Research in International Business and Finance 24 (2010) 146–157

Contents lists available at ScienceDirect

Research in International Business and Finance

journal homepage: www.elsevier.com/locate/ribaf

Does volume help in predicting stock returns? An analysis of the Australian market

Emawtee Bissoondoyal-Bheenick[a,*], Robert D. Brooks[b]

[a] Department of Accounting and Finance, Monash University, Melbourne, Australia
[b] Department of Econometrics and Business Statistics, Monash University, Melbourne, Australia

ARTICLE INFO

Article history:
Received 21 May 2009
Received in revised form 19 October 2009
Accepted 4 November 2009
Available online 13 November 2009

Keywords:
G12
G14
G15

Keywords:
Volume
Stock market returns
Binary probit

ABSTRACT

This paper presents an analysis of the relationship between trading volume and stock returns in the Australian market. We test this hypothesis by using data from a sample of firms listed on the Australian stock market for a period of 5 years from January 2001 to December 2005. We explore this relationship by focusing on the level of trading volume and thin trading in the market. Our results suggest that trading volume does seem to have some predictive power for high volume firms and in certain industries of the Australian market. However, for smaller firms, trading volume does not seem to have the same predictive power to explain stock returns in Australia.

1. Introduction

It is well known that stock market investing is risky. Both practitioners and academics recommend holding a well-diversified portfolio to reduce risk. To date, there has been a very large focus on the benefits of diversification and a large number of studies have focused on the correlation between stock markets. In line with this area of research, there have been numerous studies on the relationship between trading volume and stock returns. In particular some researchers have argued that low volume stocks are typified by momentum while high volume stocks exhibit mean reverting behavior in returns. The relationship between volume and future returns suggests that volume can in fact be used

* Corresponding author at: Department of Accounting and Finance, Monash University, PO Box 197, Caulfield East, VIC 3145, Australia. Tel.: +61 3 99032957; fax: +61 3 9903 2422.
E-mail address: Banita.Bissoondoyal-Bheenick@buseco.monash.edu.au (E. Bissoondoyal-Bheenick).

0275-5319/$ – see front matter © 2009 Published by Elsevier B.V.
doi:10.1016/j.ribaf.2009.11.001

as a threshold variable to forecast stock returns. However, it should be highlighted that there have been studies where a weak relationship between volume and returns has been found, for example see Karpoff (1987). Karpoff (1987) identified several key reasons as to why it is important to analyse the price–volume relationship. He identifies four reasons which are as follows: (1) it provides an insight into the structure of the financial markets, (2) the price–volume relation is important for event studies that use a combination of price and volume data to draw conclusions, (3) this relationship is critical to the debate of speculative prices and lastly price–volume relations have significant implications for research into futures markets.

Following these reasons, the relationship between return and volume has been explored extensively in the finance literature rather than just focusing on the relationship between volume and volatility of returns. There is equally a wide literature which explores the relation between trading volumes and skewness. Chen et al. (2001) develop a series of cross-sectional specifications to forecast skewness in the daily returns of individual stocks in the US market for the period of July 1962 to December 1998. Their findings suggest that negative skewness is greater in stocks that have an increase in trading volume and those stocks that have had positive returns over the prior 36 months. However, when they extend the analysis to market level data, the relationship between volume and skewness does not seem to be the same. It should be highlighted though that studies following Chen et al. (2001) on the volume–skewness relationship at the market level suggest contradictory results (see for example, Hueng and Brooks, 2003; Charoenook and Daouk, 2004; Hueng and Mcdonald, 2005). Recently, Hutson et al. (2008) extended the work by Chen et al. (2001) to examine the relationship of trading volume and skewness in 11 international stock market using both daily and monthly data from January 1980 to August 2004. Hutson et al. (2008) suggest that the differences in prior studies are due to categories of variables analysed, in particular, a set of incomplete variables that have been used. They motivate their work by using variables suggested by Chordia et al. (2001) who examined the impact of the variability in trading activity on expected returns. Overall their market level findings are consistent with the firm level analysis of Chen et al. (2001) that this high trading volume leads to negative skewness in returns.

