Investor sentiment and the cross-section of stock returns : evidence from China

Guo, Xia

2012

Guo, X. (2012). Investor sentiment and the cross-section of stock returns : evidence from China. Doctoral thesis, Nanyang Technological University, Singapore.

https://hdl.handle.net/10356/50587

https://doi.org/10.32657/10356/50587

Investor Sentiment and the Cross-Section of Stock Returns: Evidence from China GUO Xia 2012

Investor Sentiment and the Cross-Section of Stock Returns: Evidence from China

GUO Xia

NANYANG BUSINESS SCHOOL

2012

GUO Xia

Investor Sentiment and the Cross-Section of Stock Returns: Evidence from China

GUO Xia

Nanyang Business School

A thesis submitted to the Nanyang Technological University in partial fulfilment of the requirement for the degree of Master of Business

2012

Acknowledgement

I would like to express my deepest gratitude and appreciation to my supervisor, Associate Professor Xin Chang (Simba), who kindly accommodated me despite his busy schedule. I am grateful for his invaluable advice throughout the whole of this thesis, without which the completion of this thesis would not have been possible. I also would like to thank all those who have helped me in one way or another.

Contents

Abstract

The thesis studies how investor sentiment affects the cross-section of stock returns in china stock market. I construct an investor sentiment index, which is based on the common variation in four underlying proxies for sentiment: turnover of tradable share, the number and average first-day return of IPOs, and the number of newly opened accounts. I predict that investor sentiment has more pronounced effects on stocks which are more difficult to value and riskier to arbitrage. Consistent with this prediction, I find that when sentiment is high (low), the returns are relatively higher (lower) for small size stocks, high volatility stocks, unprofitable stocks, and extreme growth stocks. When sentiment is low, these categories of stock earn relatively lower returns.

1. Introduction

Traditional finance theories assume rational investors and market efficiency. In equilibrium, price should reflect the fundamental value of the asset. Any price deviation from its fundamental value will be corrected by arbitrageurs (Fama (1970), Sharp (1964)). However, it is well documented in the recent literature that the traditional finance theories fail to explain many financial market anomalies, such as the long-term reversals in stock returns, momentum effects (Jegadeesh and Titman (1993)), and total accruals (Sloan (1996)), etc.

Several recent studies replace the traditional rationality assumptions with behavioral assumptions and have shown great success. For example, Barberis, Shleifer, and Vishy (1998) build a behavioral model which captures long-term reversals, momentum, and post-earnings announcement drift. Daniel, Hirshleifer, and Subrahmanyam (1998, 2001) construct a behavioral model with emphasis on self-attribution bias which explains momentum and post-earnings announcement effect. Given the extensive amount of empirical evidence, it is no longer a debate of whether behavioral finance matters in the financial markets but to what extent it impacts the financial markets.

Two core building blocks of behavioral finance in asset prices are investor sentiment and limits to costless arbitrage. Investor sentiment, defined broadly, is a belief about future cash flows and investment risk that is not justified by the facts at hand (Baker and Wurgler (2007)). Investor sentiment can drive asset price away from its fundamental value. Yet, in an imperfect capital markets, arbitrageurs may find it too costly and risky, if not impossible, to engage in arbitrage activities. Thus, asset prices can be persistently affected by investment sentiment. For example, Internet bubble and the subsequent Nasdaq and telecom crashes validate the two premises of behavioral finance. Extraordinary investor sentiment pushed the prices of speculative and difficult-to-value technology stocks to unfathomable levels in the late 1990s. Instead of creating opportunity for contrarian arbitrageurs, the market forced many out of business, as prices that were merely high went higher still before an eventual crash(Markus and Stefan (2004), Robin and Stefan (2009)).

In this study, I analyze the impact of investor sentiment on stock returns using China stock market data. Baker and Wurgler (2006a) claim that two basic premises of behavioral finance are investor sentiment and the limit of arbitrage. China stock markets fit these two criteria nicely due to a strong retail investor sentiment and strict short-sell prohibitions. In other words, China stock markets offer an ideal setting for studying the importance of investor sentiment because the markets are dominated by individual (sentimental) investors, stock prices are highly speculative, and betting against sentimental investors is costly and risky due to the strict short-sell constraints.

In the language of modern behavioral finance, there are severe limits to arbitrage in China stock markets. Thus, I expect that the impact of investor sentiment is more pronounced in China stock markets than any other more established markets. Liu & Shrestha (2008) document the dominance of individual investors. Jianping, Jose, Wei (2005) analyze above mentioned theories and the implications for Chinese A-B share price fluctuation and premia.

The thesis consists of two parts. In the first part, I construct an investor sentiment index, which is based on the common variation in four underlying proxies for sentiment: turnover of tradable share, the number and average first-day return of IPOs, and the number of newly opened accounts.

In the second part of analysis, with the aid of this index, I identify stocks that are likely to be most affected by investor sentiment. The mispricing is the result of an uninformed demand shock along with limits to costless arbitrage and, in fact, these two channels lead to similar results because stocks which are likely to be most sensitive to speculative demand, those with highly subjective valuations, also tend to be the riskiest ones to arbitrage. Therefore, I expect that small, high volatility, unprofitable, high growth, and distressed firms are more subject to investor sentiment.

Prior work suggests a number of proxies for investor sentiment. There are no definitive or uncontroversial measures, however. I form a composite index of sentiment (*SENTIMENT*) based on four proxies for sentiment: turnover of tradable share (*TURN*), the number of initial public offerings (*NIPO*), the average first-day return of IPOs (RIPO), and the number of newly opened accounts (NACC).

TURN is the ratio of reported share volume to average share listed. Baker and Stein (2004) suggest that turnover, or more generally liquidity, can serve as a sentiment index. In a market with short-sales constraints, irrational investors participate, and thus add liquidity, only when they are optimistic. Hence high liquidity is a symptom of overvaluation.

RIPO and NIPO are included because the IPO market is often viewed as sensitive to sentiment, and high first-day returns on IPOs are cited as a measure of investor enthusiasm. The theoretical motivation for using the number of IPOs is that insiders and long-run shareholders have strong incentives to time the equity market for when valuation are greatest, which is presumably when sentiment is highest. Low long-run returns to IPOs have been noted by Stigler (1964), Ritter (1991), and Loughran, Ritter, and Rydkvist (1994), which is ex post evidence of successful market timing relative to a market index.

NACC reflects the sentiment of investor directly. The majority of investors in

China stock market are Individual ones and, the number of newly opened accounts reflects the focus from individual investors. Therefore, NACC indicates investor sentiment directly.

