Revisiting the empirical linkages between stock returns and trading volume

Chen, Shiu-Sheng

3 February 2012

Revisiting the empirical linkages between stock returns and trading volume

Shiu-Sheng Chen[*]

Department of Economics, National Taiwan University, No. 21, Hsu-Chow Road, Taipei, Taiwan

This version: February 3, 2012

Abstract

This paper investigates whether the empirical linkages between stock returns and trading volume differ over the fluctuations of stock markets, i.e., whether the return–volume relation is asymmetric in bull and bear stock markets. Using monthly data for the S&P 500 price index and trading volume from 1973M2 to 2008M10, strong evidence of asymmetry in contemporaneous correlation is found. As for a dynamic (causal) relation, it is found that the stock return is capable of predicting trading volume in both bear and bull markets. However, the evidence for trade volume predicting returns is weaker.

---

[*] Corresponding author. Tel.: +886 2 2351 9641 ext 481; fax: +886 2 2351 1826.
E-mail address: sschen@ntu.edu.tw

1. Introduction

The relation between stock price changes and trading volume (return–volume relation) has received considerable attention in the field of finance over the past two decades. As discussed in Karpoff (1987), evidence on the return–volume relation not only enhances the knowledge on financial market structure, but also provides information to discriminate between competing theoretical models. For instance, Campbell et al. (1993) show that the return–volume relation helps solve the identification problem for testing different models.

Based on market folklore, it is generally believed that trading volume is positively associated with stock returns. As the old Wall Street adage asserts, "it takes volume to move prices". However, an early empirical study by Granger and Morgenstern (1963) fails to find a correlation between movements in a Securities and Exchange Commission composite price index and the aggregate level of volume on the New York Stock Exchange (NYSE). Succeeding studies have found more evidence of a positive correlation, but several findings remain that are inconsistent with a positive correlation. See Karpoff (1987) for a thorough survey of empirical evidence before the 1990s.

Note that what is investigated is the *contemporaneous* correlation in the studies mentioned above. Since the 1990s, the focus has moved to *dynamic* (*causal*) correlation between price changes and trading volume. That is, studies have started to examine the causal relation by asking questions such as, "does volume help forecast stock returns" or "do investors trade more when stock prices go up"? Typically, bivariate vector autoregressive (VAR) models and Granger causality tests are applied in most studies investigating the dynamic return–volume relation. Lee

2

and Rui (2002) find that trading volume does not Granger-cause stock returns using daily data from three stock markets: New York, Tokyo and London. Statman et al. (2006) use monthly data from the NYSE/AMEX from 1962 to 2001, and provide evidence that trading activity is positively related to lagged returns for many months. Griffin et al. (2007) examine data from 46 developed and developing countries, and show a strong positive relation between turnover and past returns in many markets. Using data from emerging stock markets (six Latin American markets), Saatcioglu and Starks (1998) fail to find strong evidence of stock price changes leading to volume changes. On the other hand, they find that volume seems to lead to stock price changes. Eleanor Xu et al. (2006) use a time-consistent VAR model to test the dynamic return volatility-volume relationship, and find that volatility and volume are persistent and highly correlated with past volatility and volume. Hutson et al. (2008) examine the relation between the first three moments of market returns and trading volumes, and find significant evidence that higher trading volumes trigger subsequent greater negative market return skewness. Finally, Chuang et al. (2009) use quantile regressions to investigate the causal relations between stock return and volume, and show that causal effects of volume on return are usually heterogeneous across quantiles and those of return on volume are more stable.

Departing from the framework of linear models, Hiemstra and Jones (1994) apply nonlinear Granger causality tests to examine the dynamic relation between daily Dow Jones stock returns and percentage changes in NYSE trading volume. They find evidence of significant bidirectional nonlinear causality between returns and volume. Moreover, McMillan (2007) finds that lagged volume can be used as a threshold to improve the performance of nonlinear return forecasting models.

In this paper, we add to the existing literature by examining whether the return–volume relation differs during different phases of stock market cycles, i.e., whether or not the relation is asymmetric in bull and bear stock markets. The motivation for such an asymmetric relation is intuitive. First, cyclical variations in stock returns are widely reported in the literature. See, for example, Hamilton and Lin (1996), and Perez-Quiros and Timmermann (2000). Thus, it is empirically evident that nonlinear models of the stock return with switches across bull and bear market regimes fit the data better than do linear models. Second, as the return–volume relation reflects the structure of financial markets, and various factors—such as how investors behave—may change in bull and bear markets, we should expect that the return–volume relation would also change across different phases of market cycles. For instance, in a bull market, overconfidence may grow with long-lasting past success in the market, which would result in a strong positive return–volume correlation. As shown in Hong et al. (2006), overconfidence can lead to stock market bubbles with heterogeneous beliefs and short-sales constraints.

It is worth noting that the proposed asymmetric return–volume relation here is different from the asymmetric relation proposed by Karpoff (1987). Karpoff (1987) hypothesizes that volume is positively correlated with positive price changes, and negatively correlated with negative price changes (see Figure 1 in Karpoff (1987)). That is, Karpoff (1987) proposes that the return–volume relation is fundamentally different for positive and negative price changes. However, what we aim to investigate here is that the return–volume relation is fundamentally different for bull and bear markets. Such an asymmetric return–volume relation has been examined in Ning and Wirjanto (2009) for emerging economies. Using a copula approach, they find significant and asymmetric return-volume dependence at extremes for six emerging East-Asian

4

equity markets.

