

Can Investor Sentiment Be a Momentum Time-Series Predictor?

Evidence from China

Xing Han^{a,*}, Youwei Li^b

^a *Department of Financial Economics, Ghent University, Sint-Pietersplein 5, 9000 Ghent, Belgium*

^b *School of Management, Queen's University Belfast, 185 Stranmillis Road, BT9 5EE Belfast, UK*

First version: 20 Apr. 2015

This version: 26 Aug. 2016

Abstract

This paper challenges the prevailing view that investor sentiment is a contrarian predictor of market returns at nearly *all* horizons. As an important piece of "out-of-sample" evidence, we document that investor sentiment in China is a reliable momentum signal at monthly frequency. The strong momentum predictability is robust under both single- and multi-regressor settings, and is statistically and economically significant both in and out of sample, enhancing portfolio performance as shown by our numerical examples. More importantly, we find a striking term structure that local sentiment shifts from a short-term momentum predictor to a contrarian predictor in the long run. Cross-sectional analysis reveals that sentiment is more of a small-firm effect. Finally, we confirm that global sentiment spills over to the local Chinese market as it predicts negatively future returns over the longer horizons and in the cross section.

JEL classification: C22; C53; G11; G12; G17

Keywords: Investor Sentiment; Return Predictability; Bias Correction; China

* Corresponding Author: Department of Financial Economics, Ghent University, Sint-Pietersplein 5, 9000 Ghent, Belgium. Tel.: +32-(0)9 264 78 95. Fax: +32 (0)9 264 89 95. E-mail: xing.han@ugent.be (X. Han).

Can Investor Sentiment Be a Momentum Time-Series Predictor?

Evidence from China

Abstract

This paper challenges the prevailing view that investor sentiment is a contrarian predictor of market returns at nearly *all* horizons. As an important piece of "out-of-sample" evidence, we document that investor sentiment in China is a reliable momentum signal at monthly frequency. The strong momentum predictability is robust under both single- and multi-regressor settings, and is statistically and economically significant both in and out of sample, enhancing portfolio performance as shown by our numerical examples. More importantly, we find a striking term structure that local sentiment shifts from a short-term momentum predictor to a contrarian predictor in the long run. Cross-sectional analysis reveals that sentiment is more of a small-firm effect. Finally, we confirm that global sentiment spills over to the local Chinese market as it predicts negatively future returns over the longer horizons and in the cross section.

JEL classification: C22; C53; G11; G12; G17

Keywords: Investor Sentiment; Return Predictability; Bias Correction; China

“Studies of the sentiment of investors are important for two reasons. First, they teach us about biases in the stock market forecasts of investors. Second, they teach us about opportunities to earn extra returns by exploiting those biases.”

—Kenneth L. Fisher and Meir Statman (2000)

1. Introduction

The effect of investor sentiment on asset prices has long been a debatable issue in financial economics. On the theoretical side, proponents of the Efficient Market Hypothesis (EMH) argue that the impact of sentiment is “at most” negligible as rational investors ruthlessly arbitrage away any sentiment-induced mispricing. Theorists from the behavioral camp, however, contend that sentiment induces systematic deviations from fundamental values due to the well-known “limit to arbitrage” argument ([De Long et al. 1990a](#)). Besides, [Brown and Cliff \(2005\)](#) and [Baker and Wurgler \(2006\)](#) point out that sentiment tends to have a long-lasting effect so that demand shocks of uninformed investors (eg. noise traders) may be correlated over time to give rise to strong and persistent mispricing. Intuitively, the impact of sentiment on returns should be stronger for markets with binding arbitrage constraints ([Kling & Gao 2008](#)) or that are more culturally collectivistic and highly prone to herding ([Schmeling 2009](#)). Still, whether investor sentiment matters for asset returns is an empirical issue.

This article relates to the above classical, yet contestable topic by investigating thoroughly the return predictability of investor sentiment over multiple horizons in China, the largest and most liquid emerging stock market. Prior works on the predictive power of sentiment mainly focused on US and other developed markets ([Brown & Cliff 2004, 2005](#); [Baker & Wurgler 2006, 2007](#); [Huang et al. 2014](#)), which uniformly find that sentiment is a strong long-term *contrarian* predictor ([Baker & Stein 2004](#); [Brown & Cliff 2005](#); [Baker et al. 2012](#)).¹ High sentiment is linked with low market returns in the long-run, implying (possible) market revision over time. Moreover, the contrarian predictability is more pronounced in the cross section ([Baker & Wurgler 2006](#); [Baker et al. 2012](#)). [Baker and Wurgler \(2007\)](#) attribute the contrarian

¹ It should be noted that the strikingly large long-run predictability documented in the empirical literature is not without controversy. See [Boudoukh et al. \(2008\)](#) for the caveats in long-horizon regressions.

predictability to the mean-reverting pattern in sentiment. That is, sentiment wanes in the long-run and market corrects its mispricing induced by sentiment.

Unfortunately, empirical evidence on the near-term return predictability, compared to the strong long-term predictability, is both thin and unclear. Using a wide range of sentiment proxies, including technical indicators and survey type data, [Brown and Cliff \(2004\)](#) find little, if any, evidence that sentiment forecasts returns at the weekly and monthly frequency in the US. [Huang et al. \(2014\)](#), however, find some statistical evidence that investor sentiment forecasts negatively future monthly returns with the aligned sentiment index, but not with the [Baker and Wurgler \(2006\)](#) sentiment index, which is more commonly used in the empirical studies.

The somewhat ambiguous empirical evidence on the short-term return predictability of investor sentiment, motivates our study as the widely perceived contrarian predictability of investor sentiment in the long run does *not* necessarily carry over to the short run. Moreover, compared to developed markets, emerging markets tend to have binding arbitrage constraints, pervasive local retail investors, and less openness to international (institutional) investors. Therefore, whether we could find similar contrarian predictability pattern in alternative market settings remains an interesting, yet unanswered question. The Chinese stock market provides a natural experiment for such an “out-of-sample” test in the spirit of [Lo and MacKinlay \(1990\)](#). It is well-known for its distinctive features from the US and other developed market: First, it is generally regarded as a highly speculative market dominated by the trading of individual investors. Individual investors (noise traders) have more market power in China than in developed markets. Second, it is highly restrictive for its stringent institutional settings such as strong short-sales constraints. The high short-selling costs also hinder institutional investors (rational investors) to engage in price stabilizing activities by trading against noise traders in China. All of these unique features, which are different from the US market (and other major developed markets), seem to tilt more towards the possibility that irrational sentiment leads to the expectation of increased future returns in the short-term.

It should be noted that a number of prior works have shed some lights on the sentiment impact on China’s stock market. For example, [Chi et al. \(2012\)](#) use mutual fund flow as a sentiment proxy and find that high sentiment stocks earns high average returns compared to low sentiment stocks between 2004 and 2008. [Kling and Gao \(2008\)](#), however, document that a survey-based

institutional investor sentiment measure offers no return predictability at the daily frequency. [Chen et al. \(2014\)](#) construct a market-based Chinese sentiment index and find their sentiment index has good forecasting power for the market movements. However, their sentiment index contains both rational and irrational components, which makes it difficult to interpret the sources of the forecasting power. Alongside these prior works, our article contributes to the evolving literature on sentiment in a number of distinctive ways:

Firstly, we motivate a thorough discussion on the possible mechanisms through which the sentiment of investors might impact on subsequent returns. In particular, we emphasize the competing hypotheses regarding the short-run return predictability of sentiment. For emerging markets with pervasive irrational speculation (eg. positive feedback trading) and binding arbitrage constraints (eg. stronger short-sales constraints and less openness to international investors), the level of investor sentiment is very likely to forecast positively stock returns in the short run, while still predicts negatively the stock returns over longer horizons. Obviously, our perceived view challenges the conventional view that investor sentiment is a contrarian predictor at nearly *all* horizons, which is almost drawn exclusively from the evidence in the US and other developed markets.