The Sentiment Index has significant positive correlation with stock market index, which reaches 90% between them. I then examine how the cross-sectional stock returns vary with monthly sentiment.

To this end, I collect financial statement data and the stock prices and returns from CSMAR for the period 2003-2010. Every month, I sort stocks into ten portfolios by one of stock characteristics, and then according to high and low sentiments, respectively. At last, I calculate the equal-weighted portfolio returns and plot them in bar charts.

The empirical results reveal that small size, risky, unprofitable, and high growth stocks are more affected by investor sentiment. Specifically, these stocks exhibit higher returns than those their counterparts with large size, low risk, high profitability, and low growth when sentiment is high. In addition, those stocks achieve lower returns when sentiment is low, consistent with the conjecture that these stocks are difficult to value and risky to arbitrage. Taken together, these results are consistent with the theories on investor sentiment and the limits of arbitrage.

I then perform the regression analysis to examine the effects of investor sentiment on stock returns after controlling for traditional asset pricing factors, including size, book-to-market, and momentum factors. Specifically, I regress various high-minus-low portfolio returns on the sentiment indexes (in terms of sensitivity to sentiment) and SMB, HML and UMD. The high-minus-low portfolios are formed on firm characteristics, such as firm size (Size), Volatility (Risk), etc. High portfolios include firms in the top one third in all A-share stocks according to one characteristic in a certain month; Low portfolios include firms in the bottom one third.

Consistent with previous non-parametric analysis, the regression results reinforce the findings that the impact of investor sentiment on stock returns is more pronounced for small, unprofitable stocks and those with high volatility and high growth potentials. The results provide practical insight for both researchers and practitioners. In asset pricing, the results suggest that models and expected returns should incorporate a role for investor sentiment.

The remainder of the thesis is organized as follows. Section 2 details the distinctive characteristics of China stock market. Section 3 reviews the literature reviews on investor sentiment. Section 4 describes methodology and data for the empirical analysis. Section 5 presents the empirical results. Section 6 summarizes and concludes the whole thesis.

2. Background of China Stock Market

The Shanghai Stock Exchange was established on December 19, 1990, followed by the Shenzhen Stock Exchange on July 3, 1991. The Chinese Securities Regulatory Committee (CRSC), the regulatory body that supervises new stock listing and daily trading activities was set up in 1987. At that time, A-Shares were only allowed to be traded by domestic residents and were dominated in the Chinese currency, Renminbi. In 1992, in hope of attracting international capital, B-shares were introduced to boost the Chinese stock market. B-Shares can only be traded by foreign investors and they are denominated in US dollars and HK dollars on the SHSE and SZSE respectively.

Although the two types of shares own the same rights for voting and dividends, they are subjected to different disclosure standards. For companies issuing A-shares, only one set of financial statements are required. These financial statements are prepared based on the Chinese GAAP and are audited by domestic CPA firms. However, B-shares must prepare an additional set of financial statements which is based on the International Accounting Standards (IAS) and will be audited by the Big Four firms. Trading restrictions were later removed on February 20, 2001 to allow domestic investors to trade in B-shares.

The domestic class A-shares is segmented into three categories: state shares held by the central or local government or solely state-owned enterprises that are not tradable; legal person or institutional shares held by joint stock companies and non-bank financial institutions that are also not traded on the stock exchanges; and publicly tradable shares mostly held by individual investors. This segmentation has resulted in strong governmental influence in the market. Usually, when a state-owned company becomes listed, only one third of its shares are issued to the public while the rest are owned by the government or remain with the business.

When the stock exchanges were first instituted, the government imposed price-change limits to both: 0.5% for SZSE and 1% for SHSE. The price limits affected the stock markets negatively. In May 1991, SZSE was completely deregulated while SHZE relaxed its price change to 5% on February 18, 1992. On May 20, 1992, it was announced that the price-change limits would be completely removed. However on December 16, 1996. New daily price-change limits of 10% were imposed on both exchanges again to reduce volatility.

Capital controls in China do not allow the transmission of international price movements to China. This results in different pricing policy in China as compared to the international markets. Hence China does not have the opportunity to learn and adopt international standards, leading to a weak disclosure and legal system. Despite the introduction of Company Law and Securities law, the Chinese stock market remains dominated by speculation, as can be seen on the SHSE. Whereby IPOs are found to be severely underpriced (Mok and Hui(1998)).

To summarize, deregulation and reregulation have been a major feature of the Chinese market. IPOs are strictly regulated, with a quota being set for new listings each year. The government completely controls the size of the market through segmentation of shares, pace of issuance and allocation of resources.

3. Literature Review

Black (1986) comment that noise in the sense of a large number of small events is often a cause factor much more powerful than a small number of large events can be. Noisy traders increase the liquidity of stock market but also decrease market efficiency. In 1990, De Long, Shleifer, Summers, Waldmann(1990) present a model of an asset market in which irrational noise traders with erroneous stochastic beliefs both affect prices and earn higher expected returns, named DSSW model and conclude that investor sentiment is the systematic factor to affect stock price in the market with limited arbitrage.

Barberis, Shleifer, and Vishy (1998) build a behavioral model which successfully captures long-term reversals, momentum, and post-earnings announcement drift. Daniel, Hirshleifer, and Subrahmanyam (1998, 2001) construct another successful behavioral model with emphasis on self-attribution bias which produces momentum and post-earnings announcement effect. Hong and Stein(1999,2000) model a market populated by two groups of boundedly rational agents: “news watchers” and “momentum traders.” and provide some empirical implications, such as in what stocks do momentum strategies work best, differential dynamics in response to public versus private news shocks, etc.

Besides the theoretical model, many empirical research focus on relation between investor sentiment and stock returns. Otoo(1999) examines the relationship between movements in consumer sentiment and stock prices. At the aggregate level, the two share a strong contemporaneous relationship—an increase in equity values boosts sentiment. Brown(1999) shows that unusual levels of individual investor sentiment are associated with greater volatility of closed-end investment funds. Furthermore, this volatility occurs only when the markets is open and is associated with heightened trading activity. It persists after controlling for market wide volatility and changes in fund discounts. Fisher and Statman(2000) show that the sentiment of Wall Street strategists is unrelated to the sentiment of individual investors or that of newsletter writers, although the sentiment of the last two groups is closely related. They also found a negative relationship between the sentiment of each of these three groups and future stock returns, and the relationship is statistically significant for Wall Street strategists and individual investors. Brown and Cliff (2004) investigate investor sentiment and its relation to near-term stock market returns. They found that many indirect measures of sentiment are related to direct measures (surveys) of investor sentiment. However, past market returns are also an important determinant of sentiment. Although sentiment levels and changes are strongly correlated with contemporaneous market returns, their tests show that sentiment has little predictive power for near-term future stock returns.

