Economics Bulletin

Volume 39, Issue 1

Individual Investors and R^2

Jinghan Cai
University of Scranton

Jia He
Southern University of Science and Technology

Jibao He
Shenzhen Stock Exchange

Weili Zhai
Shenzhen University

Abstract

Some behavioral view of R^2 in the literature argues that lower R^2 may imply that the prices are less efficient, since lower R^2 may be the results of higher noise trader participation, and therefore the sum of squared errors is higher. This paper uses a novel dataset from Chinese stock market and directly checks the relationship between R^2 and noise trader participation. Cross-sectionally, we find no evidence supporting the negative relationship between R^2 and noise trader participation. Time-series wise, we find a case where R^2 positively comoves with noise trader participation. This paper casts doubt on the prediction that noise trader participation will lower the R^2.

Citation: Jinghan Cai and Jia He and Jibao He and Weili Zhai, (2019) "Individual Investors and R^2", *Economics Bulletin*, Volume 39, Issue 1, pages 159-165
Contact: Jinghan Cai - jinghan.cai@scranton.edu, Jia He - hej@sustc.edu.cn, Jibao He - jbhe@szse.cn, Weili Zhai - zhaiweili2006@126.com.
Submitted: December 05, 2018. **Published:** February 02, 2019.

1. Introduction

The $R^2$of a stock is the goodness-of-fit measure, derived from regressing the stock's returns on one or multiple market index or common factors. It has been widely adopted in the finance literature as a proxy of "price inefficiency" (see Morck, Yeung and Yu, 2000, among others). However, some behavioral finance literature holds the opposite opinions, that $R^2$of a stock may refer to a better "price efficiency", since more noise trader participation may reduce R^2. These two competing opinions echo Roll (1988), which concludes that the majority of returns are explained either by private information or a "frenzy" unrelated to specific information. While recent literature has reconciled some seemingly discordant findings[1], much more has remained unexplained about the two competing opinions.

How to measure the trading "frenzy" in Roll (1988)? Existing papers try to use measures such that medium-term price momentum and long-term price reversal (see Hou, Peng and Xiong, 2013), etc., to capture the degree of noise trading, which are nonetheless not direct measures of noise trading. Do the stocks traded by more noise traders tend to have lower R^2? This question still remains incompletely answered.

This paper uses a novel dataset from Chinese stock market which is able to directly proxy the degree of noise trading. The dataset has the ratio of buy (sell) volume from individuals for stock i on day t, to the total trading volume in the same day. This individual buy-ratio (sell-ratio) directly proxies the degree of noise trading. Using the novel proxies, we find that; (1) Cross-sectionally, there is no evidence supporting the negative correlation between individual investor (noise trader[2]) participation and R^2. (2) On the time-series dimension, we a use a natural experiment of bull-bear market switching, and figure out that individuals investors become less active in trading when the bull market becomes bearish, while R^2 tends to decrease at the same time. That is, a single stock's R^2 may positively comove with individual investors' participation. This comovement of individual investors and R^2 may reflect that more individual investors tend to participate in trading when market goes up. They are overconfident and attribute the capital gain to their ability, therefore they are trading along with the index and participate more. When market goes down, individual investors quit from trading, leaving a larger fraction of institutions, who do not trade along with the index. This paper casts doubt on the behavioral viewpoints on the R^2 and individual investors.