Second, we piece together a number of imperfect sentiment proxies and extract their common component to form a real-time market-based sentiment index, which spans a period of 17 years (1997-2013). Obviously, the much longer and more recent sample period guarantees that our analysis on the possible effect on China's stock market is the most comprehensive and the most up-to-date. For example, [Chi et al. \(2012\)](#) use a sample which covers only from 2004 to 2008. The relatively larger sample size, compared to the prior studies, will undoubtedly increase the statistical power for the time-series regressions employed in the article.

Third, we properly address the [Stambaugh \(1999\)](#) estimation bias inherent in the predictive regression with a persistent predicting variable, while this important econometric issue has often been overlooked in prior works such as in [Chi et al. \(2012\)](#). Our simulation evidence reveals that the OLS slope coefficient on sentiment tends to be downward biased towards the (spurious) finding of contrarian predictability. Therefore, we employ a number of refined econometric techniques, including bias-reduction in coefficients and wild bootstrap for test statistics, to provide compelling evidence that investor sentiment is indeed a reliable momentum predictor for

subsequent monthly market returns in China, confirming the momentum hypothesis. A one-standard-deviation increase in sentiment translates into a 1.50% rise in equity premium per month, an economically significant result for all market participants. The strong positive return predictability is robust under both single- and multi-regressor settings, and is statistically and economically significant both in and out of sample.

Fourthly, the strong out-of-sample predictability at monthly frequency renders us to construct simple trading strategies based on the forecasting power of sentiment. In addition to the increase of the portfolio performance from a mean-variance perspective, our numerical example also indicates that the investor would reap much higher profits if she follows the trading signals generated by sentiment rather than a naïve buy-and-hold strategy.

Fifthly, we investigate the term structure of the return predictability of investor sentiment in China. Evidence from multi-horizon predictive regressions yields a striking term pattern: Investor sentiment converts from a momentum predictor in the short run (eg. within the first year) to a contrarian predictor in the subsequent, longer horizons (from 2 to 5 years). The fact that sentiment forecasts negatively market returns in the long run confirms the well-known contrarian predictability of sentiment, as sentiment mean-reverts over an unknowingly long horizon. We also examine the potential spillover effect of global sentiment on China's stock market. In contrast to local sentiment, we find global sentiment does not have an immediate impact, but predicts returns negatively over much longer horizons.

Last but not least, we perform a batch of cross-sectional analyses. Using size-sorted portfolios as proxies for stocks with different sentiment sensitivity, we find that local investor sentiment is more of a size story ([Lee et al. 1991](#)): the short-term momentum pattern of sentiment matters more of small-cap stocks than for large-cap ones. In contrast, there seems less cross-sectional difference for the impact of global sentiment, as it uniformly predicts negatively long-run stock returns across the size groups.

The structure of the paper is as follows. Section 2 first introduces the Chinese stock market and then reviews the relevant literature and puts forth the competing testable hypotheses. Section 3 describes the data sources, the construction of sentiment index, and descriptive statistics. Section

4 provides the main results of the return predictability over various horizons. Section 5 performs cross-sectional analyses. Section 6 discusses the implications of our findings and concludes.

2. Literature Review

2.1. The Introduction of the Chinese Stock Market

There are two major security exchanges in mainland China, the Shanghai Stock Exchange (SHSZ) and the Shenzhen Stock Exchange (SZSE). The two exchanges have no functional difference, except that SHSZ is larger than SZSE in terms of market capitalization. At the end of 2013, both exchanges are ranked among the top 12 stock exchanges in the world based on the total value of market capitalization (see **Table A.2** in the appendix). The combined market capitalization of SHSZ and SZSE is equivalent to 42% of China's GDP in 2013. For historical reasons, common shares in the two exchanges are classified as A-shares and B-shares, which are denominated in local currency and foreign currencies (USD or Hong Kong dollar), respectively. As A-shares comprise the lion's share (96%) of the market, we focus exclusively on the A-share market for our empirical analyses.

Several distinctive features regarding the Chinese A-share market are worth mentioning:

First and foremost, the Chinese market is well known for its speculative nature with a huge amount of young and inexperienced retail investors. According to the 2013 annual report of China Securities Depository and Clearing Corporation, there are more than 53 million valid individual investor accounts in SHSE and SZSE, among which 44% of the account holders are less than 40 years old. Less than 20% of the retail investors have an education background of bachelor degree or above. A breakdown on the stock holdings and trading activities in SHSE (**Panel B of Table A.2**) provides more insights: Individual investors hold directly more than 21% of the total market capitalization of the stocks in SHSZ. In comparison, stock holdings by professional institutions—including investment funds, pension funds, security companies, insurance companies, asset under management (AUM) firms, and qualified foreign institutional investors (QFII)—is less than 15% as of 2013. More strikingly, trading activities by individual investors account for 82.24% of the total trading volume in 2013.² It is well known in the

² Unfortunately, we do not have the breakdown data for SZSE. However, there are no reasons to believe that individual investors play a lesser role in SZSE as compared to SHSE. In fact, the percentage of stocking holding and the composition of trading volume by individual investors is more likely to be higher for SZSE than SHSE, as individual investors tend to hold small and speculative stocks, which are mostly listed in SZSE. This is also supported by the much higher turnover ratio in SZSE than in SHSE (see the last column of Panel A of Table 1).

financial literature that retail investors are highly influenced by sentiment. They hold less diversified portfolio, have more incentives to trade speculative stocks, and engage in unsophisticated trading strategies such as trend following or correlated trading ([Feng & Seasholes 2004](#); [Kumar & Lee 2006](#)).

Secondly, as one of the leading emerging markets, the Chinese stock market is extremely liquid and has unprecedented trading volume. As is shown in the last column of **Panel A of Table A.2**, on average, stocks in SHSE and SZSE have turned over at least 1.49 and 2.65 times in 2013, which is even higher than the average turnover ratio for many of the developed markets.

Thirdly, the Chinese stock market is characterized by heavy regulation: Short-sales of stocks are prohibited by law. The stringent constraints on short selling make it very difficult to arbitrage away the mispricing at the market level as well as the stock level ([Mei et al. 2009](#)).

Fourthly, the market is featured with less openness to international investors. It only opens its door to foreign investors in November 2002. Only approved qualified foreign institutional investors (QFII) can invest in the Chinese A-share market and they are subject to a number of rules and regulations (eg. capital control) imposed by the local regulatory body.

Overall, given their large share holdings and frequent trading volume, retail investors (noise traders) have very strong market power in China's stock market. These unique investor features, which is distinct from developed markets, seems to weigh more on the sentimental demand from the noise traders. It provides an interesting "out-of-sample" test for the sentiment effect on stock returns under alternative settings.