In this paper, we first use Markov-switching models to identify the bull and bear regimes in the stock market and then examine the possible asymmetric return–volume relation. However, as discussed in Candelon et al. (2008), there is no consensus in the academic literature on what bear and bull markets actually are, even though recessions and booms in stock returns are widely acknowledged. One main alternative approach to identify stock market fluctuations is based on a nonparametric methodology. For instance, Candelon et al. (2008) use the Quarterly Bry–Boschan method to examine monthly stock price series. Therefore, in order to check the robustness of our main findings, we also employ the Bry–Boschan dating method as well as a naive moving average approach to identify the fluctuations in the stock markets. Moreover, as well as examining the contemporaneous return–volume correlation, we use a joint two-state Markov-switching model to study the dynamic (causal) link between returns and volume. We would like to know if lagged volume is able to predict stock returns, and vice versa.

The main empirical results can be summarized as follows. (1) A Markov-switching model identifies the bull and bear regimes in the stock market well. (2) Strong evidence of an asymmetric contemporaneous return–volume relation is found: stock returns are negatively correlated with volume in bear markets, whereas the correlation is positive in bull markets. (3) In terms of the dynamic (causal) relation, it is evident that the stock return is able to predict volume in both bear and bull markets, but the evidence for predictability of returns from volume is weaker.

The paper is structured as follows. Section 2 presents the econometric framework. Section 3 describes the data and reports the preliminary empirical results. Section 4 presents the results on the contemporaneous return–volume relation, and robustness checks are provided in section 5.

In section 6, we consider alternative identifications of stock market fluctuations. The dynamic relationship between returns and volume is examined in section 7. Finally, concluding remarks are offered in section 8.

2. Econometric framework

2.1. Characterizing the stock market fluctuations

Before exploring the link between returns and volume over stock market fluctuations, first we need to identify the recessions (bears) and booms (bulls) in the stock market. Following Maheu and McCurdy (2000) and Frauendorfer et al. (2007), we use a modified version of the Markov-switching model developed by Hamilton (1989) to identify the bear and bull stock markets.

Let r_t represent the stock return, and let V_t be trading volume. Consider a two-state Markov-switching autoregressive model of stock returns of order q (MS-AR(q)):

$$\varphi(L)r_t = \mu_{s_t} + \beta_{s_t} V_t + \varepsilon_t, \ \ \varepsilon_t \sim \text{i.i.d.} \mathcal{N}(0, \sigma_{s_t}^2), \tag{1}$$

where $\varphi(L) = 1 - \varphi_1 L - \varphi_2 L^2 - \cdots - \varphi_q L^q$ and L is the lag operator. Terms μ_{s_t} and $\sigma_{s_t}^2$ are the state-dependent mean and the variance, respectively, of r_t. The unobserved state variable s_t is a latent dummy variable set at either 0 or 1. Let $s_t = 0$ indicate the bear market and let $s_t = 1$ indicate the bull market. Stock returns are assumed to follow a two-state Markov process with a fixed transition probability matrix:

$$P = \begin{bmatrix} p^{00} & 1 - p^{11} \\ 1 - p^{00} & p^{11} \end{bmatrix}, \tag{2}$$

where:

$$p^{00} = P(s_t = 0 | s_{t-1} = 0), \tag{3}$$

$$p^{11} = P(s_t = 1 | s_{t-1} = 1). \tag{4}$$

3. Data and preliminary empirical results

3.1. Data

Using the monthly returns on the S&P 500 price index and trading volume from 1973M2 to 2008M10, this paper focuses on the US stock market. The sample period is chosen due to the availability of data on trading volume. We collected the data on the stock price index (p_t) and trading volume (vo_t) from Datastream database.[1] Stock returns are expressed in percentages: $r_t = \log(p_t/p_{t-1}) \times 100$. First, unit root tests are conducted to investigate whether these series are stationary. The results of the Augmented Dickey–Fuller (ADF) test, the Phillips–Perron (PP) test and the Elliott et al. (1996) DF-GLS test are reported in Table 1, with some other descriptive statistics. Clearly, the hypothesis of a unit root process is rejected for each series, with the exception of the trading volume series (in log). Because of the nonstationarity property of the trading volume, we consider the percentage changes in volume in our empirical analysis: $V_t = \log(vo_t/vo_{t-1}) \times 100$. According to Table 1, the percentage change in trading volume is stationary. We plot all the series in Figure 1.

[Insert Table 1 Here]

[Insert Figure 1 Here]

---

[1]The data codes in Datastream are `S&PCOMZ(PI)` and `S&PCOMZ(VO)` for the price index and trading volume, respectively.

3.2. Linear regression and Granger causality

First, we examine the return–volume relation without considering the possible asymmetry over the stock market fluctuations. The contemporaneous correlation is simply estimated as follows:

$$r_t = a + b_1 r_{t-1} + c_0 V_t + e_t. \tag{5}$$

Moreover, to consider the dynamic correlation, we conduct the following bivariate VAR model:

$$r_t = a + \sum_i b_i r_{t-i} + \sum_i c_i V_{t-i} + u_t, \tag{6}$$

$$V_t = d + \sum_i g_i r_{t-i} + \sum_i h_i V_{t-i} + v_t. \tag{7}$$

Table 2 shows the estimation results. Columns (1) and (2) indicate that the correlation between returns and volume changes is positive, but without statistical significance. Moreover, the correlation between returns and lagged volume changes is also nonsignificantly positive. The Granger causality test results are shown in columns (3) and (4). Clearly, the F-statistics suggest that Granger causal relations do not exist between volume and return, either in the direction of trading volume to returns, or from returns to trading volume. That is, in linear setting, there is no statistically significant relation between returns and volume either contemporaneously or causally. The results seem to contrast with some previous studies that find strong evidence of a return–volume relation, such as Lee and Rui (2002), who also investigate monthly data.

