movements on return volatility, which is called the leverage effect in the finance literature. By doing this test, we can check whether the negative relation between trading volume and subsequent volatility is influenced by the sign of returns, and we can know whether trading volume can partly explain the leverage effect as well.

Specifically, we examine the possible asymmetric effect of trading volume on subsequent volatility with the following model,

$$Abs\,\mathrm{Re}t_t = \alpha_0 + \sum_{i=1}^{3} \alpha_i Abs\,\mathrm{Re}t_{t-i} + \beta_1 TO_t + \beta_2 TO_{t-1} + \beta_3 D * TO_{t-1} + \varepsilon_t \quad (6),$$

where $Abs\,\mathrm{Re}t_t$ and TO_t have the same definitions as in (1), D is the dummy variable with the value 1 when the return of period t-1 is negative, and 0 otherwise. We also change the GARCH model settings in (2) similarly. We estimate the models with all the daily and weekly data, including individual stock data and portfolio data. Our concern is the estimated value of β_3. Again to save space we only report the results of regressions with the daily size-based portfolio data in Table 18 with model (6). Other estimation results will be provided upon request.

From Table 18 we can see that the estimated values of β_1 and β_2 do not change materially after considering the possible asymmetric effect of trading volume. And the coefficients of the term $D * TO_{t-1}$, β_3, are significantly positive consistently, indicating that when the return is negative, trading volume will contribute less negatively to subsequent volatility.

One plausible explanation for this asymmetric effect of trading volume is the short-selling constraints. Positive and significant β_3's indicate that when the return is negative, trading volume will contribute less negatively to subsequent volatility, which can possibly be explained in the way that when there is bad news about the stock, informed traders have the information but they are restrained from trading by short-selling constraints, thus there is less informed trading. According to our previous conclusion, we expect that when there is less information-based trading, trading volume should contribute less negatively to subsequent volatility. Thus these positive estimated β_3's are consistent with the short-selling constraint hypothesis.

Table 17: Examining the non-linear relationship between trading volume and subsequent volatility.

This table reports the following model estimation results with the daily size-based portfolio data.

$$ret_t = \mu + \varepsilon_t$$

$$\varepsilon_t \mid (TO_t, TO_{t-1}, ..., TO_{t-j}, \varepsilon_{t-1}, \varepsilon_{t-2}, ...) \sim Distr(0, h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 h_{t-1} + \beta_1 TO_t + \beta_2 TO_{t-1} + \beta_3 TO_{t-1}^2$$

where μ is the capitalization-weighted portfolio return, TO_t is the turnover of period t, h_t is the conditional volatility. The coefficients are estimated by FIML, with t-statistics in the next column. These approximate t-statistics under FIML do make sense since our sample is large enough (3533 observations used in each of the estimations).

	Portfolio 1 (Largest)		**Portfolio 2**		**Portfolio 3**		**Portfolio 4**		**Portfolio 5(Smallest)**	
	Coeff.Est	T stat.	Coeff.Est	T stat.	Coeff.Est	T stat.	Coeff.Est	T stat.	Coeff.Est	T stat.
α_0	0.000007578	5.99	0.000001583	2.82	0.000003307	3.49	0.000004414	3.59	0.000003915	3.29
α_1	0.048119000	5.53	0.116443000	7.53	0.140920000	8.94	0.178366000	9.44	0.170602000	9.11
α_2	0.925115000	76.03	0.849723000	43.83	0.799887000	40.04	0.725577000	28.02	0.715974000	23.04
β_1	0.000003320	20.65	0.000002101	13.48	0.000002936	11.97	0.000003376	9.77	0.000005219	8.37
β_2	-0.000004190	-18.36	-0.000002150	-14.62	-0.000003180	-12.14	-0.000003930	-7.76	-0.000006070	-6.85
β_3	**0.000000029**	**7.05**	**0.000000005**	**2.23**	**0.000000015**	**2.51**	**0.000000058**	**2.23**	**0.000000189**	**3.08**

Table 18: Examining the asymmetric effect of trading volume on subsequent volatility.