2.2. Relevant Literature

In recent years, a prevailing view in finance is that investor sentiment, defined as the irrational belief formed by a subset of uninformed investors (eg. noise traders), is a contrarian indicator of future stock returns at nearly *all* horizons ([Fisher & Statman 2000](#); [Baker & Wurgler 2007](#)):

The contrarian hypothesis: the net effect of investor sentiment on subsequent returns is negative.

The contrarian view— high (low) level of sentiment is linked with low (high) market returns in the subsequent periods—is rationalized by the theoretical work of [Barberis et al. \(1998\)](#), in which the representative, sentiment-prone agent extrapolates the current trend far into the future (the representative bias), makes erroneous bets on future market movements by pushing prices far

away from the fundamental values, and eventually faces adverse consequences as “fads” do not live up to the facts (price revision).

A similar view is noted in [Warther \(1995\)](#). He formulates the “price-pressure” hypothesis that unwarranted sentimental demand (via aggregated mutual fund flows), which reflects no real information, exerts contemporaneous price pressure and is linked with lower expected returns in subsequent periods (price reversals). [De Long et al. \(1990a\)](#) argue that the sentiment-induced mispricing provides profitable opportunities for arbitrageurs to buck the trend, the so-called “arbitrage effect”, which also leads to a contrarian effect of investor sentiment on subsequent returns. Therefore, a negative relation between investor sentiment and subsequent market returns can be broadly perceived as indirect evidence that 1) noise traders systematically bet wrongly on the future direction of market movements, the “dumb money effect”, and 2) rational speculators engaging price stabilizing activities (eg. arbitrage) to bring the market prices back to fundamental levels, the “smart money effect”.

Empirically, ample evidence on long-run time-series predictability is provided from the US market ([Baker & Stein 2004](#); [Brown & Cliff 2005](#)) and other developed markets ([Schmeling 2009](#)) to reassure the contrarian view. Moreover, the contrarian predictability of sentiment is more pronounced in the cross section ([Baker & Wurgler 2006](#); [Frazzini & Lamont 2008](#); [Baker et al. 2012](#)).

Despite its wide popularity, powerful tests in the above sentiment-induced return reversal pattern are, however, inherently difficult as there are no firm guidelines as to the appropriate time frame to be used in these tests ([Warther 1995](#)). At one extreme, researchers look for the forecasting power of sentiment at really short horizons (eg. weekly or monthly) as in [Brown and Cliff \(2004\)](#). At the other extreme, return predictability of sentiment from one up to five year horizons are also heavily investigated by empiricists ([Brown & Cliff 2005](#); [Schmeling 2009](#)).

The uncertainty about the appropriate time frame to test the relation between investor sentiment and future market returns, at the very least, points to the fact that the documented long-run contrarian predictability of investor sentiment does *not* necessarily carry over to the short horizons, such as the weekly or monthly frequency. On one hand, high levels of investor sentiment, which is unjustifiable by the facts, can pose a warning sign for stock returns in the

near-term (*a negative relation between sentiment and subsequent returns*). On the other hand, the sentiment of investors is also likely to exert a persistent, positive effect on subsequent returns in the short run (*a positive relation between sentiment and subsequent returns*). A number of behavioral reasons or real-world mechanisms could contribute to such a positive mechanism between investor sentiment and future returns in the short run. Though the literature offers no fully specified model on the linkage between sentiment and short-term returns, a number of works give hints on the possible directions in which sentiment might impact on future returns in the short run, which we summarize in the next subsection of testable hypotheses.

2.3. The Momentum Hypothesis

The contrarian predictability of sentiment builds heavily on the efficiency of financial markets. That is, rational speculators (informed investors) stabilize the asset prices, bucking the trend induced by sentiment. It should be noted, however, that a number of real world mechanisms or frictions, which prevail in the financial markets, limit the arbitrage ability of rational speculators. As a result, mispricing could prevail for extended periods of time, and the level of sentiment could also serve as a momentum predictor for subsequent returns in the short run. This should not be surprising as [Kling and Gao \(2008\)](#) find that the investor sentiment follows a positive feedback process in the short run in China. Adding to this argument is the empirically observed yearly-long self-feeding price bubbles in the history of Chinese stock market (eg. the 2007 price bubble). We provide below a number of reasons for the alternative, momentum predictability of investor sentiment.

First, arbitrage—the activity to bring mispriced assets back to their fundamental level—is limited in practice ([Shleifer & Summers 1990](#)). Unlike the theoretical assumption of infinite horizon, rational speculators, such as institutional investors, typically have shorter investment horizons than noise traders (retail investors) in practice. The shorter horizon makes them concerned about the resale prices, in particular. The uncertainty due to noise trader risk suggests that betting against noise traders run the risk, at least in the short-run, could incur huge financial losses if investor sentiment turns more extreme, and prices move further away from fundamental levels, the well known “create-space effect” ([De Long et al. 1990a](#)). Therefore, the concern about the resale price limits rational speculators’ ability to take on price stabilization activities (arbitrage) in the short run. Adding to the nerve of rational institutional investors are the fundamental risk of

the underlying ([De Long et al. 1990a](#)), the regular performance evaluation with their peers, the short-sales constraints (or the barrier of short-sales costs). All of these factors could contribute to the possibility that noise traders induced mispricing gets more pronounced over a protracted period of time, leading to a positive relation between investor sentiment and subsequent returns in the short turn.

Second, [Shleifer and Summers \(1990\)](#) warn that one of the strongest behavioral tendencies is to extrapolate or chase the trend. Investment is a social process: sentiment investors are reinforced by each other hopping on the bandwagon (the so called “bandwagon effect”). Put differently, noise traders *herd* in and out of the market to engage in the trend-following strategy, causing prices to deviate further from their fundamental level. More dramatically, the “bandwagon effect” could persist for extended periods of time, as sentiment gets higher and higher over time and sequential (new) investors flooded into the market, which leads to the phenomenal “deposit drawdown effect” in China as documented in [Burdekin and Redfern \(2009\)](#). Using a sequential decision model, [Banerjee \(1992\)](#) shows that it makes sense for subsequent investors to underweight their own information and overweight the information of existing market participants, thereby leading to the herd behavior. The herd behavior also implies a broad wave of trend chasing activities which widens the sentiment-induced mispricing in subsequent periods. Therefore, the trend-chasing behavior postulates a positive effect of sentiment on future returns.

A third reason is that even the rational speculators jump onto to the bandwagon and engage in price destabilizing activities, given the prevalence of the positive feedback traders ([De Long et al. 1990b](#)). In anticipation of future buying power of positive feedback traders as assumed in [De Long et al. \(1990b\)](#)’s model, rational speculators buys ahead of others to reap the profits in subsequent selling against positive feedback traders. The essence of the model is that informed rational speculators exploit the systematic bias of others who extrapolate trends far into the future by purchasing ahead of these positive feedback traders, pumping up the asset price, and triggering anticipated sentimental demand at many horizons. As long as positive feedback traders focus on a sufficient long investment horizons (based on their erroneous extrapolation), a series of positively correlated realized returns are expected in conjunction with their sentiment. Therefore, the implication of the model is consistent with a positive effect of sentiment on future returns.

Overall, the possibility that misperception about future market movements gets more pronounced over a protracted period of time suggests a momentum predictability of investor sentiment for subsequent returns, at least in the short-term. In other words, *high* level of sentiment leads to *high* stock returns in subsequent periods.

The momentum hypothesis: the net effect of investor sentiment on subsequent returns is positive.