Lemmon and Portniaguina (2006) found that investor sentiment measured using consumer confidence forecasts the returns of small stocks and stocks with low institutional ownership. Qiu and Welch (2006) have similar conclusion and found that consumer confidence correlates well with direct investor sentiment survey data from UBS/Gallup.

Above research focus on US data mainly. Schmeling (2009) examines whether consumer confidence – as a proxy for individual investor sentiment – affects expected stock returns internationally in 18 industrialized countries. When sentiment is high, they found that future stock returns tend to be lower and vice versa. This relation also holds for returns of value stocks, growth stocks, small stocks, and for different forecasting horizons. Schmeling (2007) also show that institutional and individual sentiments seem to proxy for smart money and noise trader risk, respectively. Using bias-adjusted long-horizon regressions, they show that institutional sentiment forecasts stock market returns at intermediate horizons correctly, whereas individuals consistently get the direction wrong.

There are fewer researches on cross-section of stock returns relatively. Lee, Shleifer, Thaler (1991) focus on closed-end fund puzzle and found that both closed-end funds and small stocks tend to be held by individual investors, and that the discounts on closed-end funds narrow when small stocks do well. Baker and

Wurgler (2006; 2007) construct an investor sentiment index and test how the index affects cross-section of stock return. They use yearly data and find that highly speculative stocks are more affected by investor sentiment, such as small size stock, young stock, high volatility stock, and extreme growth stocks. They also found that investor sentiment has predictability for time series market return.

How investor sentiment affects stock return volatility and return is another research area. French (1987) find evidence that the expected market risk premium (the expected return on a stock portfolio minus the Treasury bill yield) is positively related to the predictable volatility of stock returns. Lee, Jiang, Indro (2002) show that sentiment is a systematic risk that is priced. Excess returns are contemporaneously positively correlated with shifts in sentiment. Moreover, the magnitude of bullish (bearish) changes in sentiment leads to downward (upward) revisions in volatility and higher (lower) future excess returns. Verma (2007) find significant positive (negative) effects of investor sentiments on stock returns (volatilities) for both individual and institutional investors. There are greater positive effects of rational sentiments on stock returns than irrational sentiments. Conversely, there are significant (insignificant) negative effects of irrational (rational) sentiments on volatility.

4. Empirical Approach and Data

4.1 Characteristics and Return

The firm-level data are collected from CSMAR database. The sample includes all A-share stock in Shanghai Security Exchange and Shenzhen Security Exchange from 2003 to 2010. Following Fama and French (1992), I match accounting data for fiscal year-ends in calendar year t-1 to (monthly) returns from July t through June t+1, and I use their variable definitions when possible.

I use market value of tradable share as size and match size of June of year t to monthly return from July of year t through June of year t+1. Volatility (Sigma, Risk) is the standard deviation of monthly returns over the 12 months ending in June of year t. Then volatility (sigma, risk) is matched to monthly returns from July of year t through June of year t+1. Prior work argues that volatility is likely to be a good proxy for the difficulty of both valuation and arbitrage. Another risk indicator is the debt to asset ratio. I match Dec of year t-1 the debt to asset ratio to returns from July t through June t+1.

Profitability characteristics include the return on equity (E/BE). Similarly, I match Dec of year t-1 E/BE to returns from July t through June t+1. E is net profit attributable to owners of the parent company and BE is total equity attribute to owners of the parent company.

Characteristics indicating growth opportunities, distress, or both include book-to-market equity (B/M), sale growth rate (SG) and earning price ratio (E/P). I match Dec of year t-1 B/M to returns from July t through June t+1 as before. B is total shareholders' equity and M is total market value of the stock. SG is monthly growth rate of total operating revenue, and I match monthly growth rate to contemporaneous monthly returns. Finally, in terms of E/P, I match earning of Dec of year t-1 to the price from July t to June t+1. A high B/M value means distress and low value may indicate high growth opportunities. Similarly, High SG values reflect growth opportunities but low value may mean distress.

Table I shows summary statistics.

4.2 Investor Sentiment

Prior work suggests a number of proxies for investor sentiment. There are no definitive or uncontroversial measures, however. I form a composite index of sentiment (*SENTIMENT*) based on four proxies for sentiment: turnover of tradable shares (*TURN*), the number of initial public offerings (*NIPO*), the average first-day return of IPOs (*RIPO*), and the number of newly opened account(NACC). The sentiment proxies are collected monthly from 2003 to 2010. I first introduce each proxy separately, and then analyze how they are calculated into investor sentiment index.

TURN is the ratio of reported share value to average share listed from WIND database 2003-2010. Baker and Stein (2004) suggest that turnover, or more generally liquidity, can serve as a sentiment index. In a market with short-sales constraints, irrational investors participate, and thus add liquidity, only when they are optimistic. Hence high liquidity is a symptom of overvaluation. Supporting this, Jones (2001) has similar conclusion.

RIPO and NIPO are included because the IPO market is often viewed as sensitive to sentiment, and high first-day returns on IPOs are cited as a measure of investor enthusiasm. The theoretical motivation for using the number of IPOs is that insiders and long-run shareholders have strong incentives to time the equity market for when valuation are greatest, which is presumably when sentiment is highest. Low long-run returns to IPOs have been noted by Stigler (1964), Ritter (1991), and Loughran, Ritter, and Rydkvist (1994), which is ex post evidence of successful market timing relative to a market index. NIPO is equal-weighted monthly average first-day return, and NIPO is the monthly IPO number. Because of the change of IPO policy, there are no IPO between Oct. 2004-Jan. 2005, Jul. 2005-May 2006, Oct. 2008-Nov. 2008, and Jan. 2009-Jun.2009.

NACC reflects the sentiment of investor directly. The majority of investors in China stock market are Individual ones and, the number of newly opened account reflects the focus from individual investors. Therefore, NACC indicates investor sentiment directly.

Each sentiment proxy includes a sentiment component as well as unique, non-sentiment-related components. I use principal components analysis to isolate the common component. Another concern is the lead-lag issue. Some proxies may reflect a given shift in sentiment earlier than others.

I form a composite investor sentiment index from the four proxies using principal component analysis and the procedure is as follows. I estimate the first principal component of the four proxies and their lags. This gives us a first-stage index with 8 loadings, one for each of the current and lagged proxies. I then compute the correlation between the first-stage index and the current and lagged values of each of the proxies. Finally, I define SENTIMENT as value-weighted average of the first three principal component of the correlation matrix of four variables—each respective proxy's lead or lag, whichever has higher correlation with the first-stage index—rescaling the coefficients so that the index has unit variance. The first three eigenvalues of the correlation matrix of four variables work as weight. This is what I improve from Baker and Wurgler (2006 2007), who only use the first principal component. I also use first principal component as robust test and there are no differences in conclusion. The final result is as follows.