This table reports the following model estimation results with the daily size-based portfolio data.

$$Abs\,\mathrm{Re}\,t_t = \alpha_0 + \sum_{i=1}^{3} \alpha_i Abs\,\mathrm{Re}\,t_{t-i} + \beta_1 TO_t + \beta_2 TO_{t-1} + \beta_3 D * TO_{t-1} + \varepsilon_t$$

where $Abs\,\mathrm{Re}\,t_t$ is the volatility measure of the portfolios in period t, TO_t is the portfolio trading volume measure in period t, which is the equal-weighted average of the daily turnover of the component stocks in the portfolio, D is a dummy variable with the value 1 when the return of period t-1 is negative and 0 otherwise . The coefficients are estimated by GLS, with t-statistics in the next column.

	Portfolio 1 (Largest)		**Portfolio 2**		**Portfolio 3**		**Portfolio 4**		**Portfolio 5(Smallest)**	
	Coeff.Est	T stat.	Coeff.Est	T stat.	Coeff.Est	T stat.	Coeff.Est	T stat.	Coeff.Est	T stat.
α_0	0.00516	15.95	0.00384	17.54	0.00305	15.6	0.00231	9.86	0.00239	13.51
α_1	0.1509	9.23	0.20124	12.12	0.19357	11.8	0.20553	12.58	0.1754	10.41
α_2	0.12031	7.05	0.21721	12.81	0.25463	15.01	0.25707	15.39	0.25159	14.63
β_1	0.00038	13.31	0.000098	6.12	0.000084	4.34	0.000216	6.24	0.000083	2.99
β_2	-0.00044	-14.33	-0.00018723	-9.7	-0.00015	-6.43	-0.00024	-6.01	-0.000056	-1.45
β_3	**0.000061**	**4.2**	**0.000036**	**2.63**	**0.000077**	**4.43**	**0.000112**	**4.47**	**0.000117**	**3.19**

E. Conclusions

In this paper we examined the relation between trading volume and volatility. We find that trading volume not only contributes positively to the contemporaneous volatility, as indicated in the previous literature, but also contributes negatively to the subsequent volatility. And this dynamics relation between trading volume and volatility are consistently held among individual stocks, volume-based portfolios, size-based portfolios, and market index and in the US., Japan and China stock markets. We find this empirical regularity in both the daily data and weekly data.

No existing theory directly addresses our empirical settings and findings. We only find that some existing models have some implications about the relation between trading volume and volatility. We summarize these implications into two conflicting hypothesis, the Information-Driven-Trade (IDT) hypothesis, which predicts a negative relation between trading volume and subsequent volatility, and the Liquidity-Driven-Trade (LDT) hypothesis, which predicts a positive relation between trading volume and subsequent volatility.

Our empirical findings tend to support that the Information-Driven-Trade (IDT) hypothesis is more pervasive and powerful in explaining trading activities in the stock market than the Liquidity-Driven-Trade (LDT) hypothesis. And this is also supported by the international evidence.

Our additional tests have three interesting findings. First, liquidity and the degree of information influence the relation between trading volume and the following volatility. Second, the effect of volume on subsequent volatility and volume size have a non-linear relationship. Our empirical results indicate that there exists a most information-intensive volume size for each stock. Third, the effect of volume on subsequent volatility is not symmetric when the stock price moves up and when the stock price moves down. We attribute this asymmetry to the short-selling constraints.

The relation between trading volume and volatility need to be theorized and up to now this still remains an open question. And furthermore, there is no theoretical guidance about the horizon during which our empirical findings generally hold. Further research can go in these two directions. And another possible attempt is to examine the dynamic volume and volatility based on intra-day data.

F. Appendix

Table 7: Daily volume-based portfolios regression results

This table reports the results from estimating the following model with the five volume-based portfolios data,

$$Abs\,\mathrm{Re}t_t = \alpha_0 + \alpha_1 Abs\,\mathrm{Re}t_{t-1} + \alpha_2 Abs\,\mathrm{Re}t_{t-2} + \beta_1 TO_t + \beta_2 TO_{t-1} + \varepsilon_t$$

where $Abs\,\mathrm{Re}t_t$ is the volatility measure of the portfolios in period t, TO_t is the portfolio trading volume measure in period t, which is the equal-weighted average of the daily turnover of the component stocks in the portfolio. The coefficients are estimated by GLS, with t-statistics in the next column.