3. Data and Empirical Design

3.1. Sample Periods and Market Return Proxies

The paper employs a comprehensive data set of the Chinese A-shares, which is carefully constructed from several reliable sources (see **Table 1**). The sample period spans 17 years from 1997 to 2013. To properly characterize the overall performance of the market, we follow [Fama and French \(1993, 2012\)](#) and [Frömmel and Han \(2014\)](#) by calculating the total return (including dividends) of the market portfolio as the value-weighted average of all available individual stocks, which is free of survival bias. Following the convention, we use the monthly rate of the one-year bank time-deposit in China as the risk-free rate. The excess market return is then the difference between the return of the market portfolio and the risk-free rate. Our proxy for the market excess returns covers both capital gains and dividend income, which is better than the Shanghai and Shenzhen Composite Indices that are pure price indices.

[Insert Table 1 here]

3.2. The Investor Sentiment Index

To properly measure the aggregated market sentiment in China, we start by searching for a number of reliable individual sentiment proxies: Namely, market turnover ratio (*TURN*), number of newly opened individual investor accounts in Shanghai Stock Exchange (*IIA*), and value-weighted price-earning ratio (*PE*) of the market. We are aware that there are no definitive or uncontroversial measures for market sentiment. In many cases, the sentiment proxies are country-

specific and are subject to data availability.³ The inclusion of these individual sentiment measures for our study is mainly based on sound economic reasons, data consistency, and data availability. For example, market-wide turnover is considered as a direct measure of investor sentiment according to the theoretical justification in [Baker and Stein \(2004\)](#). High market turnover ratio indicates high demand from the sentiment investors who push the rational investors out of the market, destabilizing asset prices. The number of new investor accounts, is also highly reflective of investor sentiment in China ([Chen et al. 2014](#)): Given the strong market power of retail investors in China, an increasing number of new accounts implies a larger (irrational) demand from retail investors. The inclusion of the price-earnings ratio follows the same logic of [Indro \(2004\)](#). As we do not have mutual fund flow data to estimate increased money flowing to the market, we use the valuation ratio as a direct proxy for the flow data. The logic is quite straightforward, as long as (more) money is pouring into the market, valuation ratio will power higher. Historical episodes of dramatic price bubbles in China's stock markets are uniformly linked with unreasonably high valuation (PE) events. In principle, all the three individual measures are good candidates for investor sentiment as they have greater values in bull markets and lower values in bear markets, reflecting a positive correlation with the investor mood. Moreover, these sentiment proxies are available for the entire sample period and have no artificial breaks in the time series.

To properly form a valid market-based sentiment index in China, we then follow carefully a number of statistical procedures as summarized below.

First, as noted in [Baker and Wurgler \(2007\)](#), individual sentiment proxies may have a deterministic trend, which is not related to investor sentiment. For example, both turnover ratio and the number of newly opened individual investor accounts have an upward trend in our sample period, which reflects the unprecedented growth of China's capital market and the increased market participation by retail investors over time. Therefore, we follow the detrend procedure as suggested in [Baker and Wurgler \(2007\)](#) and [Chen et al. \(2014\)](#) to scale individual

³ For example, [Baker and Wurgler \(2006, 2007\)](#) construct a US sentiment index based on six individual components: the close-end fund discount, the NYSE share turnover, the number of IPOs, the average first day return of IPOs, the equity share in new issues, and the dividend premium. However, in a later international study, [Baker et al. \(2012\)](#) admit data limitations and could only construct the country-specific sentiment index with four components: the volatility premium, the volume of IPOs, the average first day return of IPOs, and market turnover. Other scholars advocate the use of alternative measures for sentiment such as technical indicators and/or survey-based measures ([Brown & Cliff 2004, 2005](#); [Lemmon & Portniaguina 2006](#); [Schmeling 2007, 2009](#)).

sentiment proxies by their prior 6-month moving average. The detrend procedure removes the (undesirable) deterministic trend and also helps to ensure the stationarity of the time series.

Second, as is noted in prior literature that shifts in these sentiment proxies has both a rational component and an irrational one. The rational component relates to the changes in macro fundamental, as sentiment investors receives a noisy signal about the overall economy ([Verma & Soydemir 2009](#)). To remove, or at least alleviate, the “rational” effect, we follow [Baker and Wurgler \(2006\)](#) and [Verma and Soydemir \(2009\)](#)’s orthogonalization procedure by regressing each individual proxy on a set of macro variables (namely the growth of industrial production, growth of money supply, levels of short-term interest rates and foreign exchange rates). The residuals of the orthogonalization procedure are then the proxies for the irrational part of investor sentiment.

Third, we adopt the principal component approach (PCA) of [Baker and Wurgler \(2007\)](#) by extracting the first principal component from the three individual sentiment measures (the residuals) as the market-based investor sentiment index.⁴ In principle, the first principal component, which captures the common variation of the individual measures, should represent properly the time-variation of (unobserved) market-wide investor sentiment in China.

The principal component procedure leads to the following sentiment index, denoted as S^{PCA} .

$$S_t^{PCA} = 0.570TURN_t^O + 0.642IIA_t^O + 0.510PE_t^O \quad [3.1]$$

The first component explains 67% of the sample variance, well capturing the co-movement of individual sentiment proxies. We use the subscript O to denote that all the three components are orthogonalized to the macroeconomic variations, capturing the irrational sentiment effects. All the three proxies have the expected sign, confirming the positive relation with market sentiment. The PCA sentiment index places the most importance on IIA^O as its coefficient is the largest in magnitude, reinforcing that the number of newly opened accounts gauges the irrationally optimism among retail investors. For comparison purpose, we also construct a naïve sentiment index, denoted as S^{Naive} , which imposes equal weights on the three sentiment composites, and another PCA sentiment index which excludes the less commonly used component PE^O , denoted

⁴ The PCA proceeds as follows: First, we standardize the three sentiment proxies and obtain the eigenvalue and eigenvector of their covariance matrix. Second, the composite sentiment index is constructed as a linear combination of the three variables using the eigenvector associated with the largest eigenvalues as the corresponding weight.

as S^{ExPE} . In fact, these two alternative sentiment indices are highly correlated with the original PCA sentiment index with correlation coefficients of 98% and 94%, respectively.

Figure 1 plots the fluctuation of the PCA sentiment index together with the market excess returns in our sample periods. The most salient feature is that the sentiment index tracks closely the direction of market returns, implying a high correlation between the innovations of the two series. Moreover, a (preliminary) eye-ball test indicates the appropriateness of our sentiment index, as the peaks and troughs of the index are in concurrent with the major episodes in China's stock market history. For example, the spike in June 1999 is probably a response to the reduction of the stamp tax of security transaction within that month, while the spikes in May 2006 and January 2007 are in concurrent with the then-record daily transaction value of 55.3 billion and 100 billion Chinese yuan, respectively. The largest price crash (the post 2007 period) is also consistent with the consecutive pessimistic sentiment wave from October 2007 to October 2008.