The first principal component:

$$Com1 = 0.5584 \times RIPO_t + 0.5527 \times TURN_{t-1} + 0.0211 \times NIPO_t + 0.6183 \times NACC_t$$

The second principal component:

$$Com2 = -0.1566 \times RIPO_t - 0.0666 \times TURN_{t-1} + 0.9710 \times NIPO_t + 0.1678 \times NACC_t$$

The third principal component:

$$Com3 = 0.6721 \times RIPO_t - 0.7371 \times TURN_{t-1} + 0.0491 \times NIPO_t + 0.0502 \times NACC_t$$

And the final result is:

$$SENTIMENT_t = 0.5836 \times Com1 + 0.2760 \times Com2 + 0.1504 \times Com3$$

$$SENTIMENT_t = 0.3837 \times RIPO_t + 0.1933 \times TURN_{t-1} + 0.2877 \times NIPO_t + 0.4147 \times NACC_t$$

Here each of the index components has first been standardized. We can see that every proxy has corrected sign, which means positive relation for each. The first three principal components explain 95% of the sample variance, so I conclude that one factor captures much of the common variation. The correlation between the 8-term first-stage index and the SENTIMENT index is 0.97, suggesting that little information is lost in dropping the six terms with other time subscripts.

Because the above sentiment cannot distinguish between a common sentiment component and a common business cycle component, I construct a second index that explicitly removes macro-index from each of the proxies prior to the principal components analysis.

Specifically, I regress each of the four raw proxies on china macro-index to eliminate business cycle effect. The residuals are labeled with a superscript$'$ and can be cleaner proxies for investor sentiment. I form an index of the orthogonalized proxies following the same procedure as before. The result is as follows:

The first principal component:

$$Com1' = 0.5565 \times RIPO_t{}' + 0.5467 \times TURN_{t-1}{}' - 0.1961 \times NIPO_t{}' + 0.5941 \times NACC_t{}'$$

The second principal component:

$$Com2' = -0.1144 \times RIPO_t{}' + 0.1396 \times TURN_{t-1}{}' + 0.9401 \times NIPO_t{}' + 0.2891 \times NACC_t{}'$$

The third principal component:

$$Com3' = 0.6760 \times RIPO_t' - 0.7140 \times TURN_{t-1}' + 0.1643 \times NIPO_t' + 0.0782 \times NACC_t'$$

And the final result is:

$$SENTIMENT_t' = 0.5859 \times Com1' + 0.2675 \times Com2' + 0.1466 \times Com3'$$

$$SENTIMENT_t' = 0.3945 \times RIPO_t' + 0.2529 \times TURN_{t-1}' + 0.1606 \times NIPO_t' + 0.4368 \times NACC_t'$$

Here, the first three principal components explain 94% of the sample variance of the orthogonalized variables. Furthermore, in terms of the sign of proxies, SENTIMENT' remains the same sign as SENTIMENT.

Table II summarizes and correlates the sentiment measures, and Figure 1 plots them. Table II shows that the orthogonalized proxies are more correlated with each other than are the raw proxies. If the raw variables were driven by macro-factor, which I remove in the second construction, rather than investor sentiment, the result should be the opposite.

There are also other proxies that are reasonably included in the sentiment index construction. The main constraint is availability and controversies. For example, close-end fund discount is widely used as one proxy in US market to measure investor sentiment. However, there are few relevant data during test period and full of controversies about close-end fund discount in china stock market, where fund industry is far from maturity.

5. Empirical Tests

5.1 Sorts

Table III shows conditional characteristics effects in a simple, nonparametric way. Every month, I sort stocks into ten portfolios by one of the stock characteristics, and then sort according to high and low sentiments, respectively. I compute the equal-weighted average firms across portfolios and look for patterns. In particular, I identify time-series changes in cross-sectional effects from the conditional difference of average returns across deciles. Figure 2 graph bar charts for Table III. The blue (positive) bars are returns in positive $\mathrm{SENTIMENT}'$ periods, and the red (negative) bars are returns in negative sentiment periods, and the solid line is the difference.

The first row of Table III shows the effect of size, conditional on sentiment. We note that returns of small size stocks are high than large size stocks when sentiment is high, suggesting small size stocks are preferred by investor compared with large size stock when sentiment is high, consistent with our previous prediction. In the other hand, when sentiment is low, large size stocks earn higher returns. If we look at the solid line, which represents difference across high sentiment and low sentiment, in figure 2, we can see that small size stocks are more affected by sentiment compared with large size stocks, also consistent with our previous prediction, because small size stocks are more difficult to value and arbitrage.

The next rows of Table III indicate that the cross-sectional effect of return volatility is conditional on sentiment in the hypothesized manner. High volatility stocks are chased by investor when sentiment is high and earn returns of 4.5% per month, compared with 3.9% for low volatility stocks. When sentiment is low, the pattern reveres: high risky stocks earn lower returns. In terms of difference line, high volatility stocks are more affected by sentiment. Loosely speaking, when sentiment is high, “riskier” stocks earn higher returns. When sentiment is low, they earn lower returns. A natural interpretation is that riskier stocks are relatively hard to value and relatively hard to arbitrage, making them more affected by sentiment. Another risky indicator Debt Asset ratio shows a similar pattern.

The next rows examine profitability indicator return of equity (E/BE). When sentiment is high, monthly returns of high profit firms are 0.8% lower than low profit firms. When sentiment is low, monthly returns of high profit firms are 1.3% higher than low profit firms. Again, this is consistent with low profitable firms being generally more difficult to value and to arbitrage, thus exposing them more to sentiment fluctuations.

The remaining variables book-to-market ratio, earning price ratio, and sale growth rate-also show intriguing patterns. They all indicate that high growth stocks earn high returns when sentiment is high and high growth stocks are more affected by fluctuations in sentiment, because high growth firms are harder to value, and perhaps to arbitrage.

In all, all indicators support our hypothesis that small size, risky, and extreme growth stocks are more affected by sentiment because they are harder to value and to arbitrage.

5.2 Predictive Regressions for Long-Short Portfolios

Next, I explore how sentiment affects equal-weighted portfolios that are long on stocks with high values of a characteristic and short on stocks with low values. For example, we can see that high-risk stocks earn higher returns than low-risk ones when sentiment is high, so sentiment likely has effect on a long-short portfolio formed on risk. A regression analysis allows us to conduct formal significance test and incorporate the continuous nature of the sentiment indexes.