[Insert Table 2 Here]

However, we should notice that the results from linear regressions and Granger causality tests are sensitive to the sample period chosen. Columns (5) to (8) in Table 2 report the estimation results using data from 1973M2–1999M12, as in Lee and Rui (2002). Clearly, a significant

8

positive contemporaneous correlation between returns and volume emerges. Moreover, accord-

ing to columns (7) and (8) of Table 2, trading volume does not Granger-cause stock returns,

but returns Granger-cause volume, which is consistent with the findings in Lee and Rui (2002).

Therefore, under the linear framework, it appears that empirical results are not robust to the

choice of sample period.

Clearly, if the relation between returns and volume is asymmetric for bull and bear markets,

one would expect that the empirical tests that specify linear relations would yield statistically

weak and nonrobust results. Bearing these findings in mind, next we will examine whether the

results change when the effects of a bull/bear market are taken into account, and we will show

that the results from Markov-switching models are robust to different sample periods.

4. Contemporaneous relationship

4.1. Estimation results from Markov-switching models

Table 3 presents the estimation results for the Markov-switching model. In order to show the

superior performance of a Markov-switching model over a linear model in fitting stock return

data, we first estimate a linear model (random walk model):

$$\Delta r_t = \mu + \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.} \mathcal{N}(0, \sigma^2),$$

and an MS-AR(0) model without including regressor V_t. Columns (1) and (2) present the es-

timation results for the linear and Markov-switching models. First, it is obvious that the naive

mean/variance Markov-switching model (MS-AR(0)) yields a higher value of the likelihood

function than does the linear model. The likelihood-ratio (LR) statistic is 80.18. Therefore,

9

although the conventional LR test is not applicable because of the nuisance parameter problem, Garcia (1998) tabulates critical values for the simple two-means, two-variances model. The LR statistic is much larger than the 99%-critical value, 14.02. This finding may suggest that the nonlinear Markov-switching model performs better than the linear random walk model.

[Insert Table 3 Here]

Next, we compute the information criteria proposed by Psaradakis and Spagnolo (2003) for Markov-switching autoregressive models to determine the optimal lag length, q, of the MS-AR(q) model in equation (1). According to the Psaradakis–Spagnolo Bayesian information criterion (PSBIC), an MS-AR(1) model is chosen, while the Psaradakis–Spagnolo Akaike information criterion (PSAIC) suggests an MS-AR(6) model. In column (3) of Table 3, the MS-AR(1) model, where the regressor is the percentage changes in trading volume, identifies a regime with a higher mean ($\mu_1 = 1.15$) and lower variance ($\sigma_1 = 3.00$), and a regime with a lower mean ($\mu_0 = -0.67$) and greater variance ($\sigma_0 = 6.76$). This result conforms with the findings in Maheu and McCurdy (2000) and in Perez-Quiros and Timmermann (2000), who investigate returns from the portfolio provided by the Center for Research in Security Prices (CRSP). The high-return stable and low-return volatile states in stock returns are conventionally labeled as bull markets and bear markets, respectively. Obviously, the Markov-switching model has identified the bull and bear markets well in stock returns. Finally, the transition probabilities show that both bull-market and bear-market states are highly persistent. The bull-market regime persists, on average, for $1/(1 - p^{11}) = 1/(1 - 0.96) = 25$ months, and it is expected that the bear-market regime will persist for $1/(1 - p^{00}) = 1/(1 - 0.91) = 11.11$ months.

Figure 2 plots the smoothing probabilities of state 1 (bull market), the high-return/low-volatility state. The Markov-switching model is observed to clearly identify the stock market cycles, which, in general, are consistent with the historical perspective of the stock market. For instance, a long bull market is identified from 1983 to 1999, which is the period of "irrational exuberance", as Robert J. Shiller remarked. Moreover, bear markets are recognized after the 1973 crash (oil crisis), and the 2000 burst of the dot-com bubble. For some eras, the impacts of a crash are short lived, such as for the 1987 crash. Finally, if we estimate MS-AR(6) models, the results do not change substantially. See column (4) in Table 3.

[Insert Figure 2 Here]

Now, we turn to the contemporaneous return–volume relation. Interestingly, in the bear market, the return–volume correlation is statistically *negative* ($\hat{\beta}_0 < 0$), whereas in the bull market, the correlation is significantly *positive* ($\hat{\beta}_1 > 0$). The result is robust to different model specifications (MS-AR(1) vs. MS-AR(6)). We can further test $\beta_0 = \beta_1$ to check if the asymmetric return–volume relation is significant. For MS-AR(1) and MS-AR(6) models, the Wald statistics (p-values) are 7.62 (0.00576) and 9.21 (0.00241), respectively. The test result suggests a rejection of the hypothesis of equal coefficients between bull and bear markets.

Such an asymmetric phenomenon can be explained as follows. In the bear market, most people have already lost considerable sums of money, so they will not sell their stocks even when the price goes up. The main characteristic of the bear market is that the stock price increases while trading volume decreases. The driving force behind stock price changes is reductions in supply, rather than increases in demand. Thus, the return–volume correlation is

11

negative.

On the other hand, in the bull market, there are two possible explanations for the positive correlation. First, in the bull market, overconfidence may grow with long-lasting past success in the market, which would result in a strong positive return–volume correlation. For instance, Hong et al. (2006) have shown that overconfidence can lead to a stock market bubble with heterogeneous beliefs and short-sales constraints. Second, momentum or positive feedback trading (buy high and sell low) may also cause a positive return–volume relation under short-sales constraints or under the circumstance that momentum investors choose not to short securities (see discussions in Griffin et al. (2007)). However, generally, momentum investing is based on the belief that an extended bull market is in effect. Hence, we would expect a positive correlation between price changes and volume in a bull market.