	Portfolio 1 (Largest vol.)		**Portfolio 2**		**Portfolio 3**		**Portfolio 4**		**Portfolio 5 (Smallest Vol.)**	
	Coefficients	t. Stat.	Coefficients	t. Stat.	Coefficients	t. Stat.	Coefficients	t. Stat.	Coefficients	t. Stat.
α_0	0.00771	17.63	0.00008356	0.13	-0.00147	-3.21	-0.00122	-3.95	0.000866	4.23
α_1	0.15807	9.66	0.12068	7.36	0.10192	6.11	0.11495	6.87	0.0773	4.82
α_2	0.15269	8.97	0.15101	9	0.13537	8	0.14075	8.33	0.21779	13.03
β_1	0.000227	11.42	0.00373	19.78	0.00589	20.03	0.00974	20.81	0.02015	15.46
β_2	-0.0003	-14.57	-0.00248	-12.78	-0.00284	-9.2	-0.00452	-9.14	-0.01218	-8.97

Table 8: Modified GARCH model estimation results with daily volume based portfolio data

The table reports the results from estimating the modified GARCH model with five daily volume-based portfolios. The modified GARCH model has the following specification,

$$ret_t = \mu + \varepsilon_t$$

$$\varepsilon_t \mid (TO_t, TO_{t-1}, ..., TO_{t-j}, \varepsilon_{t-1}, \varepsilon_{t-2}, ...) \sim Distr(0, h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 h_{t-1} + \beta_1 TO_t + \beta_2 TO_{t-1}$$

where μ is the average portfolio return, TO_t is the turnover of period t, h_t is the conditional volatility. The coefficients are estimated by FIML, with t-statistics in the next column. These approximate t-statistics under FIML do make sense since our sample is large enough (3533 observations used in the estimation).

	Portfolio 1 (Largest vol.)		**Portfolio 2**		**Portfolio 3**		**Portfolio 4**		**Portfolio 5 (Smallest vol.)**	
	Coeff. Est.	T stats.	Coefficients Est.	T stats.	Coefficients Est.	T stats.	Coefficients Est.	T stats.	Coefficients Est.	T stats.
α_0	0.00000081	2.1	-0.00000160	-3.91	-0.00000148	-4.34	-0.00000052	-5.09	-0.00000019	-7.1
α_1	0.04691500	7.47	0.02524800	5.71	0.03264400	6.23	0.02153900	3.99	0.02759700	6.67
α_2	0.94569700	135.76	0.96850900	176	0.95572000	140.27	0.97099500	136.84	0.97564900	288.94
β_1	0.00000220	20.22	0.00001000	13.63	0.00001900	16.1	0.00002500	16.28	0.00003400	9.39
β_2	-0.00000219	-21.15	-0.00000958	-14.2	-0.00002000	-17.41	-0.00002000	-16.65	-0.00003000	-9.24
μ	0.00172500	11.06	0.00081600	6.63	0.00016000	1.51	-0.00013000	-1.52	-0.00011000	-1.75

Table 9: Modified ARCH model estimation results with daily volume based portfolio data

The table reports the results from estimating the modified ARCH model with five daily volume-based portfolios. The modified ARCH model has the following specification,

$$ret_t = \mu + \varepsilon_t$$
$$\varepsilon_t \mid (TO_t, TO_{t-1}, ..., TO_{t-j}, \varepsilon_{t-1}, \varepsilon_{t-2}, ...) \sim Distr(0, h_t)$$
$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 TO_t + \beta_2 TO_{t-1}$$

where μ is the capitalization-weighted portfolio return, TO_t is the turnover of period t, h_t is the conditional volatility. The coefficients are estimated by FIML, with t-statistics in the next column. These approximate t-statistics under FIML do make sense since our sample is large enough (3533 observations used in the estimation).