Table IV graphs correlations among the average monthly returns on various long-short portfolios over time.

Next, I will explore whether sentiment affects various long-short portfolios return analyzed in Table IV and also whether lagged sentiment predicts various long-short portfolio returns. Here, I also control for traditional asset pricing factors, including market risk premium (RMRF), size (SMB), book-to-market (HML), and momentum (UMD) factors. As described in Fama and French (1993), SMB is the return on portfolio of small and big ME stocks that is separate from returns on HML, where HML is constructed to isolate the difference between high and low BE/ME portfolios. I exclude SMB and HML from the right side when they are the portfolios being forecast. The specific regressions are as follows:

$$R_{i=High,t} - R_{i=Low,t} = \alpha + \beta Sentiment_t + \varepsilon_t$$

$$R_{i=High,t} - R_{i=Low,t} = \alpha + \beta^{'} Sentiment_{t-1} + \varepsilon_t$$

$$R_{i=High,t} - R_{i=Low,t} = \alpha + \beta_1 Sentiment_t + \beta_2 RMRF_t + \beta_3 HML_t + \beta_4 SMB_t + \beta_5 UMD_t + \varepsilon_t$$

$$R_{i=High,t} - R_{i=Low,t} = \alpha + \beta_1^{'} Sentiment_{t-1} + \beta_2^{'} RMRF_t + \beta_3^{'} HML_t + \beta_4^{'} SMB_t + \beta_5^{'} UMD_t + \varepsilon_t$$

$$R_{i=High,t} - R_{i=Low,t} = \alpha + \beta Sentiment\_ex\_macro_t + \varepsilon_t$$

$$R_{i=High,t} - R_{i=Low,t} = \alpha + \beta^{'} Sentiment\_ex\_macro_{t-1} + \varepsilon_t$$

$$R_{i=High,t} - R_{i=Low,t} = \alpha + \beta_1 Sentiment\_ex\_macro_t + \beta_2 RMRF_t + \beta_3 HML_t + \beta_4 SMB_t + \beta_5 UMD_t + \varepsilon_t$$

$$R_{i=High,t} - R_{i=Low,t} = \alpha + \beta_1^{'} Sentiment\_ex\_macro_{t-1} + \beta_2^{'} RMRF_t + \beta_3^{'} HML_t + \beta_4^{'} SMB_t + \beta_5^{'} UMD_t + \varepsilon_t$$

Here i is various firm characteristics, and $R_{i=High,t} - R_{i=Low,t}$ are monthly equal-weighted returns of long-short portfolios sorted on the characteristics i.

Table V shows the results. The results provide formal support to our previous conclusion. For example, the negative coefficient of SENTIMENT ($SENTIMENT'$) in the first panel shows that returns of small firms are relatively high when sentiment is high (last month sentiment is high). In terms of magnitudes, one-standard deviation increases in sentiment induce a -1.2% lower monthly return on the large minus small portfolio. The Risk indicator, Volatility and Debt/Asset, also support our previous conclusion: high risk stocks earn relatively higher returns than low risk stocks when sentiment is high.

For profitability, I run regressions to show the difference between the high profitable firm stock returns and low profitable firm stock returns. The results show that sentiment indeed has effect on the portfolio. High-profit firms earn lower return relatively compared with low-profit firms when sentiment is high, no matter if I control for RMRF, SMB, HML and UMD.

Finally, book to market ratio, sale growth rate and earning price ratio also show consistent result as using nonparametric method. High growth firms earn relatively higher return compared with low growth firms when sentiment is high.

In summary, the regression analysis confirms significance of the patterns suggested in the sorts. When sentiment is high, small, high volatility, unprofitable, and high growth firms earn higher returns. Furthermore, the results support that sentiment has stronger effect on stocks that are hard to value and hard to arbitrage.

6. Conclusion

Traditional finance theories assume rational investors and market efficiency. In traditional finance theory, investor sentiment should not affect cross-section of stock returns. In this paper, I use theoretical arguments, investor sentiment index and empirical analysis to show that investor sentiment has cross-sectional effects.

Speculation activities can drive asset price away from its fundamental value. Yet, in an imperfect capital markets, arbitrageurs may find it too costly and risky, if not impossible, to engage in arbitrage activities. Thus, asset prices can be persistently affected by investment sentiment. Because some certain categories of stocks are more difficult to value and arbitrage, I expect these certain categories of stocks are more exposed to investor sentiment. Consistent with our hypothesis, I find that small size, high volatility, low-profit and high-growth firms are more susceptible to investor sentiment and earn higher returns when sentiment is high. In the other hand, such categories of stocks earn lower returns than those their counterparts with large size, low risk, high profitability, and low growth when sentiment is low.

References

Baker, Malcolm, and Jeremy Stein, “Market Liquidity as a sentiment indicator”, *Journal of financial markets*, Vol. 7, 271-299, 2004

Brunnermeier, Markus, and Lasse Pedersen, “predatory trading”, *Journal of Finance*, Vol. 60, 1825-1863, 2005

Charles M. C. Lee, Andrei Shleifer and Richard H. Thaler, “Investor Sentiment and the Closed-End Fund Puzzle”, *Journal of Finance*, Vol. 46, No. 1, 1991

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert Waldmann,, “Noise trader risk in financial markets”, *Journal of Political Economy*, Vol. 98, 703-738, 1990

Elton, Edwin J., Martin J. Gruber, and Jeffrey A. Busse, “Do investors care about sentiment?”, *Journal of Business*, Vol. 71, 477-500, 1998

Fama, Eugene F., "Efficient Capital Markets: A Review of Theory and Empirical Work". *Journal of Finance*, Vol. 25, No. 2, 383–417, 1970.