To sum up, we have shown evidence that there exists an asymmetric contemporaneous relation between stock returns and trading volume: in bear markets, returns and volume are negatively correlated, whereas the correlation is positive in bull markets.

5. Robustness

In this section, we consider several modifications to check the robustness of our main empirical results. First, we use different measures of trading volume. Then, we check whether the results change when considering a different stock price index. Finally, we check whether the conclusion changes when different data frequency and subsample periods are used.

Detrended Volume: Following Campbell et al. (1993) and Lee and Rui (2002), an alternative measure of trading volume is detrended volume. Here, we consider the detrended volume adjusted for a linear and quadratic time trend as well as an MA(5) trend. Evidence from ADF, PP and DF-GLS tests suggest that the detrended volume is stationary. Empirical results using detrended volume are reported in columns (1) and (2) of Table 4, and show no substantial change from our main findings. It remains the case that a negative return–volume correlation is found in bear markets, whereas the correlation is positive in bull markets.

Alternative Stock Market Indicators: The S&P 500 index is used as our benchmark as it is one of the most commonly used indexes for the overall US stock market. At one time, the Dow Jones Industrial Average (DJIA) was the most renowned index for US stocks, but it now contains only thirty companies. Therefore, it is commonly agreed that the S&P 500 is a better representation of the US market. Nevertheless, we consider DJIA as a robustness check and report the result in column (3) of Table 4. Because of the availability of the volume data, the sample period for DJIA is from 1985M1 to 2008M10. Clearly, the asymmetric return–volume relation remains, which suggests that our main empirical results are robust for different stock-price indexes.

Data Frequency: To check whether data frequency may alter the findings, we consider higher frequency data, such as weekly and daily data. Our weekly data are from 1973/2/7 to 2008/10/29 with 1865 observations, while the daily data are from 1973/2/1 to 2008/10/31 with 9026 observations. The results are presented in columns (4) and (5) of Table 4. It turns out that using higher frequency data results in a similar asymmetric return–volume relation. However, the

estimates of β_0 are not significant, which may be related to the heteroscedasticity of the data

with higher frequency such as weekly or daily data are usually not homoscedastic.[2]

Subsample Periods: It is of interest to know whether the fluctuation characteristics of the

stock market, as well as the return–volume relation, change across different sample periods.

As in Table 2, we first consider a subsample period from 1973M2 to 1999M12 as a robustness

check. Moreover, it is argued that the growth in financial globalization may cause structural

changes in the US stock market because of factors such as increasing international comove-

ments in stock prices. For instance, Arshanapalli and Doukas (1993) show that the degree of in-

ternational comovements among stock price indices has increased substantially after 1987M10.

Thus, we also consider a subsample period from 1987M11 to 2008M10. Empirical results for

the two subsample periods are shown in columns (6) and (7) of Table 4. Clearly, the recessions

and booms in the stock market are also well identified in the subsample periods (low mean/high

variance vs. high mean/low variance). Moreover, the return–volume relation is asymmetric

over the stock market fluctuations, which is consistent with our main findings.

[Insert Table 4 Here]

6. Alternative identifications of stock market fluctuations

Although recessions and booms in stock returns are widely acknowledged, Candelon et al.

(2008) argue that there is no consensus in the academic literature on what bear and bull markets

---

[2]We follow Engle (1982) to test for ARCH effects in weekly and daily returns. The F statistics (p-value) are
35.15 (0.00) for weekly returns and 127.37 (0.00) for daily returns, respectively. The test results suggest significant
ARCH effects. Therefore, considering a Markov-switching model with GARCH effects for high frequency data
appears a promising avenue for future research.

actually are. Parametric and nonparametric methods have both been employed to identify recessions and booms in the stock market. In particular, bull and bear markets are explicitly identified in Maheu and McCurdy (2000) using parametric models (Markov-switching models), whereas nonparametric approaches are used in Candelon et al. (2008). Following the above literature, we use both a nonparametric approach and a naive moving average approach as alternative methods to our Markov-switching models.

Model A: A Nonparametric Approach   Candelon et al. (2008) note that the key feature of nonparametric dating algorithms is the location of turning points (peaks and troughs), which correspond to the local maxima and minima of the series. We follow the setting in Candelon et al. (2008) to identify a peak (or trough) in the stock market when r_t reaches a local maximum (or minimum) in a six-month window, using the monthly Bry–Boschan algorithm. That is, a local peak occurs at time t whenever $\{r_t > r_{t\pm k}\}$, $k = 1, 2, \ldots, 6$. Likewise, there will be a trough at time t if $\{r_t < r_{t\pm k}\}$, $k = 1, 2, \ldots, 6$. Once turning points are obtained, the peak-to-trough period and the trough-to-peak period are identified as the bear ($D_t = 1$) and the bull ($D_t = 0$) markets, respectively. D_t is a binary dummy variable to indicate the recessions and booms in the stock market.

Model B: A Naive Moving Average Approach   Under the naive moving average approach, the bull or bear market is decided by the mean return over the last couple of periods. We may define $\bar{r}_t^k$ as the moving average of the last k values of the stock returns, $\bar{r}_t^k = \frac{r_{t-1} + r_{t-2} + \cdots + r_{t-k}}{k}$.

Then, we define a dummy variable D_t as follows:

$$D_t = \begin{cases} 1 & \text{(bear market)} \quad \text{if } \bar{r}_t^k \leq 0, \\ \\ 0 & \text{(bull market)} \quad \text{if } \bar{r}_t^k > 0. \end{cases} \quad (8)$$

That is, if the mean return over the last k periods is negative, we identify the current market status as a bear market. On the other hand, a bull market is defined by a positive mean return over the last k periods.