2. The Data

The data used in this paper is from Shenzhen Stock Exchange, China. It covers the trading data of all the common stocks in the mainboard of Shenzhen Stock Exchange. The sample period is from March 1, 2007 to June 30, 2008. In additional to the daily price and trading volume, the dataset

[1] ***For example, Chan and Hameed (2006) and Piotroski and Roulstone (2004)'s finding that stocks followed by more analysts commove more can be explained by Veldkamp's (2006) information production procedure. More discussions can be found in Morck, Yeung and Yu (2013).***

[2] ***In this paper, we use "individual investors" and "noise traders" interchangeably.***

also provides two items: $BuyRatio_{it}$ is the total number of shares purchased by individual investors, divided by the total share volume for stock i on date t. $SellRatio_{it}$ is the total number of shares sold by individual investors, divided by the total share volume for stock i on date t. There are altogether 462 stocks that have trading records in the sample period. The mean $BuyRatio_{it}$ is 92.16%, while the mean of $SellRatio_{it}$ is 91.98%, implying that it is an individual investor driven market. Also, not surprisingly the correlation $BuyRatio_{it}$ and $SellRatio_{it}$ is as high as 0.99. More details are listed in Table I.

Table I: Descriptive Statistics for the Sample Stocks

Panel A: Basic information for the Individual Investors Per Stock

	Individual BuyRatio	Individual SellRatio	Daily buy Yuan Vol.	Daily buy share Vol.	Daily sell Yuan Vol.	Daily sell share Vol.
Mean	92.16	91.83	$1.07*10^8$	$0.838*10^7$	$1.06*10^8$	$0.831*10^7$
Median	96.19	95.69	0.770	0.638	0.764	0.638
P25	91.20	90.81	0.442	0.384	0.444	0.385
P75	97.79	97.34	1.30	1.07	1.31	1.07
# of stocks	462					

Panel B: Correlation for Individual investors' Participation

	Individual BuyRatio	Individual SellRatio	Daily buy Yuan Vol.	Daily buy share Vol.	Daily sell Yuan Vol.	Daily sell share Vol.
Individual BuyRatio	1					
Individual Sellratio	0.999	1				
Daily buy Yuan Vol.	0.457	0.452	1			
Daily buy share Vol.	0.598	0.595	0.790	1		
Daily sell Yuan Vol.	0.455	0.451	0.999	0.790	1	
Daily sell share Vol.	0.592	0.590	0.786	0.999	0.788	1

3. Empirical Results

We first try to uncover the cross-sectional relationship between R^2 and investor participation. The R_i^2 is calculated as follows: in the whole sample period, we regress the continuously compounded daily individual stock i's return on the Shanghai A Share Composite Index daily return. Since none of these Shenzhen-based stocks in the sample is a component stock in the index, there exists no spurious relationship documented in Barberis, Shleifer and Wurgler (2005). Therefore, we do not need to exclude the underlying stocks from the index. We also take the average of the $BuyRatio_{it}$ and $SellRatio_{it}$ for each stock i across different date t, and construct the series of $BuyRatio_i$, $SellRatio_i$. Then we run the following cross-sectional regression:

$$Syn_i = \alpha + \beta_1 BuyRatio_i + \beta_2 SellRatio_i + \gamma Control_i + \varepsilon_i \quad (1)$$

where Syn_i is the return synchronicity measure, proxied by R_i^2, as well as the log transformation of R_i^2, defined as $\Psi_i = \log\left(\frac{R_i^2}{1-R_i^2}\right)$. Control variables include the average daily buy/sell share volume and Yuan volume from individual investors for stock i. The results are shown in Table II.

Table II: R2 and Individual Investor Participation: Cross-sectional Comparison
Table II shows the results of the following regression model:

$$Syn_i^2 = \alpha + \beta_1 BuyRatio_i + \beta_2 SellRatio_i + \varepsilon_i$$

where Syn_i^2 is the R_i^2 from regression stock i's daily return on the Shanghai A Composite Index Return for all trading days in the whole sample period. We also adopt the log variation, i.e., and $\Psi_i = \log\left(\frac{R^2}{1-R^2}\right)$ as robustness check. $BuyRatio_i$ is the mean of $BuyRatio_{it}$ which the total number of shares purchased by individual investors, divided by the total share volume for stock i on date t, in the whole sample period. $SellRatio_i$ is defined analogously. $Control_i$ is the control variable, including the mean daily share volume and dollar volume for stock i, both of which are averaged over the whole sample period.