Fama, Eugene F. and Kenneth R. French, “Common risk factors in the returns on stocks and bonds”, *Journal of Financial Economics*, Vol. 33, 3-56, 1993

Fischer Black, “Noise”, *Journal of Finance*, Vol. 41, No. 3, 529-543,1985

Gregory W. Brown, “Volatility, Sentiment and Noise Traders”, *Financial Analysts Journal*, Vol. 55, No. 2, 1999

Gregory W. Brown and Michael T. cliff, “investor sentiment and the near-term stock market”, *Journal of Empirical Finance*, Vol. 11, No. 1, 2004

Hong,H., Lim,T., and Stein, “bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies”, *NBER Working Paper*, No. 6553, 2000

Hong, H., and stein, J.C. (1999) “A unified theory of underreaction, momentum trading, and overreaction in asset markets”, *Journal of Finance*, Vol. 54, No 6, 2143-2184, 1999

Jianping Mei, Jose Scheinkman and Wei Xiong, “Speculative Trading and Stock Prices: Evidence from Chinese A-B Share Premia", *NBER Working Paper*, No. 11362, 2005

Jegadeesh and Titman, "Returns to buying winners and selling losers: Implications for stock market efficiency", *Journal of Finance*, Vol. XLVIII, No. 1, 1993

Jones, Charles, and Owen Lmont, "Short sale constraints and stock returns", *Journal of Financial Economics,* Vol. 66, 207-239, 2002

Kenneth L. Fisher and Meir. Statman, "Investor sentiment and stock returns", *Financial Analysts Journal*, Vol. 5, No. 2, 2000

Ljungqvist, Alexander, and William Wilhelm, "IPO pricing in the dot-com bubble", *Journal of Finance*, Vol. 58, 723-752, 2003

Maik Schmeling, "Institutional and individual sentiment: Smart money and noise trader risk?", *International Journal of Forecasting*, Vol. 23, No. 1, 2007

Maik Schmeling, "Investor sentiment and stock returns: Some international evidence", Journal of Empirical Finance, Vol. 16, No. 3, 2009

MARKUS K. BRUNNERMEIER, and STEFAN NAGEL, "Hedge Funds and the Technology Bubble", *Journal of Finance*, Vol. 59, 2013-2040, 2004

Michael Lemmon and Evgenia Portniaguina, "Consumer Confidence and Asset Prices: Some Empirical Evidence", The review of Financial Studies, Vol. 19, No 4, 2006

Ming-hua Liu and Keshab M. Shrestha, "Analysis of long term relationship of macro-economic variables and the Chinese stock market", *Managerial Finance*, Vol.34, No. 11 744-755, 2008

Mitchell, Mark L., Todd Pulvino, and Erik Stafford, "Limited arbitrage and equity markets", *Journal of Finance*, Vol. 27, 551-581, 2002

Ofek, Eli, and Matthew Richardson,"DotCom mania: The rise and fall of internet stocks", *Journal of Finance*, Vol. 58, 1113-1138, 2002

Pan Guoling, "A Study of the Stock Market Bubble", *Journal of Finance,* Vol. 7, No. 18, 2000

Rahul Verma and Priti Verma, "Noise trading and stock market volatility", Journal of Multinational Financial Management, Vol. 17, No 3, 2007

Robin Greenwood and Stefan Nagel, “Inexperienced investors and bubbles”, *Journal of Financial Economics*, Vol. 93, 239-258, 2009

Sharpe, William F. "Capital Asset Prices - A Theory of Market Equilibrium under Conditions of Risk". Journal of Finance XIX (3): 425–42. 1964

Sloan, R.G., “Do stock prices fully reflect information in accruals and cash flows about future earnings?”, *Accounting Review* 71, 289-315, 1996

Wayne Y Lee, Christine X Jiang and Daniel C Indro, “Stock market volatility, excess returns, and the role of investor sentiment”, *Journal of Banking & Finance*, Vol. 26, No. 12, 2002

Wurgler, Jeffrey, and Katia Zhuravskaya, “Does arbitrage flatten demand curves for stocks?”, *Journal of Business*, Vol. 75, 583-608, 2002

Table I: Summary Statistics, 2003-2010

Panel A summarizes the Size and Risk variables. Size is market value of tradable A-share, and Sizes in June of year t are matched with returns from July t to June t+1. Volatility is the standard deviation of monthly returns over the 12 months ending in June of year t, and Volatility in June of year t is matched with returns from July t to June t+1. Debt Asset ratio in Dec. of year t-1 is matched with returns from July t to June t+1. EBE is return of equity, and EBE in Dec. of year t-1 is matched with returns from July t to June t+1. BM is book to market ratio, and BM in Dec. of year t-1 is matched with returns from July t to June t+1. SG is monthly growth rate of total operating revenue, and SG is matched to contemporaneous monthly returns. EP is earning price ratio. And Earning in Dec of year t-1 is matched to the price from July t to June t+1.

Variable	Obs	Mean	Std. Dev.	Min	Max
	Panel A: Size and Risk				
Size	144649	2680530	1.26e+07	8488.97	1.02e+09
Volatility	142394	0.1436999	0.1074535	0.0069082	6.379451
DEBTASSET	141809	0.6079167	2.895785	0.0017253	229.27
	Panel B: Profitability				
EBE	141797	0.0390911	2.437433	-134.7938	100.7154
	Panel C: Growth Opportunities and Distress				
BM	140666	1.035031	4.206641	-74.36741	117.5585
SG	149780	0.569916	90.72037	-21.01122	29274.96
EP	140339	5.105467	786.964	-73059.99	14048.95

Table II: Investor Sentiment Data, 2003--2010

Means, standard deviations, minimum, maximum and correlations for measures of investor sentiment. In the first graph, I present raw sentiment proxies. RIPO is average monthly first-day return of IPO. TURN is the reported share volume to average shares listed. NIPO is the monthly number of IPO. NACC is monthly new opening account number. SENTIMENT is value weighted first three principal component of the above four proxies, and the first three eigenvalues of the correlation matrix of four proxies work as weight. SENTIMENT' is value weighted first tree principal component of the above four orthogonalized proxies.

Variable	Mean	Std. Dev.	Min	Max
	Raw Data			
$RIPO_t$	95.33419	64.6761	12.03078	334.6403
$TURN_{t-1}$	0.3797853	0.2185405	0.1174301	1.003343
$NIPO_t$	10.09524	6.272609	2	24
$NACC_t$	1103232	1285988	51306	5594364
	Controlling for Macro-Index			
$RIPO_t$	0	61.57144	-115.5573	219.9008
$TURN_{t-1}$	0	0.2140577	-0.2917455	0.5946835
$NIPO_t$	0	5.661416	-9.917702	11.36085
$NACC_t$	0	1172062	-1260027	4041125

Variable	Correlations with Sentiment		Correlations with Proxies			
	SENTIMENT	SENTIMENT’	$RIPO_t$	$TURN_{t-1}$	$NIPO_t$	$NACC_t$
	Raw Data					
$RIPO_t$	0.7957	0.7776	1			
$TURN_{t-1}$	0.7006	0.7387	0.4673	1		
$NIPO_t$	0.3418	0.0164	-0.0928	-0.0423	1	
$NACC_t$	0.9156	0.8335	0.6796	0.6625	0.1630	1
	Controlling for Macro-Index					
$RIPO_t'$	0.6782	0.8169	1			
$TURN_{t-1}'$	0.6145	0.7542	0.4350	1		
$NIPO_t'$	0.1450	0.0182	-0.2614	-0.1459	1	
$NACC_t'$	0.7834	0.9145	0.6381	0.6492	0.172	1