For Models A and B, we consider the following regression model:

$$r_t = \alpha_0 + \alpha_1 D_t + \varphi_1 r_{t-1} + \delta_0 V_t + \delta_1 (V_t \times D_t) + \varepsilon_t. \quad (9)$$

That is, in bull markets, the return–volume correlation is represented by δ_0, whereas $\delta_0 + \delta_1$ shows the return–volume correlation in bear markets.

In Table 5, we report the regression results from estimating equation (9) with bear/bull dummy variables identified by the nonparametric Bry–Boschan dating algorithm and the moving average of order five, MA(5) model.[3] Clearly, departing from the Markov-switching framework and using a different measure of bear markets (models A and B) does not substantially alter the findings. It is evident that in the bull market, $\hat{\delta}_0 > 0$ implies a positive correlation between returns and trading volume. On the other hand, $\hat{\delta}_0 + \hat{\delta}_1 < 0$ suggests that the correlation is negative in a bear market. Moreover, in most cases, the estimates $\hat{\delta}_1$ are statistically significant, which provides strong evidence of an asymmetric return–volume relation. Therefore, our main findings are robust to different methods of identifying the bull and bear markets.

[Insert Table 5 Here]

---

[3]The results do not change substantially when considering different orders of the MA model.

16

7. Dynamic relationship

As discussed above, since the 1990s, the focus of financial research has moved to investigating the dynamic (causal) correlation between price changes and trading volume, i.e., to investigating whether trading volume precedes stock returns, or vice versa. First, it is important to know if trading volume provides useful information content that would improve stock return forecasts. For instance, Blume et al. (1994) demonstrate a model in which volume yields insights into the quality of traders' information that cannot be deduced from past price movements.

On the other hand, it is also of interest to ask if investors trade more when markets have done well in the past. As argued in Griffin et al. (2007), answering such a question may help in obtaining forecasts of trading intensity, and devising efficient trading strategies.

To investigate the dynamic relation between volume and returns, equation (1) needs to be modified. Consider the following joint two-state Markov-switching model:

$$r_t = \mu_{s_t} + \varphi r_{t-1} + \sum_{i=1}^{k} \lambda_{s_t,i} V_{t-i} + \varepsilon_t, \tag{10}$$

$$V_t = \gamma_{s_t} + \phi V_{t-1} + \sum_{i=1}^{k} \theta_{s_t,i} r_{t-i} + \eta_t, \tag{11}$$

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim \text{i.i.d.} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{s_t}^2 & 0 \\ 0 & \xi_{s_t}^2 \end{bmatrix} \right),$$

where μ_{s_t} and $\sigma_{s_t}^2$ are the state-dependent mean and variance of r_t. Coefficient $\lambda_{s_t,i}$ is the state-dependent coefficient on the lagged volume. γ_{s_t} and $\xi_{s_t}^2$ are the state-dependent mean and variance of V_t. Coefficient $\theta_{s_t,i}$ is the state-dependent coefficient on the lagged stock return.

We report the empirical results for the return equation (equation (10)) in Table 6 and the results for the volume equation (equation (11)) in Table 7. We consider the lag lengths $k = 1$

17

and $k = 2$. Table 6 suggests that in bear markets, $\hat{\lambda}_{0,1}$ is statistically significant under the case $k = 1$, and the F-statistic to test the hypothesis $\hat{\lambda}_{0,1} = \hat{\lambda}_{0,2} = 0$ is also significant for $k = 2$. Hence, trading volume is able to predict price movements only in the bear market. In contrast, it is clear in Table 7 that stock returns are capable of predicting trading volume in both bear and bull markets. For $k = 1$, both $\hat{\theta}_{0,1}$ and $\hat{\theta}_{1,1}$ are statistically significant. For $k = 2$, the F-statistics reach the same conclusion. That is, we have found strong evidence that returns are able to forecast trading volume, regardless of whether stock markets are in recessions or booms. On the other hand, the evidence for stock return forecastability from trading volume is weaker.

[Insert Table 6 Here]

[Insert Table 7 Here]

8. Concluding remarks

The relation between stock price changes and trading volume (return–volume relation) has received considerable attention over the past two decades in the field of finance. This paper investigates whether the relation is asymmetric in bull and bear stock markets. Using monthly data for the S&P 500 price index and trading volume from 1973M2 to 2008M10, we estimate a Markov-switching model to identify the different phases of stock market cycles, and then study the asymmetric return–volume relation.

In regard to contemporaneous correlation, we find that returns and volume are negatively correlated in the bear market, whereas in the bull market, the correlation is positive. The asymmetric return–volume relation is statistically significant, and is robust to differences in model

18

specifications, measures of trading volume, data frequencies, subsample periods and methods of identifying bull and bear stock markets. The explanation for such an asymmetric return–volume correlation is intuitive. In the bear market, most people have already lost considerable sums of money, so they will not sell their stocks even if the price goes up. The driving force on stock prices is from reductions in supply, instead of increases in demand. Moreover, the low-return-high volume case in bear markets may be explained by a panic overselling at the bottoms of market declines. On the other hand, in the bull market, overconfidence may grow with long-lasting past success in the market, which would result in a strong positive return–volume correlation. Moreover, in general, momentum investing is based on the belief that an extended bull market is in effect. Hence, we would expect a positive correlation between price changes and volume in the bull market.

Further, we investigate the causal link between returns and trading volume. Using a joint two-state Markov-switching model, we present strong evidence that the stock return is able to forecast volume in both bear and bull markets. There is weaker evidence regarding the information content of trading volume to forecast stock returns. The forecastability is found only in bear markets.