There are a number of other studies that highlight on the existence of a negative relationship between volume and future returns (see for example, Campbell et al., 1993; Conrad et al., 1994; Datar et al., 1998; Wang and Chin, 2004). In a recent paper by McMillan (2007), the reasons behind the negative relationship are explained from the market structure and the recent behavioral finance literature. For the market structure explanation, market makers are expected to act following high volume which signals that liquidity traders are active in the market, and hence the market makers adjust the prices. The behavioral argument is that momentum is consistent with low volume (see Barberis et al., 1998; Hong and Stein, 1999). McMillan (2007) explores the negative relationship between volume and future returns and suggests that volume can in fact be used as a threshold variable to forecast stock returns. He conducts a study using daily national stock market index data for the UK, US, France and Japan. He analyses the non-linear relationship between stock returns and lagged trading volume using the STAR class of models. The results indicate that the low volume regime is associated with positive autocorrelation and momentum behavior, while the high volume regime is associated with either randomness, or reverting behavior (negative autocorrelation).

In this study, we consider and extend some of the arguments put forward by McMillan (2007) in the Australian context. A focus of our analysis is to explore the volume–returns relationships for individual stock data. An added complication in the analysis of individual stock data is the occurrence of censoring that is zero return observations in the data due to thin trading. In the Australian context, this has been found to be an issue in a range of applications including volatility modelling (Brooks et al., 2001), beta estimation (Brooks et al., 2004a, 2005) and CAPM testing (Brooks et al., 2004b). In the international context Lesmond (2005) found censoring to be a significant issue in emerging market returns. The study by McMillan (2007) considers the countries which form part of the top 10 stock exchanges of the world, and with a focus on index data does not need to explicitly consider the possibility of any censoring in the data. The issue of censoring is very important, as even if we are dealing with the large stock exchanges, the movements of the value weighted indices are in fact driven by trading of the larger stocks in the market and hence it becomes important to consider an analysis by focusing on only those stocks where there has been substantial movement in trading volume to assess the importance of the relationship between return and volume. It is well known in the literature that thin trading and firm size are strongly related (*e.g.* see Roll, 1981; Reinganum, 1982). What this means therefore, is that the most thinly traded stocks are the smallest stocks and so the smallest stocks are the most likely to exhibit severe thin trading problems. In extreme cases of thin trading, we may observe a zero change in price and thus a zero return; hence in this analysis we analyse these stocks to assess whether trading volume can be used to predict returns in the Australian market across varying degrees of censoring. The degree of censoring in a dataset with n observations may be measured by the proportion of observations piled up at the censoring point (in our case at zero). This proportion is given by $c = n_1/n$, where n_1 is the number of observations at the censoring point, zero.

The remainder of this paper is organized as follows. The next section outlines the data and the models that have been considered in this study. Section 3 then presents and discusses the empirical results of findings across our sample of Australian stocks. The final section provides a summary and conclusion.

2. Data and modelling framework

Daily stock return index and volume data were obtained from DataStream over a sample period of 2 January 2000 to 21 December 2005. The return was then calculated by taking the first difference of the natural log of the index. The volume data is the average number of shares traded in the market for a particular day. The initial sample obtained for the Australian market comprised of 480 firms. The data was cleaned, that is eliminating those firms for which complete data was not available for this period and the final sample comprised of 357 firms only. The average daily trading volume of the company varies from 4200 shares for Danks Holding Limited and to 23.6 million shares for Telstra Corporation for the 5-year period. Table 1 provides some summary statistics on the volume data for each of the sectors that have been analysed using the GICS classification code. The summary statistics shows the usual characteristics of these types of data that is the mean, median and the standard deviation with the volume measured in thousands of shares. The results in Table 1 reveal the following patterns. First, all of the industries show a positive skew in the data with the mean exceeding the median. This is characteristic of high volume stocks being spread across the industries. The telecommunications industry has the smallest number of stocks, but the highest mean volume, primarily because of Telstra Corporation.