3.3. Descriptive Statistics

Table 2 provides the summary statistics of the dataset. The arithmetic mean and standard deviation of the monthly market excess return are 0.16% and 8.07% respectively, indicating a monthly Sharpe ratio of 0.02 and a market price risk (defined as the ratio of market excess return to its variance) at around 0.25. One important note is that the relatively low market price of risk, which reflects the implied risk aversion of an average investor, suggests that Chinese investors are generally less risk averse as compared to investors in developed markets.⁵ The high return volatility, low Sharpe ratio, and lower degree of risk aversion are, however, in line with the general perception of a speculative market as described in **section 2.1**. The sentiment indices, the three individual sentiment components, and other economic variables are all standardized with mean zero and unit variance, which will facilitate their comparison of economic significance in later predictive regression. Consistent with the literature, all the predicting variables seem quite persistent, indicated by their first-order autocorrelation coefficient. In contrast, the market excess return exhibits no serial correlation effect, as its first-order autocorrelation coefficient is virtually zero. The time-series plots of these economic and financial variables are available in **Figure A.1** in the appendix.

⁵ The historical market price of risk in the United States is between 1.5 and 4 in [Bodie et al. \(2012\)](#).

The lower part of **Table 2** presents the cross-correlation coefficients of the dataset. Several prominent features are worth noting. First, the market excess return is highly correlated with the PCA sentiment index and its three components. The positive contemporaneous correlation is consistent with **Figure 1** that increased sentiment level will push up asset price and lead to increased return during the current period (the “price pressure”). It also reinforced the need to adopt bias-reduction technique in our later estimation because the innovations of both sentiment and market return are highly correlated. Second, the three sentiment components are highly correlated with each other, reinforcing their validity as the proxies for market sentiment. The orthogonalized *PE* ratio has only a correlation coefficient of 0.26 with the unorthogonalized *PE* ratio, much less than its correlation with *IIA* sentiment proxy, which lends further support for its validity as a proxy for the (irrational) sentiment measure. Third, the consumer confidence index (CCI), a survey-based sentiment measure, is not correlated with contemporaneous returns. Nor does it directly correlated with all the market-based sentiment proxies as the coefficients are slightly negative, which further challenges its validity as a sentiment proxy in China. While some studies tend to use CCI as a proxy for investor sentiment ([Schmeling 2009](#)), we should point out that CCI concerns consumers’ general expectations about the overall prospects of the economy rather than the stock market. This is indeed the case as CCI is positively correlated with other economic variables (such as BCI, IP, and PE ratio). We have more to say about CCI in our later analysis.

[Insert Table 2 here]

4. Time-series Analysis

4.1. Short-term Return Predictability

4.1.1. Single-Factor Predictive Regression

The predictive regression with single predictor variable is by far the most heavily studied and commonly used in the literature. Hence, it forms the baseline model in establishing the relation between sentiment and market returns.

$$R_{t+1} = \alpha + \beta S_t + u_{t+1} \quad [4.1]$$

where R_{t+1} is the market return in excess of risk-free rate in period $t + 1$, S_t is the (lagged) sentiment proxy (or other predicting variable) measured at the end of period t . We test the null hypothesis ($H_0: \beta = 0$) that investor sentiment has no predictive ability by evaluating the slope coefficient.

As our interest is to perform hypothesis testing on the slope coefficient in the predictive model, several econometric issues may well arise if we use the conventional ordinary least square (OLS) estimation procedures: First, there is potentially a spurious regression concern when the predictor is highly persistent ([Ferson et al. 2003](#)). Second, OLS method produces biased coefficient in finite samples, the well-known [Stambaugh \(1999\)](#) small sample bias, when the regressor is (highly) persistent and its innovations are correlated with those of the dependent variable. Note also that in our case the beta coefficient is downward biased (towards zero), as the innovations of the two time series are positively correlated. The magnitude of the potential downward bias in OLS estimate is illustrated by the simulation evidence in **Appendix A**.

To properly tackle the above mentioned econometric issues, we apply the multiple Augmented Regression Method (mARM hereafter) of [Amihud et al. \(2009\)](#) to obtain the bias-reduced estimates of the slope coefficient on sentiment. We are aware that there are a number of alternative bias-reduction techniques proposed in the literature, such as the univariate Augmented Regression Method (uARM hereafter) of [Amihud and Hurvich \(2004\)](#) and the Moving-block Jackknife Method (MBJK hereafter) of [Zhu \(2013\)](#). Our choice of the mARM method is based on the good finite sample performance as indicated in the simulation evidence in **Appendix A**.

The mARM method in [Amihud et al. \(2009\)](#) proceeds by running an auxiliary AR(1) regression on the predictor variable, S_t .⁶

$$S_{t+1} = \theta^c + \rho^c S_t + v_{t+1}^c \quad [4.2]$$

where the superscript c denotes that the coefficients are bias-corrected estimates, which are obtained by iterating on the [Nicholls and Pope \(1988\)](#) expression for the analytical bias of the OLS regression for the AR(1) process.⁷ According to [Amihud et al. \(2009\)](#), the bias-reduced

⁶ Throughout the article, we use the term uARM and mARM to refer to the bias-reduction method in [Amihud and Hurvich \(2004\)](#) and [Amihud et al. \(2009\)](#), respectively. It should be noted that both methods are flexible enough to accommodate predictive regressions with single regressor and multiple regressors as well.

⁷ For brevity, we refer readers to [Amihud et al. \(2009\)](#) regarding the technical details on the iterative procedure.

estimator of the slope coefficient β in the predictive regression is achieved by augmenting Eq. [4.1] with the bias-corrected error term v_{t+1}^c , obtained from the auxiliary AR(1) regression in Eq. [4.2]. The statistical significance of the slope coefficient is based on the t -statistics associated with the Newey-West adjusted standard errors.

For robustness of our hypothesis testing, we also report separately the empirical p -values based on the wild bootstrap procedure used in [Huang et al. \(2014\)](#). The wild bootstrap procedure takes into account the persistence in predictors, correlations between the innovations of market returns and those of predictors, and general forms of return distribution (eg. heavy tails).⁸ In unreported analysis, we also perform the t -tests based on the adjusted standard errors according to the analytical expressions suggested in [Amihud and Hurvich \(2004\)](#) and [Amihud et al. \(2009\)](#). Results for the significance of the slope coefficients are very similar, and thus omitted for brevity purpose.

Table 3 presents the estimation results of the predictive regression. **Panel A** reports the baseline results using our PCA sentiment index, S^{PCA} , constructed from the three individual sentiment components. Contrary to the prevailing view that sentiment is a contrarian predictor, we find that the PCA sentiment index is a momentum predictor over the entire sample period: lagged sentiment is positively associated with the excess market return in the subsequent month with an OLS slope coefficient of 1.38%. The positive coefficient is statistically significant at 5% level based on the Newey-West t -statistics as well as the wild bootstrapped t -statistics (see the empirical p -value). When correcting for the finite sample bias using mARM method, the slope coefficient becomes even larger in magnitude than the OLS estimator. The downward bias in OLS estimation is as expected, given that the forecast error in returns and innovations in investor sentiment are positively correlated (see also our simulation evidence in **Appendix A**). The bias adjusted slope coefficient is around 1.50% under [Amihud et al. \(2009\)](#)'s mARM method. That is, a one-standard-deviation rise (drop) in sentiment increases (reduces) expected market returns over the next month by approximately 1.50%. Recall (from **Table 2**) that the sample average of the monthly excess market return in China is only 0.16%, thus the magnitude of the slope coefficient implies that the expected excess market return varies by about nine times larger than its average level, signaling strong economic significance. It seems tempting to extrapolating this

⁸ A walk-through of the wild bootstrap method is detailed in the Appendix of [Huang et al. \(2014\)](#).

return predictability to an annual basis by a multiplication factor of 12, which indicates an annualized value of around 18%. This, however, seems unreasonable because sentiment is mean-reverting and high (low) sentiment level indicates lower return in the long-run, which we will analyze further in **Section 4.2**.