Acknowledgements

Table 1: Descriptive Statistics and Unit Root Tests

	Stock Returns	Volume in Logarithm	Percentage Changes in Volume
Mean	0.48	14.77	1.69
Standard Deviation	4.78	2.08	18.58
Observations	428	429	428
ADF	-20.276	0.018	-10.860
PP	-20.353	0.004	-42.412
DF-GLS	-12.240	4.649	-2.376

Note: ADF, PP and DF-GLS are Augmented Dickey-Fuller, Phillips-Perron and Elliott-Rothenberg-Stock DF-GLS test statistics, respectively. In each test, the null hypothesis is that the series has a unit root. Test critical values for ADF and PP are -3.44 (1%), -2.87 (5%) and -2.57 (10%). Test critical values for DF-GLS are -2.58 (1%), -1.95 (5%) and -1.62 (10%). Lags in ADF and DF-GLS tests are chosen by Akaike Information Criterion (AIC).

Table 2: Linear Regression and Granger Causality Tests

Dependent Variable	1973M2–2008M10				1973M2–1999M12			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	r_t	r_t	r_t	V_t	r_t	r_t	r_t	V_t
Constant	0.51**	0.48**	0.54**	3.33***	0.77***	0.82***	0.90***	3.06***
	(0.23)	(0.23)	(0.25)	(0.77)	(0.26)	(0.26)	(0.27)	(0.91)
r_{t-1}	-0.04	-0.04	-0.04	0.20	-0.05	-0.05	-0.04	0.46**
	(0.05)	(0.05)	(0.05)	(0.16)	(0.06)	(0.06)	(0.06)	(0.19)
r_{t-2}			0.04	0.19			0.06	0.24
			(0.05)	(0.16)			(0.06)	(0.19)
r_{t-3}			0.002	-0.11			-0.05	-0.07
			(0.05)	(0.16)			(0.06)	(0.19)
V_t	0.0001				0.031**			
	(0.01)				(0.014)			
V_{t-1}		0.019	0.01	-0.67***		0.01	-0.01	-0.67***
		(0.013)	(0.02)	(0.05)		(0.01)	(0.02)	(0.06)
V_{t-2}			-0.02	-0.41***			-0.02	-0.40***
			(0.02)	(0.06)			(0.02)	(0.07)
V_{t-3}			-0.02	0.02			-0.01	0.005
			(0.02)	(0.05)			(0.02)	(0.06)
F-stat			1.48	1.10			0.55	2.40*
p-value			0.22	0.35			0.65	0.07

Note: The entries in brackets are the standard errors. The F-stat and p-value are tests of Granger causality. Asterisks *, ** and *** indicate rejection at the 10%, 5% and 1% level, respectively.

Table 3: Contemporaneous Relationship: Markov-switching Models

	(1)	(2)	(3)	(4)
μ	0.48** (0.23)			
μ_0		-0.51 (0.64)	-0.67 (0.73)	-0.61 (0.69)
μ_1		1.00*** (0.20)	1.15*** (0.21)	1.19*** (0.26)
σ	4.77*** (0.17)			
σ_0		6.86*** (0.53)	6.76*** (0.47)	6.78*** (0.48)
σ_1		3.03*** (0.21)	3.00*** (0.21)	2.94*** (0.21)
φ_1			-0.096* (0.052)	-0.10** (0.05)
φ_2				0.01 (0.05)
φ_3				-0.003 (0.05)
φ_4				0.001 (0.05)
φ_5				-0.03 (0.04)
φ_6				0.04 (0.05)
β_0			-0.07* (0.04)	-0.07* (0.04)
β_1			0.04*** (0.01)	0.04*** (0.01)
p^{00}		0.92	0.91	0.92
p^{11}		0.96	0.96	0.96
LogLik	-1276.26	-1236.17	-1225.23	-1210.99

Note: The entries in brackets are the standard errors. The dependent variable is the stock returns. The model is $\varphi(L)r_t = \mu_{s_t} + \beta_{s_t}V_t + \varepsilon_t$ with mean/variance (μ_0, σ_0^2) in regime 0 and (μ_1, σ_1^2) in regime 1. Asterisks *, ** and *** indicate rejection at the 10%, 5% and 1% level, respectively. Columns (1) and (2) reports the results for the linear and MS-AR(0) models without including V_t as a regressor. Column (3) and (4) show the results from MS-AR(1) and MS-AR(6) model with V_t.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
μ_0	-0.31	-1.33*	-1.21	-0.29*	-0.10**	-0.21	-0.01
	(0.64)	(0.71)	(0.91)	(0.17)	(0.04)	(0.96)	(0.66)
μ_1	1.45***	1.26***	1.69 ***	0.29***	0.06***	1.20***	1.15***
	(0.24)	(0.22)	(0.30)	(0.05)	(0.01)	(0.25)	(0.24)
σ_0	6.60***	6.11***	5.59***	3.22***	1.79***	6.92***	6.09***
	(0.47)	(0.49)	(0.63)	(0.16)	(0.04)	(0.70)	(0.46)
σ_1	2.87***	3.17***	3.17***	1.57***	0.73***	2.99***	2.73***
	(0.17)	(0.19)	(0.23)	(0.05)	(0.01)	(0.24)	(0.19)
φ_1	-0.13**	-0.18***	-0.10*	-0.04**	0.04***	-0.09	-0.09
	(0.06)	(0.05)	(0.06)	(0.02)	(0.01)	(0.06)	(0.06)
β_0	-3.81*	-18.69***	-0.18***	-0.004	-0.002	-0.01	-0.108***
	(2.23)	(5.14)	(0.03)	(0.01)	(0.002)	(0.05)	(0.035)
β_1	3.02***	10.82***	0.021*	0.01***	0.01***	0.05***	0.016
	(0.92)	(1.84)	(0.012)	(0.003)	(0.0004)	(0.01)	(0.015)
p^{00}	0.92	0.86	0.77	0.94	0.97	0.84	0.97
p^{11}	0.95	0.94	0.92	0.97	0.99	0.94	0.98
LogLik	-1226.52	-1213.33	-808.61	-3980.81	-11573.70	-910.82	-704.15