The data has been split in days with positive returns, days that have a negative return and days that have zero returns. The volume, change in volume, lagged volume and the lagged volume change were then calculated for each firm for the sample period 2000–2005. The issue we are interested in is whether the volume measures can be used to predict the days with positive (or negative) returns. Thus, we do not focus on the role of the volume measures in predicting the return themselves, but instead focus on their capacity to predict the direction of returns (either positive or negative). Thus,

Table 1
Descriptive statistics: The descriptive statistics for the daily trading volume for the sample of 357 firms that have been used in the analysis.

GICS classification	No. of firms	Mean – 000s	Median – 000s	SD – 000s
Consumer Discretionary	44	569	273	822
Consumer staples	17	1137	540	1771
Energy	29	912	515	1114
Financials	92	957	160	1584
Healthcare	32	340	206	371
Industrials	51	997	229	1977
Information Technology	20	477	211	640
Materials	58	1075	387	2133
Telecommunication services	6	5097	1582	9154
Utilities	8	666	397	586

Table 2
Predictive accuracy – all firms: The predictive accuracy of the binary probit model across the whole sample of 357 firms.

All firms – 357 firms	Volume		Volume change		Lagged volume		Lagged volume change	
	0	1	0	1	0	1	0	1
Mean	1.0228	0.6630	1.0189	0.6617	0.9934	0.6406	1.0027	0.6411
Median	1.4615	0.1611	1.4686	0.1219	1.4253	0.0515	1.4263	0.0093
% < 1 – underprediction	36%	64%	37%	63%	37%	63%	37%	63%
% > 1 – overprediction	64%	36%	63%	37%	63%	37%	63%	37%

we define an indicator variable (y_{it}) which takes the values of:

$$\begin{aligned} y_{it} \quad &= 1 \text{ if } r_{it} > 0 \\ &0 \text{ if } r_{it} < 0 \end{aligned} \tag{1}$$

where r_{it} is the return on stock i on day t. We can then model using the binary probit model (see, for example, Greene, 2003) of the form

$$Pr(y_{it}) = x_{it}\beta_i + \varepsilon_{it} \tag{2}$$

where ε_i is a normally distributed random error with $E(\varepsilon_i) = 0$ and $V(\varepsilon_i) = 1$. To estimate the model, the variables that constitute x_{it}, must be chosen and in our case, these variables are volume, change in volume, lagged volume and lagged change volume.

3. Results

We first consider the analysis of all firms. We estimate four specifications of the binary probit model with the four different measures of volume (current volume, lagged volume, change in volume and lagged change in volume) as explanatory variables. For each firm, we then determine the degree of predictive accuracy for each of the models. There are two possible outcomes that can be predicted, a positive return ($\ddot{y}_i = 1$) or a non-positive return ($\ddot{y} = 0$). For each firm, we calculate the degree of predictive accuracy, and we then report the mean (median) across the cross-section of 357 firms. Where the predictive accuracy proportion exceeds (is less than) unity then that outcome is being over-predicted by the model. The results in Table 2 report the percentage of firms where the outcome is over-predicted (under predicted).

For all four possible measures, we find an over-prediction of the non-positive outcome indicating that the volume measures are not useful in predicting the direction of returns. For all cases, the extent of over-prediction of the non-positive outcome exceeds 60%, and all of the median extent of over-prediction exceeds 1.4. However, it is worth noting that the mean over-prediction is much lower than the median over-prediction. Thus, there are some firms where the volume measures are doing a better job in predicting positive returns. We now explore if these findings can be linked to characteristics of stocks around volume levels, industry and censoring.