The R^2 of the OLS regression for the entire sample period is around 2.95%. Empirically, this level is comparable with, or slightly higher than, that of conventional return predictors at short horizons ([Goyal & Welch 2008](#)). The typical explanatory power of a return predictor ranges between 0.3% and 5% at the monthly frequency. The R^2 of the mARM regression increases substantially as expected because the model includes the innovations from the auxiliary regression in Eq. [4.2], a contemporaneous term, which significantly increased the explanatory power of the variation in stock returns.

To check whether the predicting power of sentiment is stable over time, we divide the full sample into two subsamples with more or less equivalent length: One spans from 1997 to 2005 and the other spans from 2006 to 2013. The positive return predictability of sentiment remains statistically significant in both cases, indicating a very stable predictive pattern at the monthly frequency. Interestingly, the magnitude of the slope coefficient seems more pronounced in the second subsample (2006-2013) than the first subsample (1997-2005). The enlarged sentiment effect over time is in line with the fact that the second subsample is characterized with more dramatic market shifts (eg. the 2006-2007 price bubble and the 2007-2008 financial crisis).

Panel B reports the estimation results using each of the individual sentiment components as the regressor. The results are highly consistent with that of the PCA sentiment index: The slope coefficients of $TURN^O$, IIA^O , and PE^O all have the expected positive sign and are all statistically significant as demonstrated by the empirical p -values. The magnitude of the bias-reduced slope coefficients ranges between 0.91% and 1.77% on a monthly basis, again signaling strong economic significance. The R^2 s of the OLS predictive regression range from 1.10% to 3.64%. It seems clear that the forecasting power of our constructed PCA sentiment index comes directly from the three underlying components. Although each individual component can be used as a valid sentiment proxy and produces significant output, we weigh more on our constructed PCA sentiment index as the PCA procedure removes the idiosyncratic noises contained in the individual proxies, if any ([Baker & Wurgler 2006, 2007](#)).

For comparison, we report in **Panel C** the estimation results for two alternative sentiment indexes: The S^{ExPE} sentiment index which excludes the component PE^O , and the consumer confidence index (CCI), a survey-based measure. From **Panel B**, it seems that PE^O is a stronger predictor than the other two conventional sentiment proxies. To address the suspicion that our baseline result is mainly driven by one single predictor, we construct the S^{ExPE} index, which includes only $TURN^O$ and IIA^O . Not surprisingly, the S^{ExPE} index has very similar predictive features as the S^{PCA} index, although the slope coefficient and the R^2 of the OLS regression is a bit smaller in magnitude: The bias-adjusted coefficient is 1.12%, which is significant at 5% level, and the R^2 of the OLS regression is 1.70%. The similar predictability from the S^{ExPE} index demonstrates the robustness of our baseline results, as it is not driven by the inclusion of a single sentiment proxy. In unreported analysis, we also find very similar strong forecasting power for the equally-weighted naiver sentiment index, S^{Naive} . Finally, the estimation results for the survey-based sentiment measure, CCI, is disappointing as its slope coefficient is virtually zero under both OLS and mARM methods, indicating no return predictability at all. The inability to forecast returns in combination with its low correlation with all other (market-based) sentiment proxies (as shown in **table 2**) posing challenges on the validity of CCI as a sentiment measure. In fact, CCI might be a poor measure for investor sentiment for a number of reasons. First, by definition, CCI measures consumers' general expectations about the overall prospects of the economy rather than the stock market. Second and more critically, as a survey-based measure, there is always the concern of "what they say is not what they act", meaning that the survey respondents do not really act in the way as they indicated in their survey. Another concern is that CCI might reflect the rational expectation, rather than the irrational belief of sentiment investors. Therefore, all of these facts cast doubt on CCI as a valid sentiment proxy in our application.

[Insert Table 3 here]

4.1.2. Multi-factor Predictive Regression

We note that some criticism might arise that the predictive power of investor sentiment is mainly a manifestation of the shifts in fundamentals driven by business or economic cycles. To alleviate this concern, we first employ several macroeconomic variables, including business cycle indicator (BCI), industrial production (IP), and aggregated (value-weighted) PE ratio, as valid proxies for the economic fundamentals. We then examine whether the return predictability of

sentiment is robust after controlling these fundamental effects with multivariate predictive regressions. Following the convention in the literature, we conduct the following multiple-factor predictive regressions based on the sentiment index S_t^{PCA} and a vector of the control variables, Z_t .

$$R_{t+1} = \alpha + \beta S_t^{PCA} + \psi' Z_t + \varepsilon_{t+1} \quad [4.3]$$

Again, our interest is in the slope coefficient, β , which measures the marginal effect of the (lagged) sentiment level on subsequent market returns, after accounting for the effects of control variable(s).

The first three model specifications in **Table 4** report the bias-adjusted mARM estimation results based on bivariate predictive regression, in which we include the PCA sentiment index and one of the economic variables each at a time. Across the three model specifications, the slope coefficient β on our sentiment index is very robust, as it has the same positive sign and its magnitude is more or less similar in all three cases, ranging from 1.34% to 1.52%. A result highly similar to that in the univariate regression. The uniform results in the bivariate regression confirm that the return predictability of sentiment is not purely due to the shifts in fundamentals (eg. business cycles, industrial expansion, and market valuation effects).

The last row of **Table 4** reports the result for the so-called Kitchen Sink model, in which we include sentiment and the three economic variables all together. The advantage of the Kitchen Sink model is that it could increase explanatory power of the predictive regression. However, the downside is that the heavily parameterized model with potentially highly correlated regressors might suffer from over parameterization problem. Again, we find very similar and consistent pattern: The bias-adjusted slope coefficient on S_t^{PCA} is significantly positive and large in magnitude (1.39%), which is very similar to the case of the single-predictor regression in **Table 3**. The slope coefficients on IP and PE seem also significant with the Newey-West t -statistics. However, when checking the empirical p -values generated from the wild-bootstrap testing procedure, only the slope coefficient on S_t^{PCA} is statistically significant. To sum up, the incremental predictive effect of S_t^{PCA} are very robust, as it remains both statistically and economically significant after controlling for macro fundamentals under all the four model specifications.

[Insert Table 4 here]

4.1.3 Out-of-sample predictability

In this subsection, we address the well-known [Goyal and Welch \(2008\)](#)'s challenge, a legitimate concern that whether the strong monthly return predictability documented in the in-sample analysis will be carried over to the out-of-sample period. In principle, an over-fitted predictive model tend to deliver better in-sample forecasting performance, while such return predictability do not sustain out of sample. Therefore, the out-of-sample test is the most relevant for assessing genuine return predictability in real time as it avoids over-parameterization issue ([Goyal & Welch 2008](#)).⁹ A somewhat related aim of the out-of-sample analysis is also to assess whether bias-reduction techniques such as mARM, uARM, and MBJK are able to provide enhanced forecast performance out of sample.