Note: The entries in brackets are the standard errors. The dependent variable is the stock returns. The model is $\varphi(L)r_t = \mu_{s_t} + \beta_{s_t}V_t + \varepsilon_t$ with mean/variance (μ_0, σ_0^2) in regime 0 and (μ_1, σ_1^2) in regime 1. Asterisks *, ** and *** indicate rejection at the 10%, 5% and 1% level, respectively. In column (1), V_t is the detrended volume adjusted for linear and quadratic time trend. In column (2), V_t is detrended volume adjusted for MA(5) trend. Column (3) reports the result using the Dow Jones Industrial Average index. In columns (4) and (5), weekly and daily S&P 500 index is used. Finally, in columns (6) and (7), subsample periods are 1973M2–1999M12 and 1987M11–2008M10, respectively.

Table 5: Contemporaneous Relationship: Bry–Boschan and Moving Average Methods

	Bry–Boschan	Moving Average
α_0	0.694*	1.886***
	(0.375)	(0.266)
α_1	-0.310	-3.869***
	(0.479)	(0.465)
φ_1	-0.036	-0.161***
	(0.051)	(0.048)
δ_0	0.039**	0.039***
	(0.019)	(0.014)
δ_1	-0.067***	-0.113***
	(0.025)	(0.024)

Note: The entries in brackets are the standard errors. The dependent variable is the stock return. The regression model is $r_t = \alpha_0 + \alpha_1 D_t + \varphi_1 r_{t-1} + \delta_0 V_t + \delta_1 (V_t \times D_t) + \varepsilon_t$, where V_t is trading volume. Dummy variable D_t is constructed as $D_t = 1$ for a bear market at time t and $D_t = 0$ for a bull market. Asterisks *, ** and *** indicate rejection at the 10%, 5% and 1% level, respectively.

Table 6: Dynamic Relation: Joint Markov-switching Model

	$k = 1$	$k = 2$
μ_0	-0.79	-0.76
	(0.73)	(0.74)
μ_1	1.02***	0.97***
	(0.23)	(0.22)
σ_0	7.01***	7.06***
	(0.55)	(0.55)
σ_1	2.98***	3.00***
	(0.25)	(0.22)
φ	-0.06	-0.05
	(0.05)	(0.06)
$\lambda_{0,1}$	0.078**	0.075*
	(0.039)	(0.04)
$\lambda_{1,1}$	-0.01	-0.01
	(0.01)	(0.01)
$\lambda_{0,2}$		-0.04
		(0.04)
$\lambda_{1,2}$		0.01
		(0.01)
p^{00}	0.78	0.70
p^{11}	0.90	0.87
LogLik	-3015.86	-3007.85
F-stat ($\lambda_{0,1} = \lambda_{0,2} = 0$)		2.66*
P-value		0.07
F-stat ($\lambda_{1,1} = \lambda_{1,2} = 0$)		1.15
P-value		0.32

Note: The entries in brackets are the standard errors. The Joint Markov-switching Models for the return equation is $r_t = \mu_{s_t} + \varphi r_{t-1} + \sum_{i=1}^{k} \lambda_{s_t,i} V_{t-i} + \varepsilon_t$ with mean/variance (μ_0, σ_0^2) in regime 0 and (μ_1, σ_1^2) in regime 1. Asterisks *, ** and *** indicate rejection at the 10%, 5% and 1% level, respectively.

Table 7: Dynamic Relation: Joint Markov-switching Model

	$k = 1$	$k = 2$
γ_0	6.69*	7.99***
	(4.02)	(3.22)
γ_1	-0.19	-0.80
	(1.53)	(1.26)
ξ_0	19.22***	19.39***
	(1.85)	(1.43)
ξ_1	13.63***	13.32***
	(1.36)	(0.84)
ϕ	-0.49***	-0.50***
	(0.04)	(0.04)
$\theta_{0,1}$	-0.73**	-0.82**
	(0.36)	(0.37)
$\theta_{1,1}$	0.99***	0.97***
	(0.35)	(0.28)
$\theta_{0,2}$		-0.09
		(0.38)
$\theta_{1,2}$		0.36
		(0.22)
p^{00}	0.78	0.70
p^{11}	0.90	0.87
LogLik	-3015.86	-3007.85
F-stat ($\theta_{0,1} = \theta_{0,2} = 0$)		2.53*
P-value		0.08
F-stat ($\theta_{1,1} = \theta_{1,2} = 0$)		7.40***
P-value		0.00

Note: The entries in brackets are the standard errors. The Joint Markov-switching Models for the volume equation is $V_t = \gamma_{s_t} + \phi V_{t-1} + \sum_{i=1}^{k} \theta_{s_t,i} r_{t-i} + \eta_t$ with mean/variance (γ_0, ξ_0^2) in regime 0 and (γ_1, ξ_1^2) in regime 1. Asterisks *, ** and *** indicate rejection at the 10%, 5% and 1% level, respectively.

References

Arshanapalli, B., Doukas, J., 1993. International stock market linkages: Evidence from the pre-
and post-october 1987 period. Journal of Banking and Finance 17, 193–208.