The next stage of our analysis is to determine if different levels of predictive accuracy are obtained based on the volume characteristics of the stocks. The results are reported in Table 3. We partition our stocks into 1 of 10 different volume categories and again report the mean (median) predictive accuracy and proportions of firms with over-prediction (under-prediction). We again find that there are no major differences in results across the four different specifications of the volume measure. However, we do find differences in results depending on volume measures. For the eight smallest volume categories (average daily trading volume less than 1,500,000 shares), we again find that volume performs poorly in predicting positive returns. In contrast, for the two highest volume categories (average daily trading volume of greater than 1,500,000 shares), we find significant role for all of the different volume measures in predicting returns. This finding is consistent with the results of McMillan (2007).

We now explore whether there are differences in predictive accuracy across the different industries. We partition firms into GICS industrial classifications. The results in Table 4 report the mean (median) predictive accuracy and proportions of the firms with over-prediction (under-prediction) for each of

Table 3
Predictive accuracy – volume classification: The predictive accuracy of the binary probit model when the firms are partitioned across 10 categories on the basis of volume.

Volume – 000s	Firms		Volume		Volume change		Lagged volume		Lagged volume change	
			0	1	0	1	0	1	0	1
0–25	30	Mean	0.9596	0.6161	0.9455	0.6283	0.8513	0.5919	0.8640	0.5921
		Median	1.5164	0.0263	1.4924	0.0440	1.2542	0.0203	1.2772	0.0000
		% < 1 – underprediction	40%	60%	40%	60%	40%	60%	40%	60%
		% > 1 – overprediction	60%	40%	60%	40%	60%	40%	60%	40%
25–50	22	Mean	1.0228	0.5954	0.9975	0.6377	0.9362	0.5542	0.9432	0.5754
		Median	1.4944	0.0976	1.4527	0.1779	1.2907	0.0557	1.3291	0.0000
		% < 1 – underprediction	36%	64%	36%	64%	36%	64%	36%	64%
		% > 1 – overprediction	64%	36%	64%	36%	64%	36%	64%	36%
50–100	42	Mean	1.1059	0.5280	1.0594	0.5701	1.0480	0.5241	1.0796	0.4990
		Median	1.5471	0.0550	1.5416	0.0356	1.4520	0.0317	1.4688	0.0014
		% < 1 – underprediction	31%	69%	36%	64%	33%	67%	33%	67%
		% > 1 – overprediction	69%	31%	64%	36%	67%	33%	67%	33%
100–200	55	Mean	1.1814	0.4616	1.1932	0.4420	1.1588	0.4417	1.1813	0.4412
		Median	1.4779	0.0848	1.5000	0.0152	1.4804	0.0122	1.5034	0.0000
		% < 1 – underprediction	24%	76%	25%	75%	25%	75%	25%	73%
		% > 1 – overprediction	76%	24%	75%	25%	75%	25%	73%	25%
200–300	34	Mean	1.0664	0.6457	1.1049	0.5965	1.0947	0.5873	1.0910	0.5965
		Median	1.4713	0.1807	1.5148	0.1066	1.5208	0.0280	1.5156	0.0361
		% < 1 – underprediction	32%	68%	32%	68%	32%	68%	32%	68%
		% > 1 – overprediction	68%	32%	68%	32%	68%	32%	68%	32%
300–400	33	Mean	1.3046	0.3856	1.2765	0.4027	1.2921	0.3509	1.3040	0.3462
		Median	1.5165	0.1227	1.5254	0.0846	1.5000	0.0214	1.5254	0.0000
		% < 1 – underprediction	18%	82%	18%	82%	18%	82%	18%	82%
		% > 1 – overprediction	82%	18%	82%	18%	82%	18%	82%	18%
400–800	46	Mean	0.9973	0.7275	1.0321	0.6856	1.0151	0.6730	0.9981	0.7030
		Median	1.4459	0.2021	1.4920	0.0913	1.5051	0.0629	1.5068	0.1630
		% < 1 – underprediction	37%	63%	37%	63%	37%	63%	37%	63%
		% > 1 – overprediction	63%	37%	63%	37%	63%	37%	63%	37%