	Portfolio 1 (largest)		**Portfolio 2**		**Portfolio 3**		**Portfolio 4**		**Portfolio 5 (Smallest)**	
	Coeff. Est.	T stat.	Coeff. Est.	T stat.	Coeff. Est.	T stat.	Coeff. Est.	T stat.	Coeff. Est.	T stat.
α_0	0.00013800	20.73	0.00000067	0.11	-0.00001000	-3.48	-0.00001000	-7.85	0.00000641	7.41
α_1	0.31382000	8.4	0.14450200	5.66	0.10653000	5.15	0.12342500	5.34	0.14047600	5.96
β_1	0.00000157	5.41	0.00002300	22.28	0.00003400	26.2	0.00004800	26.4	0.00011500	12.1
β_2	-0.00000277	-17.9	-0.00000503	-3.67	-0.00000043	-0.21	0.00000423	1.87	-0.00000615	-0.68
μ	0.00187000	10.32	0.00082500	5.95	0.00007200	0.6	-0.00019000	-1.93	-0.00018000	-2.12

Table 12: Daily size-based portfolios regression results

This table reports the results from estimating the following model with the five daily size-based portfolios data,

$$Abs\,\mathrm{Re}\,t_t = \alpha_0 + \alpha_1 Abs\,\mathrm{Re}\,t_{t-1} + \alpha_2 Abs\,\mathrm{Re}\,t_{t-2} + \beta_1 TO_t + \beta_2 TO_{t-1} + \varepsilon_t$$

where $Abs\,\mathrm{Re}\,t_t$ is the volatility measure of the portfolios in period t, TO_t is the portfolio trading volume measure in period t, which is the equal-weighted average of the daily turnover of the component stocks in the portfolio. The coefficients are estimated by GLS, with t-statistics in the next column.

	Portfolio 1 (largest size)		**Portfolio 2**		**Portfolio 3**		**Portfolio 4**		**Portfolio 5 (Smallest size)**	
	Coeff. Est.	T stat.	Coeff. Est.	T stat.	Coeff. Est.	T stat.	Coeff. Est.	T stat.	Coeff. Est.	T stat.
α_0	0.00523	16.15	0.00386	17.64	0.00316	14.39	0.00248	10.68	0.00252	10.44
α_1	0.15033	9.17	0.20111	12.1	0.19794	12.23	0.20567	12.55	0.17658	10.43
α_2	0.11565	6.78	0.21679	12.78	0.25266	15.59	0.25641	15.31	0.25344	14.72
β_1	0.000379	13.25	9.72E-05	6.05	0.0001521	4.94	0.000208	5.99	8.17E-05	2.89
β_2	-0.00041	-13.74	-0.00017	-9.33	-0.0002135	-7	-0.00022	-5.6	-0.00016	-0.81

Table 13: Modified GARCH model estimation results with daily size-based portfolio data

The table reports the results from estimating the modified GARCH model with five daily size-based portfolios. The modified GARCH model has the following specification,

$$ret_t = \mu + \varepsilon_t$$

$$\varepsilon_t \mid (TO_t, TO_{t-1}, ..., TO_{t-j}, \varepsilon_{t-1}, \varepsilon_{t-2}, ...) \sim Distr(0, h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 h_{t-1} + \beta_1 TO_t + \beta_2 TO_{t-1}$$

where μ is the capitalization-weighted portfolio return, TO_t is the turnover of period t, h_t is the conditional volatility. The coefficients are estimated by FIML, with t-statistics in the next column. These approximate t-statistics under FIML do make sense since our sample is large enough (3533 observations used in the estimation).

	Portfolio 1 (Largest size)		Portfolio 2		Portfolio 3		Portfolio 4		Portfolio 5 (Smallest size)	
	Coeff. Est.	T stats.	Coefficients Est.	T stats.	Coefficients Est.	T stats.	Coefficients Est.	T stats.	Coefficients Est.	T stats.
α_0	0.0000008	2.07	0.0000001	0.41	0.0000031	5.32	0.0000020	3.63	0.0000042	4.49
α_1	0.0487170	5.71	0.1189640	8.15	0.1812320	8.85	0.1766490	9.53	0.1839420	9.91
α_2	0.9397840	90.77	0.8531620	49.41	0.7645750	31.87	0.7373410	30.58	0.6802350	22.51
β_1	0.0000023	15.11	0.0000021	10.16	0.0000001	0.21	0.0000029	8.52	0.0000049	7.57
β_2	-0.0000023	-16.47	-0.0000020	-10.31	-0.0000001	-0.44	-0.0000027	-9.16	-0.0000100	-7.22
μ	0.0009520	7.52	0.0009580	10.02	0.0012050	12.73	0.0012360	14.19	0.0015730	18.61