Table III: Return by Sentiment Index and Firm Characteristics, 2003-2010

	$SENTIMENT'$	Decile										comparisons
		1	2	3	4	5	6	7	8	9	10	10-1
Size	Positive	6%	5.4%	4.8%	4.9%	4.6%	4.4%	4.1%	4.1%	3.9%	3.2%	-2.8%
	Negative	-1%	-1.3%	-1.3%	-1.2%	-1.2%	-1.2%	-0.9%	-0.9%	-0.7%	-0.4%	0.6%
	Difference	7%	6.7%	6.3%	6.1%	5.8%	5.6%	5%	5%	4.6%	3.6%	-3.4%
Risk	Positive	3.9%	4.8%	4.4%	4.3%	4.5%	4.7%	4.8%	4.7%	4.6%	4.5%	0.6%
	Negative	-0.7%	-0.6%	-0.8%	-0.9%	-0.9%	-1.1%	-1%	-1%	-1.4%	-1.4%	-0.7%
	Difference	4.7%	5.5%	5.2%	5.2%	5.4%	5.8%	5.8%	5.8%	6%	6%	1.3%
D/A	Positive	4.5%	4.4%	4.3%	4.4%	4.4%	4.3%	4.4%	4.6%	4.7%	5.7%	1.2%
	Negative	-0.8%	-0.8%	-0.5%	-0.7%	-1%	-1%	-0.8%	-1%	-0.9%	-1.1%	-0.3%
	Difference	5.3%	5.3%	4.8%	5.1%	5.4%	5.3%	5.2%	5.5%	5.7%	6.7%	1.4%
E/BE	Positive	5.6%	4.7%	4.6%	4.6%	4.6%	4.2%	4.4%	4%	4.1%	4.8%	-0.8%
	Negative	-1.7%	-1.1%	-1%	-1%	-0.9%	-0.9%	-0.5%	-0.4%	-0.4%	-0.4%	1.3%
	Difference	7.3%	5.8%	5.7%	5.6%	5.5%	5.1%	5%	4.6%	4.6%	5.2%	-2.1%
B/M	Positive	4.6%	4.6%	4.6%	4.4%	4.6%	4.5%	4.8%	4.9%	4.6%	3.6%	1%
	Negative	-1.9%	-1.2%	-0.9%	-1.1%	-0.8%	-0.7%	-0.8%	-0.7%	-0.3%	-0.1%	1.8%
	Difference	6.6%	5.9%	5.5%	5.5%	5.4%	5.1%	5.6%	5.6%	4.9%	3.7%	2.9%
SG	Positive	4.7%	3.7%	4.2%	3.8%	4.1%	4.1%	4.1%	4.1%	4.8%	5.9%	1.2%
	Negative	0.4%	-1.3%	-1.2%	-1%	-1%	-0.9%	-1.1%	-1.1%	-1.4%	-1.6%	-2%
	Difference	4.3%	5%	5.4%	4.7%	5.1%	5%	5.2%	5.2%	6.1%	7.6%	3.3%
E/P	Positive	5.6%	6.6%	5.9%	5.8%	5.4%	4.7%	3.9%	3%	2.4%	2.4%	-4.6%
	Negative	-2%	-0.5%	-0.6%	-0.4%	-0.4%	-0.6%	-0.8%	-1.1%	-1%	-1%	1%
	Difference	7.6%	7.1%	6.5%	6.2%	5.9%	5.3%	4.7%	4.1%	3.4%	3.4%	-4.2%

Table IV: Correlations of Portfolio Returns, 2003-2010

Correlations among characteristics-based portfolios. The sample period includes monthly returns from 2003 to 2010. The long-short portfolios are formed on firm characteristics: firm size (Size), Volatility (Risk), Debt Asset ratio (Debt/Asset), Return of Equity (E/BE), Book-to-Market ratio (B/M), Sale Growth rate (SG), Earning Price ratio (E/P). High is defined as a firm in the top one third in all A-share stocks according to one characteristic in a certain month; Low is defined as a firm in the bottom one third.

		Size	Risk	Debt/Asset	E/BE	B/M	SG	E/P
Size	High-Low	1						
Risk	High-Low	-0.36	1					
Debt/Asset	High-Low	-058	0.50	1				
E/BE	High-Low	0.81	-0.64	-0.66	1			
B/M	High-Low	0.24	0.04	-0.023	-0.11	1		
SG	High-Low	0.20	-0.15	-0.1	0.23	-0.11	1	
E/P	High-Low	0.72	-0.46	-0.60	0.81	0.37	0.014	1

Table V: Time Series Regressions of Portfolio Returns, 2003—2010

$$R_{i=High,t} - R_{i=Low,t} = \alpha + \beta_1^{'} Sentiment_{t-1} + \beta_2^{'} RMRF_t + \beta_3^{'} HML_t + \beta_4^{'} SMB_t + \beta_5^{'} UMD_t + \varepsilon_t$$

Regressions of long-short portfolio returns on SENTIMENT, the market risk premium (RMRF), the Fama-French factors (HML and SMB), and a momentum factor (UMD) The sample period includes monthly returns from 2003 to 2010. The long-short portfolios are formed on firm characteristics: firm size (Size), Volatility (Risk), Debt Asset ratio (Debt/Asset), Return of Equity (E/BE), Book-to-Market ratio (B/M), Sale Growth rate (SG), Earning Price ratio (E/P). High is defined as a firm in the top one third in all A-share stocks according to one characteristic in a certain month, and Low is defined as a firm in the bottom one third. SENTIMENT′ Index is based on four sentiment proxies that have been orthogonalized to macro index; the components of SENTIMENT are not orthogonalized. The first and third sets of columns show univariate regression results, while the second and the fourth columns include RMRF, SMB, HML, and UMD. SMB (HML) is not included when SMB (HML) is the dependent variable.