800–1500	31	Mean	1.0472	0.6393	1.0541	0.6263	1.0424	0.5990	1.0502	0.5938
		Median	1.4310	0.1768	1.4145	0.1231	1.4561	0.0397	1.4460	0.0000
		% < 1 – underprediction	32%	68%	35%	65%	32%	68%	35%	65%
		% > 1 – overprediction	68%	32%	65%	35%	68%	32%	65%	35%
1500–3000	40	Mean	0.7298	1.0357	0.7112	1.0491	0.6690	1.0786	0.6784	1.0708
		Median	0.1223	1.4822	0.2475	1.4784	0.0257	1.5366	0.0927	1.5108
		% < 1 – underprediction	55%	45%	60%	40%	60%	40%	60%	40%
		% > 1 – overprediction	45%	55%	40%	60%	40%	60%	40%	60%
3000–30000	24	Mean	0.6489	1.1739	0.6260	1.1916	0.6309	1.1759	0.6351	1.1645
		Median	0.1961	1.5361	0.2579	1.4716	0.0471	1.5580	0.0285	1.5583
		% < 1 – underprediction	67%	33%	63%	38%	63%	38%	63%	38%
		% > 1 – overprediction	33%	67%	38%	63%	38%	63%	38%	63%

Table 4
Predictive accuracy – GICS classification: The predictive accuracy of the binary probit model when the firms are partitioned across the GICS industry classifications.

GICS classification	Firms		Volume		Volume change		Lagged volume		Lagged volume change	
			0	1	0	1	0	1	0	1
Consumer discretionary	44	Mean	1.1017	0.6210	1.0618	0.6546	1.0056	0.6666	1.0409	0.6362
		Median	1.5474	0.1099	1.5132	0.1070	1.4249	0.0793	1.4247	0.0345
		% < 1 – underprediction	34%	66%	36%	64%	36%	64%	36%	64%
		% > 1 – overprediction	66%	34%	64%	36%	64%	36%	64%	36%
Consumer staples	17	Mean	0.6313	1.1198	0.6061	1.1458	0.6024	1.1212	0.6145	1.1096
		Median	0.0171	1.5325	0.0352	1.6799	0.0236	1.5325	0.0733	1.5310
		% < 1 – underprediction	65%	35%	65%	35%	65%	35%	65%	35%
		% > 1 – overprediction	35%	65%	35%	65%	35%	65%	35%	65%
Energy	29	Mean	1.1963	0.5054	1.2672	0.4151	1.2068	0.4139	1.2366	0.3961
		Median	1.4645	0.2007	1.5227	0.0657	1.4747	0.0278	1.5190	0.0000
		% < 1 – underprediction	21%	79%	21%	79%	21%	79%	21%	79%
		% > 1 – overprediction	79%	21%	79%	21%	79%	21%	79%	21%
Financials	92	Mean	0.7629	0.8589	0.7424	0.8716	0.7151	0.8529	0.7252	0.8595
		Median	0.3325	1.3828	0.5084	1.3603	0.1410	1.2591	0.4839	1.2587
		% < 1 – underprediction	52%	48%	54%	46%	53%	47%	53%	46%
		% > 1 – overprediction	48%	52%	46%	54%	47%	53%	46%	53%
Healthcare	32	Mean	1.2768	0.4336	1.3007	0.4058	1.2980	0.3708	1.2736	0.4026
		Median	1.5271	0.1193	1.5277	0.0601	1.5563	0.0079	1.5516	0.0000
		% < 1 – underprediction	19%	81%	19%	81%	19%	81%	19%	81%
		% > 1 – overprediction	81%	19%	81%	19%	81%	19%	81%	19%
Industrials	51	Mean	1.0766	0.6383	1.0794	0.6406	1.0949	0.5795	1.0882	0.5956
		Median	1.5567	0.1611	1.5276	0.1796	1.5181	0.0556	1.5252	0.0685
		% < 1 – underprediction	33%	67%	33%	67%	33%	67%	33%	67%
		% > 1 – overprediction	67%	33%	67%	33%	67%	33%	67%	33%
Information technology	20	Mean	1.3164	0.3303	1.3160	0.3309	1.2980	0.2847	1.3240	0.2655
		Median	1.5344	0.0768	1.5316	0.0427	1.5508	0.0108	1.5619	0.0000
		% < 1 – underprediction	15%	85%	15%	85%	15%	85%	15%	85%
		% > 1 – overprediction	85%	15%	85%	15%	85%	15%	85%	15%
Materials	58	Mean	1.0702	0.6227	1.0684	0.6222	1.0169	0.6171	1.0458	0.5991
		Median	1.4363	0.1316	1.4589	0.1166	1.4235	0.0331	1.4715	0.0000
		% < 1 – underprediction	31%	69%	34%	66%	33%	67%	34%	66%
		% > 1 – overprediction	69%	31%	66%	34%	67%	33%	66%	34%