Table 14: Modified ARCH model estimation results with daily size-based portfolio data

The table reports the results from estimating the modified ARCH model with five daily size-based portfolios. The modified ARCH model has the following specification,

$$ret_t = \mu + \varepsilon_t$$
$$\varepsilon_t \mid (TO_t, TO_{t-1}, \ldots, TO_{t-j}, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) \sim Distr(0, h_t)$$
$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 TO_t + \beta_2 TO_{t-1}$$

where μ is the capitalization-weighted portfolio return, TO_t is the turnover of period t, h_t is the conditional volatility. The coefficients are estimated by FIML, with t-statistics in the next column. These approximate t-statistics under FIML do make sense since our sample is large enough (3533 observations used in the estimation).

	Portfolio 1 (largest)		**Portfolio 2**		**Portfolio 3**		**Portfolio 4**		**Portfolio 5 (Smallest)**	
	Coeff. Est.	T stat.	Coeff. Est.	T stat.	Coeff. Est.	T stat.	Coeff. Est.	T stat.	Coeff. Est.	T stat.
α_0	0.0000640	17.72	0.0000400	16.05	0.0000340	15.32	0.0000210	10.79	0.0000120	5.64
α_1	0.2031020	7.13	0.3799090	10.96	0.4014000	10.94	0.4181650	11.92	0.3287710	11.13
β_1	0.0000043	9.22	0.0000004	1.43	0.0000001	0.41	0.0000018	4.02	0.0000063	9.04
β_2	-0.0000040	-17.13	-0.0000007	-12.29	-0.0000006	-6.33	-0.0000014	-10.85	-0.0000087	-5.23
μ	0.0009350	6.37	0.0010650	8.96	0.0012620	11.49	0.0014120	14.72	0.0016300	18.09

Table 6: Modified ARCH model estimation results with weekly volume based portfolio data

The table reports the results from estimating the modified ARCH model with five weekly volume-based portfolios. The modified ARCH model has the following specification,

$$ret_t = \mu + \varepsilon_t$$
$$\varepsilon_t \mid (TO_t, TO_{t-1}, ..., TO_{t-j}, \varepsilon_{t-1}, \varepsilon_{t-2}, ...) \sim Distr(0, h_t)$$
$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 TO_t + \beta_2 TO_{t-1}$$

where μ is the capitalization-weighted portfolio return, TO_t is the turnover of period t, h_t is the conditional volatility. The coefficients are estimated by FIML, with t-statistics in the next column. These approximate t-statistics under FIML do make sense since our sample is large enough (722 observations used in the estimation).

	Portfolio 1 (largest)		**Portfolio 2**		**Portfolio 3**		**Portfolio 4**		**Portfolio 5 (Smallest)**	
	Coeff. Est.	T stat.	Coeff. Est.	T stat.	Coeff. Est.	T stat.	Coeff. Est.	T stat.	Coeff. Est.	T stat.
α_0	0.0008410	7.33	-0.0001700	-2.51	-0.0002300	-2.55	-0.0001000	-2.53	0.0000053	0.27
α_1	0.1629920	2.73	0.0902610	2	0.0818340	1.69	0.1739620	3.46	0.0940940	2.01
β_1	0.0000025	1.79	0.0000360	12.56	0.0000440	5.61	0.0000610	14.25	0.0000850	9.39
β_2	-0.0000045	-5.65	-0.0000100	-4.63	0.0000016	0.21	-0.0000100	-1.87	-0.0000200	-1.89
μ	0.0059610	6.36	0.0044670	3.92	0.0023660	3.87	0.0008580	0.92	0.0004650	1.07

References

Ahn, Hee-joon, Kee-hong Bae, and Kalok Chan, 2001, Limit orders, depth and volatility: Evidence from the Stock Exchange of Hong Kong, Journal of Finance LVI No.2, 767-788.