Panel A: Dependent Variable R_i^2

Models	(1)	(2)	(3)	(4)
BuyRatio	-0.0578*		-0.0381	
	[-1.73]		[-1.10]	
SellRatio		-0.0904**		-0.0638
		[-2.39]		[-1.57]
Controls	No	No	Yes	Yes
constant	0.243***	0.273***	0.207***	0.230***
	[7.90]	[7.86]	[6.33]	[5.97]
N	462	462	462	462

Panel B: Dependent Variable $\Psi_i = \log\left(\frac{R^2}{1-R^2}\right)$

Models	(5)	(6)	(7)	(8)
BuyRatio	-0.994***		-0.285	
	[-2.90]		[-0.86]	
SellRatio		-1.378***		-0.529
		[-3.35]		[-1.33]
Controls	No	No	Yes	Yes
constant	-0.711**	-0.362	-1.452***	-1.230***
	[-2.43]	[-1.03]	[-4.84]	[-3.43]
N	462	462	462	462

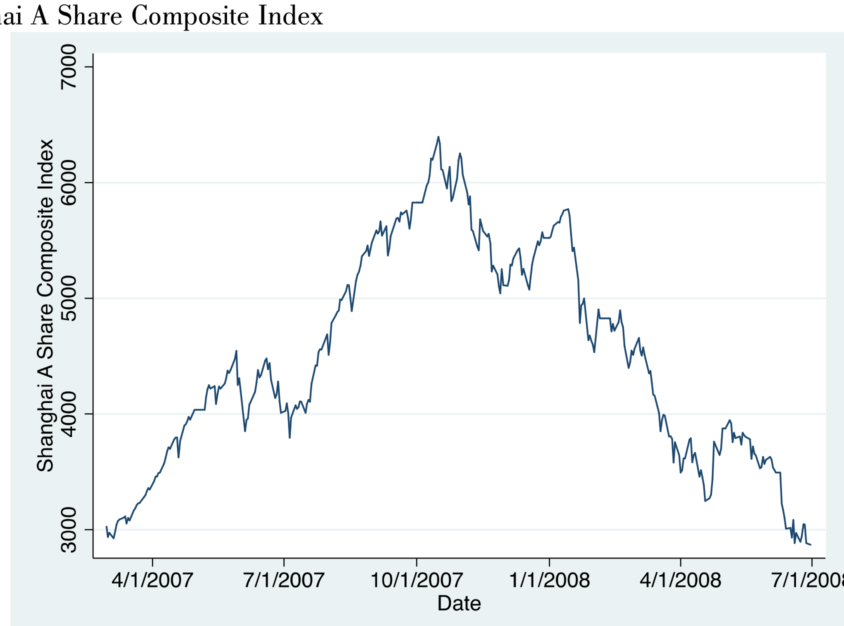
Note: *, ** and *** represent significance levels at 10%, 5%, 1%. t-values are in brackets.

Table II shows that, cross-sectionally, stocks with higher individual participation tend to have lower R_i^2 in specifications (1) and (2). However, after controlling for the trading volumes, the coefficients for BuyRatio and SellRatio both become insignificant, and there is no evidence that cross-sectionally, individual participation can impact the R^2. This find is not supportive to the behavioral literature that more noise traders participation leads to a lower R^2.

Moreover, we are also interested in the following question: for a single stock, when the individual participation rate $(BuyRatio_i, SellRatio_i)$ increases, will the R_i^2 increase or decrease? The motivation that R^2 may positively comove with the individual investors comes from Nofsinger (2014)'s explanation: It is well documented in the literature that individual investors systematically make mistakes in trading (also see Barber and Odean, 2013), including overconfidence. Individual investors tend to attribute their past good performance to their own ability. Therefore, when in bull markets when index returns go up, individual investors become more active, and they participate in trading more actively, making the underlying stock more synchronized with the market, i.e., R^2 becomes higher.