For the out-of-sample analysis we retain an in-sample period from 1997 till 2006 to parameterize our model, and use the remaining 7-year out-of-sample period to evaluate the forecast performance. In general, there is a trade-off between having sufficient in-sample observations for reliable estimation of model parameters and resuming a relatively long out-of-sample period for forecast evaluation. Given the relatively short history of China's stock market which unfortunately limits the sample size, our choice of the out-of-sample period well balances the above mentioned trade-off. Furthermore, we adopt a recursive estimation scheme for the in-sample model estimation and form the out-of-sample forecast for period $t + 1$ with the following single-factor predictive model.

$$E_t(R_{t+1}) = \hat{\alpha}_t^m + \hat{\beta}_t^m S_t \quad [4.4]$$

where $\hat{\alpha}_t^m$ and $\hat{\beta}_t^m$ are the regression coefficients estimated recursively using data from period 1 till period t . The superscript m denotes the fitted values obtained from a number of alternative estimations methods, such as OLS, mARM, uARM and MBJK. S_t denotes the predictor variable measured at the end of period t . To avoid the forward-looking bias in the original PCA sentiment index which is constructed from the entire sample, the PCA sentiment index used in Eq. [4.4] is

⁹ According to [Huang et al. \(2014\)](#), out-of-sample tests also suffers less from small-sample size distortions such as the [Stambaugh \(1999\)](#) bias.

also estimated recursively each time using real-time information up to period t . Note that the forecasting model can also be augmented to include other predictor variables.¹⁰

To take up the challenge from [Goyal and Welch \(2008\)](#), we compare the out-of-sample forecasts from the predictive model in Eq. [4.4] with the historical average as the benchmark. The historical average forecast corresponds to a constant return model, a nested version of the predictive model in Eq. [4.4] with the slope coefficient set as zero. Intuitively, if the underlying return process is generated by a constant plus the random noise term, the predictive model in Eq. [4.4] produces a noisier forecast than the historical average because it estimates the slope coefficient with zero population values. In other words, we would expect the benchmark model to generate a smaller (or equivalent) *mean squared forecast error* (MSFE) than our predictive model. To properly evaluate this assertion, we calculate the widely used out-of-sample R_{OS}^2 statistic, the [Diebold and Mariano \(1995\)](#) t -statistic, and the [Clark and West \(2007\)](#) MSFE-adjusted statistic.

The R_{OS}^2 statistic measures the proportional reduction in MSFE for the predictive regression forecast relative to historical average benchmark ([Campbell & Thompson 2008](#)). It is computed as follows.

$$R_{OS}^2 = 1 - \frac{\sum_{t=p}^{T-1} (R_{t+1} - E_t(R_{t+1}))^2}{\sum_{t=p}^{T-1} (R_{t+1} - \bar{R}_{t+1})^2} \quad [4.5]$$

where R_{t+1} denotes the realized excess market return in period $t + 1$, $E_t(R_{t+1})$ the forecast from predictive model at the end of period t . The length of the in-sample period is denoted as p , while T is total number of observations. \bar{R}_{t+1} is the historical average (benchmark) at the end of period t calculated as in Eq. [4.6].

$$\bar{R}_{t+1} = \frac{1}{t} \sum_{i=1}^t R_i \quad [4.6]$$

By construction, the R_{OS}^2 statistic lies within the range $(-\infty, 1]$. A positive value of R_{OS}^2 indicates that the forecast $E_t(R_{t+1})$ outperforms the benchmark \bar{R}_{t+1} in terms of MSFE.

¹⁰ The model specification for augmented predictive model will be as follows: $E_t(R_{t+1}) = \hat{\alpha}_t^m + \hat{\beta}_t^m S_t + \hat{\psi}_t^{m'} Z_t$, where Z_t is the vector of other predictor variables.

The Diebold and Mariano (1995) test (DM-test hereafter) evaluates whether two forecasts offer the same forecast accuracy. In this case, we test the null hypothesis that the historical average has a smaller or equivalent MSFE than that of the predictive regression model.

In a similar vein, the MSFE-adjusted statistic due to [Clark and West \(2007\)](#), also evaluates the null hypothesis that forecast error (measured by MSFE) for the historical average is less than or equal to that produced by the predictive regression model.

Panel A of **Table 5** reports the out-of-sample forecast performance of sentiment using Eq. [4.4]. In the case of PCA sentiment index, the OLS forecast model generates a positive R_{OS}^2 statistic of 3.45%, indicating a smaller MSFE than the historical average. The values of the R_{OS}^2 become even more positive, indicating more precise forecasting performance (smaller MSFE) under the two alternative, bias-reduction techniques: mARM and uARM methods than their OLS counterparts. In terms of the statistical significance of the outperformance, it is not significant for the DM-test statistic. However, the CW-test statistics are all significant at the 10% level for the four competing forecasts using OLS or bias-reduction techniques (mARM, uARM, and MBJK).

For comparison purpose, we also generate a (recursive) naïve sentiment index, which imposes equal weights on the three composites to form the sentiment index. We find a very similar pattern for the naïve sentiment index, as the predictive model delivers better forecasts than the historical average, revealed by the positive R_{OS}^2 values of 3.54%. The forecasts using mARM and uARM bias-reduction techniques are also better than the OLS counterpart. Similar to the findings in the PCA sentiment index, the CW-test statistics are all significant at the 10% level for the four competing forecasts using OLS or bias-reduction techniques (mARM, uARM, and MBJK).

Panel B of **Table 5** reports the forecast performance of bivariate predictive regressions with one predictor being fixed as the PCA sentiment index, while the other being the BCI, IP, or PE one at a time. The idea to augment the forecast model with additional predictor(s) is to exploit a (potentially) larger price-related information set. The downside, however, is that including non-relevant variable(s) might increase noise in the model estimation and leads to worse forecasting performance. For all the three bivariate forecasting models, we find very similar out-of-sample performance: The R_{OS}^2 values are positive, indicating superiority over the historical benchmark. However, a closer examination reveals that none of the R_{OS}^2 values are more positive than their

counterparts in the univariate case with sentiment index as the sole predictor. The forecasting errors (MSFE) are also larger than their counterparts in the univariate case and the DM-test and CW-test statistics also seem less significant in most situations. The weaker performance is possibly because the economic variables (BCI, IP, and PE) have very little or no predictive power for subsequent market returns and the inclusion of them only adds additional noise to the out-of-sample forecast. In unreported analysis, we also consider the return forecasts from the Kitchen Sink model in Table 4 (the multivariate case), the out-of-sample performance is even worse than the bivariate cases and are thus omitted for brevity purpose.

Summarizing the findings in **Table 5**, we find strong evidence that investor sentiment predicts subsequent market returns out of sample. In many cases, the out-of-sample forecast ability can be further improved by bias-reduction techniques, in particular the mARM methods, which is also consistent with findings in prior subsections.

[Insert Table 5 here]

4.1.4 Portfolio Implications

In this subsection, we provide further insights on the portfolio implications based on the out-of-sample predictability enhanced by investor sentiment. That is, we investigate whether simple asset allocation strategies based on return forecasts generated by the predictive regression add value for a typical mean-variance investor.