Telecommunication services	6	Mean	1.3604	0.3794	1.3676	0.3709	1.3850	0.3207	1.2549	0.4564
		Median	1.6160	0.0735	1.5554	0.0306	1.5828	0.0000	1.5308	0.0000
		% < 1 – underprediction	17%	83%	17%	83%	17%	83%	17%	83%
		% > 1 – overprediction	83%	17%	83%	17%	83%	17%	83%	17%
Utilities	8	Mean	1.0907	0.6557	1.0641	0.6400	1.0946	0.6248	1.0267	0.6758
		Median	1.4648	0.2801	1.5778	0.0141	1.5555	0.0780	1.5695	0.0076
		% < 1 – underprediction	38%	63%	38%	63%	38%	63%	38%	63%
		% > 1 – overprediction	63%	38%	63%	38%	63%	38%	63%	38%

Table 5
Predictive accuracy – censoring categories: The predictive accuracy of the binary probit model when the firms are partitioned across censoring categories.

Degree of censoring	No of firms		Volume		Volume change		Lagged volume		Lagged volume change	
			0	1	0	1	0	1	0	1
$C \leq 0.1$	147	Mean	0.3559	1.3833	0.3516	1.3881	0.3246	1.3938	0.3230	1.3988
		Median	0.0028	1.5862	0.0024	1.5753	0.0000	1.6085	0.0000	1.5912
		% < 1 – underprediction	81%	19%	83%	17%	82%	18%	83%	17%
		% > 1 – overprediction	19%	81%	17%	83%	18%	82%	17%	83%
$0.1 < C \leq 0.2$	69	Mean	1.4552	0.3023	1.4347	0.3186	1.4549	0.2762	1.4715	0.2604
		Median	1.6265	0.0893	1.6362	0.0746	1.6566	0.0224	1.6529	0.0000
		% < 1 – underprediction	12%	88%	13%	87%	13%	87%	13%	87%
		% > 1 – overprediction	88%	12%	87%	13%	87%	13%	87%	13%
$0.2 < C \leq 0.3$	65	Mean	1.5419	0.1038	1.5462	0.0801	1.5633	0.0277	1.5643	0.0336
		Median	1.5460	0.0802	1.5669	0.0159	1.5603	0.0100	1.5628	0.0000
		% < 1 – underprediction	0%	100%	2%	98%	0%	100%	0%	100%
		% > 1 – overprediction	100%	0%	98%	2%	100%	0%	100%	0%
$0.3 < C \leq 0.4$	37	Mean	1.4650	0.0983	1.4767	0.0687	1.4252	0.0656	1.4392	0.0769
		Median	1.4968	0.0379	1.5187	0.0000	1.4646	0.0068	1.4837	0.0000
		% < 1 – underprediction	3%	97%	3%	97%	3%	97%	3%	95%
		% > 1 – overprediction	97%	3%	97%	3%	97%	3%	95%	3%
$0.4 < C \leq 0.5$	15	Mean	1.4591	0.0542	1.4868	0.0084	1.3844	0.0143	1.4145	0.0000
		Median	1.4681	0.0289	1.4807	0.0000	1.3943	0.0051	1.4147	0.0000
		% < 1 – underprediction	0%	100%	0%	100%	0%	100%	0%	100%
		% > 1 – overprediction	100%	0%	100%	0%	100%	0%	100%	0%
$0.5 < C \leq 0.6$	12	Mean	1.5058	0.0397	1.4693	0.0926	1.3360	0.0161	1.3731	0.0208
		Median	1.5033	0.0038	1.4927	0.0000	1.3270	0.0082	1.3661	0.0000
		% < 1 – underprediction	0%	100%	0%	100%	0%	100%	0%	100%
		% > 1 – overprediction	100%	0%	100%	0%	100%	0%	100%	0%
$0.6 < C \leq 0.7$	9	Mean	1.5053	0.0671	1.4920	0.1045	1.3121	0.0024	1.3894	0.0000