In this paper, we have access to only 16 months' data, which makes it not an ideal setting to check how one single stock's R^2 responds to the change in the individual investor participation on a time-series basis. Instead, we use a separation between the bull market and bear market and check the above hypotheses. In our sample period March 2007 to June 2008, the Chinese stock market experienced a huge bull market. The Shanghai A Share Composite Index rocketed up from 3,026 points to a historical high 6,395 points on October 16, 2017, and then dropped back to 2,869 points on June 30, 2008. This sample period provides a perfect setting where a bear market follows a bull market. We can therefore explore the relationship between the individual investor participation and the R^2.

Figure 1: Shanghai A Share Composite Index



Following the information from Figure 1, we separate the sample period by October 16, 2017, which is the date with a historical high in Shanghai A Share Composite Index. Before this date, it is the bull market window, and after this date, it is the bear market window. In each window, there are approximately 8 months' data. We adopt the market model by regressing the continuously compounded daily individual stock's return on the Shanghai A Share Composite Index return, and obtain the R^2 in both the bull market window and the bear market window. We also calculate the mean daily buy share/Yuan volume for each stock in the bull/bear market window. The comparison between bull and bear markets are shown in Table III.

Table 3: The Bull-Bear Market Comparison

In Panel A, we define the bull market window as between March 1, 2017 to October 16, 2017, and the bear market window as between October 17, 2017 to June 30, 2018. For stock i, in the two windows (bull, bear), we calculate the R^2 from regressing stock i's daily return on the Shanghai A Share Composite Index Return for all trading days in the window (bull, bear). Another return synchronicity measure is defined as $\Psi = \log\left(\frac{R^2}{1-R^2}\right)$. $BuyRatio_{ib=bull,bear}$ is the mean of $BuyRatio_{it}$ in the above windows. $SellRatio_{ib=bull,bear}$is defined analogously. In Panel B, in each month, we run a regression of stock i's return on the Shanghai A Share Composite Index return, and obtain the $R^2_{\Im}$ for stock i in month m. If month m is from March to October, 2017, we define a bear dummy as 0, and if month m is from November 2017 to June 2018, we define the bear dummy as 1. We then run the following regression:

$Y_{\Im} = \alpha + \beta\, Beardum_m + \varepsilon_{\Im}$ where $Y_{\Im}$ includes $R^2_{\Im}$, the log transformation $\Psi_{\Im} = \log\left(\frac{R^2_{\Im}}{1-R^2_{\Im}}\right)$, $BuyRatio_{\Im}$, which is the mean of daily individual investors' $BuyRatio_{it}$ if date t is in month m, and $SellRatio_{\Im}$, which is defined analogously.

Panel A	Windows		Welch t-test
	Bull	Bear	
R_i^2	0.281	0.146	19.26***
$\Psi_i = \log\left(\frac{R_i^2}{1-R_i^2}\right)$	-1.230	-1.880	8.99***
BuyRatio	0.929	0.914	2.18**
SellRatio	0.926	0.910	2.52**

Panel B Dependent Variable	β	t-value	No of obs.	Adj R-sq (%)
$R^2_{\Im}$	-0.127***	-27.80	6861	10.1
$\Psi_{\Im} = \log\left(\frac{R^2_{\Im}}{1-R^2_{\Im}}\right)$	-0.744***	-16.66	6847	3.9
BuyRatio	-0.016***	-5.58	6879	0.4
SellRatio	-0.016***	-6.02	6879	0.5

Note: *, ** and *** represent significance levels at 10%, 5%, and 1%, respectively.

From Table III, we can see that, there is a significant drop in the R^2, from 28.1% to 14.6%. The log transformation of R^2 also experiences a significant decrease. Correspondingly, the $BuyRatio_i$ for individual investors drops. In the bull market, the average $BuyRatio_i$ is 92.9%, while in the bear market, it declines to 91.4%. The change is significant at 5% level. Similarly, the average $SellRatio_i$ also decreases significantly from 92.6% to 91.0%. The above evidences support our early hypothesis that, in the bull market, individual investors are more heavily involved in trading, owing to the overconfidence, compared with the institutional investors.