	$SENTIMENT_{t-1}$		$SENTIMENT_{t-1}$ with RMRF, SMB, HML, UMD		$SENTIMENT'_{t-1}$		$SENTIMENT'_{t-1}$ with RMRF, SMB, HML, UMD	
	β'	p(β')	β'_1	p(β'_1)	β'	p(β')	β'_1	p(β'_1)
	Panel A: Size and Risk							
Size	-1.2%	0.07	-1.2%	0.08	-0.9%	0.17	-0.9%	0.2
Risk	0.4%	0.28	1%	0.03	0.4%	0.3	0.4%	0.34
Debt/Asset	1%	0	0.7%	0.16	1%	0	0.7%	0.17
	Panel B: Profitability							
E/BM	-0.7%	0.24	-0.4	0.45	-0.6%	0.28	-0.5%	0.31
	Panel C: Growth Opportunities and Distress							
B/M	-0.5%	0.22	-0.5%	0.16	-0.2%	0.66	-0.3%	0.45
SG	0.4%	0.1	0.1%	0.77	0.6%	0.01	0.5%	0.15
E/P	-1.8%	0.001	-0.5%	0.36	-1.6%	0.004	-0.17%	0.77

	$\underline{SENTIMENT_t}$		$SENTIMENT_t$ with RMRF, SMB, HML, UMD		$\underline{SENTIMENT'_t}$		$SENTIMENT'_t$ with RMRF, SMB, HML, UMD	
	β'	p(β')	β'_1	p(β'_1)	β'	p(β')	β'_1	p(β'_1)
	Panel A: Size and Risk							
Size	-1.1%	0.11	-1.1%	0.13	-0.6%	0.42	-0.5%	0.49
Risk	0.4%	0.28	0.5%	0.26	0.3%	0.43	0.2%	0.64
Debt/Asset	1.1%	0	1%	0.06	1.1%	0	0.8%	0.04
	Panel B: Profitability							
E/BM	-0.8%	0.16	-0.9	0.09	-0.6%	0.31	-0.9%	0.03
	Panel C: Growth Opportunities and Distress							
B/M	-0.6%	0.11	-0.8%	0.05	-0.3%	0.42	-0.5%	0.24
SG	0.5%	0.02	0.5%	0.24	0.7%	0.00	0.6%	0.06
E/P	-2.3%	0.00	-2.4%	0.00	-2%	0.00	-1.5%	0.00

Note: $SENTIMENT'_t$=$SENTIMENT\_ex\_macro_t$, $SENTIMENT'_{t-1}$=$SENTIMENT\_ex\_macro_{t-1}$ for short.

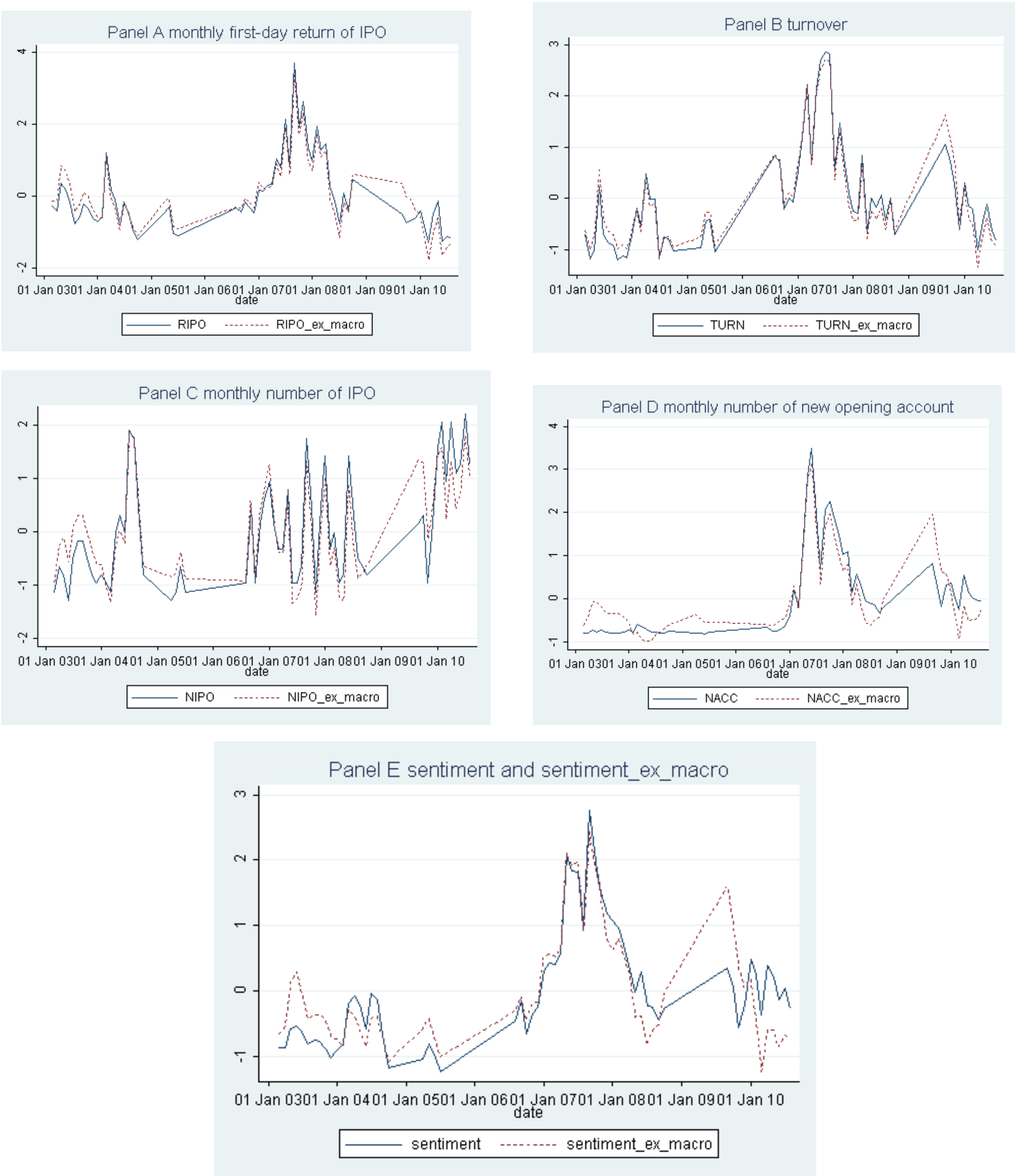

Figure 1: Investor Sentiment, 2003—2010

RIPO is average monthly first-day return of IPO. TURN is the reported share volume to average shares listed. NIPO is the monthly number of IPO. NACC is monthly new opening account number. SENTIMENT is value weighted first three principal component of the above four proxies, and the first three eigenvalues of the correlation matrix of four proxies work as weight. SENTIMENT’ is value weighted first tree principal component of the above four orthogonalized proxies. In the first four panels, the solid line is raw data after standardized. I regress each proxy on macro-index and the dashed line is the residual from the regression. The solid (dashed) line in the SENTIMENT (SENTIMENT’) respectively.

Figure 2: Two-way sorts: Returns by sentiment index and firm characteristics.

The bar charts are graphed according to Table III. The blue (positive) bars are returns in positive$SENTIMENT'$, the red (negative) bars are returns in negative $SENTIMENT'$, and the green line is the difference.