		Median	1.5274	0.0067	1.5045	0.0000	1.2895	0.0000	1.3255	0.0000
		% < 1 – underprediction	0%	100%	0%	100%	0%	100%	0%	100%
		% > 1 – overprediction	100%	0%	100%	0%	100%	0%	100%	0%
$0.7 < C \leq 0.8$	2	Mean	1.5134	0.0859	1.5936	0.1029	1.2361	0.0424	1.3112	0.0000
		Median	1.5134	0.0859	1.5936	0.1029	1.2361	0.0424	1.3112	0.0000
		% < 1 – underprediction	0%	100%	0%	100%	0%	100%	0%	100%
		% > 1 – overprediction	100%	0%	100%	0%	100%	0%	100%	0%

$0.8 < C \leq 0.9$	1	Mean	1.4469	0.0472	1.3871	0.0526	1.1293	0.0000	1.1909	0.0000
		Median	1.4469	0.0472	1.3871	0.0526	1.1293	0.0000	1.1909	0.0000
		% < 1 – underprediction	0%	100%	0%	100%	0%	100%	0%	100%
		% > 1 – overprediction	100%	0%	100%	0%	100%	0%	100%	0%
$C > 0.9$	0	Mean	0	0	0	0	0	0	0	0
		Median	0	0	0	0	0	0	0	0
		% < 1 – underprediction	0%	0%	0%	0%	0%	0%	0%	0%

the industrial classifications. The results are again similar across four volume measures. For the bulk of industries, we reach the same general finding as the overall results, which are an inability of the volume measures to predict positive returns. The exceptions are Consumer Staples and Financials where there is some ability to predict positive returns.

The final stage of our analysis is to explore the role of censoring in predictive accuracy. The results are reported in Table 5.We again find that the results are similar across all of the different specifications of the volume variable. For the lowest censoring category (firms with less than 10% of zero return observations), we find that volume measures are able to make predictions of positive returns. However, for all other censoring categories, we find that the different volume measures are not able to make predictions of positive returns.

4. Conclusion

This study considers an analysis of the relationship between trading volume and stock returns in the Australian market. We test this hypothesis by using data from a sample of firms listed on the Australian stock market for a period of 5 years from January 2001 to December 2005. We explore this relationship by focusing on the level of trading volume and thin trading in the market while there has various studies which focused on the market index data (Chen et al., 2001; McMillan, 2007; Hutson et al., 2008), in this study we focus on Australian individual stock data to explore further dimensions of the McMillan (2007) results to see whether such predictability also applies to Australian individual stocks.