For robustness sake, we use another setting to check the change of R^2 and individual investors' participation. In each month, we run a regression of stock i's return on the Shanghai A Share Composite Index return, and obtain the $R^2_{\mathfrak{Z}}$ for stock i in month m. If month m is from March to October, 2017, we define a $Beardum_m$ as 0, and if month m is from November 2017 to June 2018, we define the $Beardum_m$ as 1. We then run the following regression:

$$Syn_{\mathfrak{Z}} = \alpha + \beta\, Beardum_m + \varepsilon_{\mathfrak{Z}} \qquad (2)$$

where $Syn_{\mathfrak{Z}}$ includes $R^2_{\mathfrak{Z}}$, the log transformation $\Psi_{\mathfrak{Z}} = \log\left(\frac{R^2_{\mathfrak{Z}}}{1-R^2_{\mathfrak{Z}}}\right)$, $BuyRatio_{\mathfrak{Z}}$, which is the mean of daily individual investors' $BuyRatio_{it}$ if date t is in month m, and $SellRatio_{\mathfrak{Z}}$, which is defined analogously. The results are shown in Panel B of Table III. We can see that, all the coefficients for $Syn_{\mathfrak{Z}}$ are significantly negative, confirming the results from Panel A, Table III that, when the market turns from bull market to bear market, the R^2 and individual investors' participation both decrease.

To summarize, our evidences provide a case that, the R^2 and individual investors' participation in trading positively comove in an event study setting, which casts doubt on the behavioral explanations of R^2 that when individual investors participate more, R^2 will decrease.

4. Conclusion

This paper uses a dataset from Chinese stock market to directly test whether R^2 would be negatively correlated with individual investors participation (See Hou, Peng and Xiong, 2013). We find that, cross-sectionally, there is no evidence supporting the negative relationship between R^2 and individual investor participation. On a time-series basis, there are even some evidences showing that R^2 may positively comove with individual participation (See Nofsinger, 2014). This paper casts doubt on the behavioral viewpoints on the R^2 and individual investors.

Reference

Barber, Brad M. and Terrance Odean (2013) "The Behavior of Individual Investors" in George M. Constantinides, Milton Harris, and Rene M. Stulz, eds., Handbook of Economics of Finance, Vol. 2B: Asset Pricing, Oxford, UK: North-Holland, 1533–1570.

Barberis, Nicholas, Andrei Shleifer, Jeffrey Wurgler (2005) "Comovement" Journal of Financial Economics 75, 283–317

Chan K. and A. Hameed (2006) "Stock Price Synchronicity and Analyst Coverage in Emerging Markets" Journal of Financial Economics 80(1), 115–47

Hou, Kewei, Lin Peng, and Wei Xiong (2013) "Is R^2 a Measure of Market Inefficiency?" working paper

Morck, Randall, Bernard Yeung, and Wayne Yu (2000) "The Information Content of Stock Markets: Why Do Emerging Markets Have Synchronous Stock Price Movements?" Journal of Financial Economics 58, 215-260.

Morck, Randall, Bernard Yeung, and Wayne Yu (2013) "R^2 and the Economy" Annual Review of Financial Economics 5, 143-166

Nofsinger, John (2014) "The Psychology of Investing" Pearson Education 5th edition.

Piotroski, J. and D. Roulstone (2004) "The Influence of Analysts, Institutional Investors, and Insiders on the Incorporation of Market, Industry, and Firm-specific Information into Stock Prices" Accounting Review 79(4), 1119–51

Roll, Richard (1988) "R^2" Journal of Finance 43, 541-566.

Veldkamp, L. (2006) "Information Markets and the Comovement of Asset Prices" Review of Economic Studies 73(3) 823–45