Barclay, Michael J. and Jerold B. Warner, 1993, Stealth trading and volatility, which trades move prices?, Journal of Financial Economics 34, 281-305.

Blume, Lawrence, David Easley and Maureen O'Hara, 1994, Market statistics and technical analysis: The role of volume, Journal of Finance 49, No.1, 153-181.

Bollerslev, T., R. Y. Chou and K. F. Kroner, 1992, ARCH modeling in finance: A review of the theory and empirical evidence, Journal of Econometrics 52, 5-59.

Brennan, M., and A. Subrahmanyam, 1995, Investment analysis and price formation in securities markets, Journal of Financial Economics 38, 361-381.

Brock, William A. and Richard D. LeBaron, 1995, A dynamic structural model for stock volatility and trading volume, NBER working paper series No. 4988.

Campbell, John Y., Sanford J. Grossman and Jiang Wang, 1993, Trading volume and serial correlation in stock returns, Quarterly Journal of Economics 108, No.4, 905-939.

Chan, Kalok, and Wai-Ming Fong, 2000, Trade size, order imbalance, and the volatility-volume relation, Journal of Financial Economics 57, 247-273.

Chen, Gong-meng, Michael Firth, Oliver M. Rui, 2001, The dynamic relation between stock returns, trading volume, and volatility, The Financial Review 38, 153-174.

Connolly, Robert and Chris Stivers, 2003, Momentum and reversals in equity-index returns during periods of abnormal turnover and return dispersion, Journal of Finance, Vol. LVIII, No. 4, 1521-1555.

Conrad, Jennifer, Allaudeen Hameed, and Cathy Niden, 1994, Volume and auto-covariances in short-horizon individual stock returns, Journal of Finance 49, 1305-1329.

Epps, Thomas W., 1975, Security price changes and transaction volumes: Theory and evidence, American Economic Review 65, No.4, 586-597.

Epps, Thomas W. and Mary Lee Epps, 1976, The stochastic dependence of security price changes and transaction volumes: Implications of for the Mixture-of-distributions hypothesis, Econometrica 44, No.2, 305-321.

Easley, D., and M. O'Hara, and J. Paperman, 1998, Are financial analysts informed traders, Journal of Financial Markets 1, 175-201.

Gallant, A. Ronald, Peter E. Rossi and George Tauchen, 1992, Stock prices and Volume, Review of Financial Studies 5, No.2, 199-242.

Gervais, Simon, Ron Kaniel and Dan Mingelgrin, 2000, The high volume return premium, the Wharton School working paper series, University of Pennsylvania.

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

Lang, Larry H. P., Robert H. Litzenberger and Vicente Madrigal, 1992, Testing financial market equilibrium under asymmetric information, Journal of Political Economy 100, No.2, 317-348.

Lamoureux, Christopher G. and William D. Lastrapes, 1990, Heteroskedasticity in stock return data: Volume versus GARCH effects, Journal of Finance 45, No.1, 221-229.

Lee, C. M. C., B. Mucklow, and M. J. Ready, 1993, Spread, depths, and the impact of earnings information: An intraday analysis, Review of Financial Studies 6, 345-374.

Llorente, Guillermo, Roni Michaely, Gideon Saar, Jiang Wang, 2002, Dynamic volume-return relation of individual stocks, Review of Financial Studies 15, No.4, 1005-1047.

Lo, A. W., and A. C. Mackinlay, 1990, An econometric analysis of non-synchronous trading, Journal of Econometrics 45, 181-211.

Lo, Andrew, and Jiang Wang, 2000, Trading volume: Definitions, data analysis, and implications of portfolio theory, Review of Financial Studies 15, 1005-1047.

Morgan, I. G., 1976, Stock prices and Heteroscedasticity, Journal of Business 49, No. 4, 496-508.

Rogalski, Richard J., 1978, The dependence of prices and volume, Review of Economics and Statistics 60, No.2, 268-274.

Stickel, Scott E. and Robert E. Verrecchia, 1994, Evidence that trading volume sustains stock price changes, Financial Analysts Journal, Nov/Dec 1994, P57.

Wang, Jiang, 1994, A model of competitive stock trading volume, Journal of Political Economy 102, No.1, 127-168.