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

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

Candelon, B., Piplack, J., Straetmans, S., 2008. On measuring synchronization of bulls and
bears: The case of East Asia. Journal of Banking and Finance 32, 1022–1035.

Chuang, C.-C., Kuan, C.-M., Lin, H.-Y., July 2009. Causality in quantiles and dynamic stock
return-volume relations. Journal of Banking and Finance 33, 1351–1360.

Eleanor Xu, X., Chen, P., Wu, C., May 2006. Time and dynamic volume-volatility relation.
Journal of Banking and Finance 30, 1535–1558.

Elliott, G., Rothenberg, T. J., Stock, J. H., 1996. Efficient tests for an autoregressive unit root.
Econometrica 64, 813–36.

Engle, R., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of
U.K. inflation. Econometrica 50, 987–1008.

Frauendorfer, K., Jacoby, U., Schwendener, A., 2007. Regime switching based portfolio selec-
tion for pension funds. Journal of Banking and Finance 31, 2265–2280.

Garcia, R., 1998. Asymptotic null distribution of the likelihood ratio test in Markov switching models. International Economic Review 39, 763–788.

Granger, C. W. J., Morgenstern, O., 1963. Spectral analysis of New York stock market prices. Kyklos 16, 1–27.

Griffin, J. M., Nardari, F., Stulz, R. M., 2007. Do investors trade more when stocks have performed well? Evidence from 46 countries. Review of Financial Studies 20, 905–951.

Hamilton, J. D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57, 357–384.

Hamilton, J. D., Lin, G., 1996. Stock market volatility and the business cycle. Journal of Applied Econometrics 11, 573–593.

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

Hong, H., Scheinkman, J., Xiong, W., 2006. Asset float and speculative bubbles. Journal of Finance 61, 1073–1117.

Hutson, E., Kearney, C., Lynch, M., 2008. Volume and skewness in international equity markets. Journal of Banking and Finance 32, 1255–1268.

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

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

Maheu, J. M., McCurdy, T. H., 2000. Identifying bull and bear markets in stock returns. Journal of Business and Economic Statistics 18, 100–112.

McMillan, D. G., 2007. Non-linear forecasting of stock returns: Does volume help? International Journal of Forecasting 23, 115–126.

Ning, C., Wirjanto, T. S., 2009. Extreme return-volume dependence in east-asian stock markets: A copula approach. Finance Research Letters 6, 202–209.

Perez-Quiros, G., Timmermann, A., 2000. Firm size and cyclical variations in stock returns. Journal of Finance 55, 1229–1262.

Psaradakis, Z., Spagnolo, N., 2003. On the determination of the number of regimes in Markov-switching autoregressive models. Journal of Time Series Analysis 24, 237–252.

Saatcioglu, K., Starks, L. T., 1998. The stock price-volume relationship in emerging stock markets: The case of Latin America. International Journal of Forecasting 14, 215–225.

Statman, M., Thorley, S., Vorkink, K., 2006. Investor overconfidence and trading volume. Review of Financial Studies 19, 1531–1565.
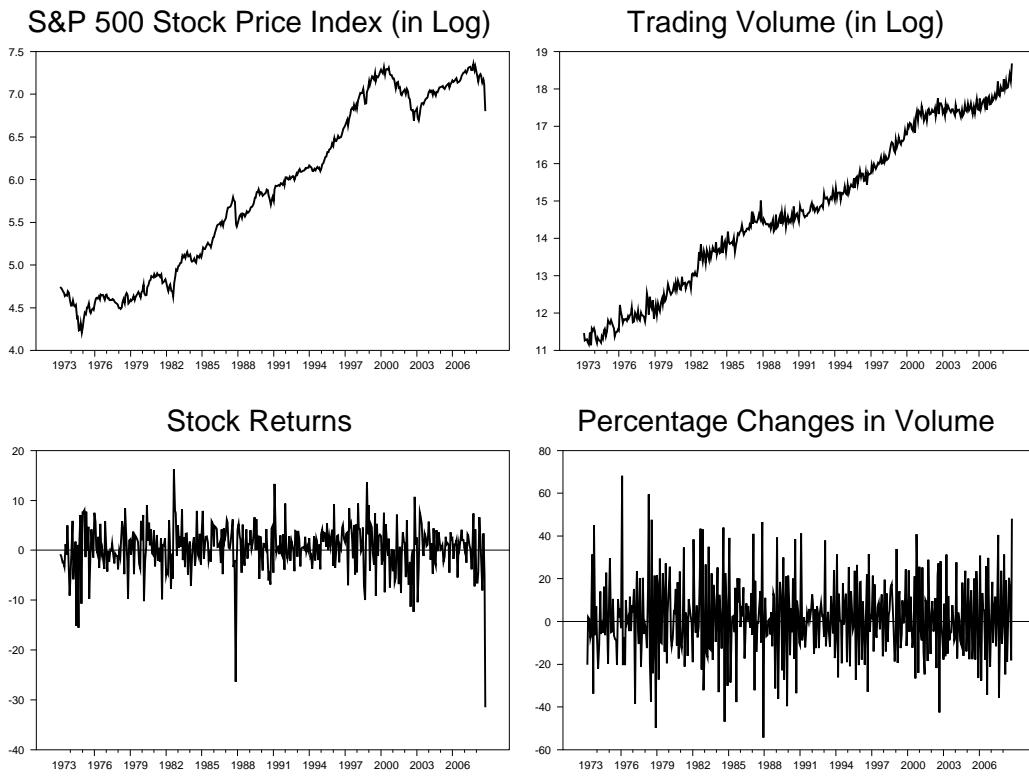
Figure 1: Data Plots

30

Figure 2: Smoothing Probabilities in State 1 (Bull Markets)