Chen et al. (2001) suggest that increased trading volumes tend to be associated with a negative skewness at individual stock level for the US market. McMillan (2007) finds evidence of such a relationship for a range of indices in developed markets. The results obtained in this study suggest that for high volume and low censoring stocks, we reach a similar conclusion to McMillan (2007), Chen et al. (2001) and Hutson et al. (2008) and find evidence of predictability. However, for other stocks, we do not find evidence of predictability. Hutson et al. (2008) include the Australian market index in their study and support Chen et al. (2001) results at firm level. However, our analysis of the individual stock data of the Australian market cannot be generalized as we find this evidence only for certain categories of stock. Specifically, segregating the market in different categories of high and low volume stocks makes a difference in our results and we find volume to have predictive power in the high volume setting.

References

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. J. Financ. Econ. 49, 307–343.

Brooks, R., Faff, R.W., Fry, T., 2001. GARCH modeling of individual stock data: the impact of censoring, firm size and trading volume. J. Int. Financ. Markets Inst. Money 11, 215–222.

Brooks, R., Faff, R.W., Fry, T., Newton, E., 2004a. Censoring and its impact on multivariate testing of the capital asset pricing model. Appl. Financ. Econ. 14, 413–420.

Brooks, R., Faff, R.W., Fry, T., Newton, E., 2004b. The selectivity corrected market model and heteroscedasticity in stock returns: yet another look at volume versus GARCH. Acc. Res. J. 17, 192–201.

Brooks, R., Faff, R., Fry, T., Gunn, L., 2005. Censoring and its Impact on beta risk estimation. Adv. Invest. Anal. Portfolio Manage. 1, 111–136.

Campbell, J., Grossman, S., Wang, J., 1993. Trading volume and serial correlation in stock returns. Q. J. Econ. 108, 905–939.

Charoenook, A., Daouk H., 2004. Conditional skewness of aggregate market returns. Working paper, Vanderbilt University.

Chen, J., Hong, H., Stein, J.C., 2001. Forecasting crashes: trading volume, past returns, and conditional skewness in stock prices. J. Financ. Econ. 61 (3), 345–381.

Chordia, T, Subrahmanyan, A., Anshuman, V.R., 2001. Trading activity and expected stock returns. J. Financ. Econ. 59, 3–32.

Conrad, J., Hameed, A., Niden, C., 1994. Volume and auto covariances in short-horizon individual security returns. J. Finance 49, 1305–1329.

Datar, V., Naik, N., Radcliffe, R., 1998. Liquidity and asset returns: an alternative test. J. Financ. Markets 1, 203–220.

Greene, W.H., 2003. Econometric analysis, Prentice Hall, 5th Ed.

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

Hong, H., Stein, J.C., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. J. Finance 54, 2143–2184.

Hueng, C.J., Brooks, R., 2003. Forecasting asymmetries in stock returns: evidence from higher order moments and conditional densities. Working paper, University of Alabama.

Hueng, C.J., Mcdonald, J.B., 2005. Forecasting asymmetries in aggregate stock market returns: evidence from conditional skewness. J. Emp. Finance 12, 666–685.
Karpoff, J.M., 1987. The relationship between price changes and trading volume: a survey. J. Financ. Quant. Anal. 22, 109–126.
Lesmond, D., 2005. Liquidity of emerging markets. J. Financ. Econ. 77, 411–452.
McMillan, D., 2007. Non-linear forecasting of stock returns: does volume help? Int. J. Forecast. 23, 115–126.
Reinganum, M., 1982. A direct test of Roll's conjecture on the firm size effect. J. Finance 37, 27–35.
Roll, R., 1981. A possible explanation of the firm size effect. J. Finance 36, 879–888.
Wang, C., Chin, S., 2004. Profitability of return and volume based investment strategies in China's stock market. Pacific-Basin Finance J. 12, 541–564.