Inspired by [DeMiguel *et al.* \(2009\)](#), we consider a simple portfolio strategy by which a mean-variance investor dynamically allocates across the risky asset (market portfolio) and risk-free asset based on the one-period ahead excess return forecast obtained by Eq. [4.4]. At the end of period t , the investor allocates w_t of the portfolio to equities and the remaining $(1 - w_t)$ of the portfolio to the risk-free asset. The equity proportion w_t depends on two factors. First, it is in direct proportion to the (expected) price of risk of the portfolio, calculated as the ratio of forecasted excess return of the portfolio ($E_t(R_{t+1})$) to its associated variance ($Var_t(R_{t+1})$). The expected variance is measured with the five-year rolling window of monthly excess returns. Second, it is inversely related to the risk aversion coefficient, γ , which gauges the degree of risk tolerance of the investor. The calculation of the equity proportion is expressed in the following equation.

$$w_t = \left(\frac{1}{\gamma}\right) \left(\frac{E_t(R_{t+1})}{Var_t(R_{t+1})}\right) \quad [4.7]$$

For realistic concerns, we place two boundary constraints on the portfolio weight on equities: The value of w_t cannot be negative and its maximum is set at 1.5. These boundary restrictions correspond to the short-sales constraint and the leverage constraint, respectively. That is, short sale is not allowed and the maximum leverage ratio is 1.5, which well reflects the real-world conditions in China's stock market. Based on these conditions, the realized excess return of the portfolio R_{t+1}^P at period $t + 1$ is calculated as

$$R_{t+1}^P = w_t R_{t+1} \quad [4.8]$$

where R_{t+1} the realized excess return of the equity market in period $t + 1$. To evaluate the portfolio performance, we first calculate (the out-of-sample) Sharpe ratio, defined as the ratio of the mean of realized excess returns $\hat{\mu}_P$ to its standard deviation $\hat{\sigma}_P$ during the out-of-sample period.

$$\widehat{SR}_P = \frac{\hat{\mu}_P}{\hat{\sigma}_P} \quad [4.9]$$

Second, we compute the certainty equivalent return (CEQ), defined as the risk-free rate that an investor willing to accept rather than adopting a particular risky portfolio strategy, given a certain level of risk aversion γ .

$$\widehat{CEQ}_P = \hat{\mu}_P - \frac{\gamma}{2} \hat{\sigma}_P^2 \quad [4.10]$$

To gauge the economic value of the out-of-sample predictability enhanced by investor sentiment, we then compute the Sharpe ratio difference and the CEQ gain between an investor who adopts the portfolio strategy based on predictive regression in Eq. [4.4] and another investor who uses the benchmark portfolio strategy with the historical average only. To access the statistical significance of the Sharpe ratio difference, we perform the test of [Jobson and Korkie \(1981\)](#) corrected by [Memmle \(2003\)](#) (JKM-test hereafter). The null hypothesis of the JKM-test is that the Sharpe ratio difference of two competing strategies is indistinguishable from zero. Similarly, we evaluate the significance of the CEQ gain by performing the test outlined in [DeMiguel et al. \(2009\)](#) (DGU-test hereafter). The null hypothesis of the DGU-test is that the CEQ gain is indistinguishable from zero between two competing strategies.

Table 6 reports the asset allocation results based on the dynamic portfolio allocation strategy. Recall (from Section 3.3) that the (implied) risk aversion coefficient (γ) of an average investor in China is only around 0.25 given the speculative nature of the market. Therefore, we consider three ad-hoc scenarios in which γ equals 0.25, 0.5, and 1, respectively. These ad-hoc values of γ more or less correspond to an investor who has average, above-average, and extreme risk aversion in China.¹¹ Panel A of the table shows that the PCA sentiment index generates large economic gains for a mean-variance investor, which is consistent with the findings in Table 5. The *CEQ* gains range between 1.95% and 2.02% under the alternative estimation techniques, which are all statistically significant at 5% levels when her risk aversion is 0.25 (the base case). The Sharpe ratio of the testing portfolio is also significantly larger than that of the benchmark portfolio (which uses only historical average as its return forecast) at the 5% or even finer significance level. The *CEQ* gains range between 1.41% and 1.44% at 10% significant level when the risk aversion coefficient is 0.50 (the above-average case). As the investor gets extremely risk averse ($\gamma=1.0$) by penalizing more on the return variability, the *CEQ* gains get smaller in magnitude and less significant. However, the Sharpe ratio remains significant and large in magnitude under all three scenarios. Specifically, the Sharpe ratio difference between the portfolio strategy based on return predictions from Eq. [4.4] and the benchmark portfolio strategy using only historical average is statistically significant at 10% or finer levels, irrespective the value of the risk-averse coefficient. We find very similar patterns of the portfolio allocation results based on the Naïve sentiment index (Panel B of **Table 6**). Both the magnitude and statistical significance of the *CEQ* gains and Sharpe ratios are very similar for the two sentiment indices.

For comparison purpose, we also report the asset allocation results based on the signals generated by the single-factor predictor model using BCI (Panel C) or IP (Panel D) as the predictor variable. We find some evidence that BCI also delivers enhanced asset allocation results compared to the benchmark strategy with historical average only. Specifically, the *CEQ* gains range between 0.73% and 1.46% under alternative estimation methods, and are all statistically significantly under the base case ($\gamma=0.25$). The Sharpe ratios are also significant at 10% or finer levels under the base case. As the investor gets more risk averse, both the *CEQ* gain and Sharpe ratio get smaller and

¹¹ The three scenarios actually correspond to a hypothetical investor, who is willing to invest in 100%, 50%, and 25% of her wealth in the risky portfolio (market portfolio) while keeping the remaining in the risk-free assets.

become less significant, in general. Again, we find no evidence that IP delivers any improved asset allocation results, as the *CEQ* gains are even negative in all cases.

Summarizing the results in **Table 6**, we find consistent evidence that the improved forecast ability with investor sentiment translates into economically relevant gains for an investor who uses the predictive regression, with or without bias-reduction techniques, to time the market. In unreported analysis, we also check the asset allocation implications based on forecasts generated by multivariate predictive regressions (which includes both sentiment and other economic variables) to explore a larger set of price-related information. The forecast results are in general slightly better than the univariate case, which are omitted for brevity purpose.

[Insert Table 6 here]

4.1.5 Market Timing Strategies

While prior subsection provides compelling evidence that the return forecasts from the “sentiment model” can enhance significantly portfolio performance during the out-of-sample period, a legitimate concern is that the portfolio allocation towards the risky asset is (partially) dependent on the risk tolerance of the investor. We, therefore, provide further evidence from a modified market timing strategy, which is independent on risk aversion, to demonstrate the significant economic gain from the predictive power of investor sentiment.

The modified market timing strategy proceeds as follows: At the end of each month, the investor will go long the market portfolio over the next month if the trading signal she receives is a buy signal (*i.e.* above the threshold). Otherwise, she will liquidate the existing position if she holds (goes long) the market portfolio in the prior month. The restriction that the investor can only liquidate her existing position mimics the short-sales constraints imposed on retail investors in China. We apply two different sets of trading signals for the empirical exercise: The first set is based on the buy signal when the lagged value of the sentiment index above zero (the threshold), while the second set is based on the buy signal when the expected market excess return exceeds zero (the threshold). Following the convention, we use the buy-and-hold strategy as the benchmark to evaluate the economic gain of our trading strategy. The out-of-sample evaluation period is from January 2007 to December 2013.

Table 7 reports the return performance of the market timing strategy as well as the benchmark strategy. The benchmark buy-and-hold strategy (1st column) incurs an average loss of 12 basis points per month during the out of sample period. In comparison, the market timing strategy based on the buying signal when lagged sentiment is above zero (2nd column) earns a monthly return of 90 basis points per month before considering transaction cost. In total, such a market timing strategy will incur 23 transactions during the 7-year evaluation period, which is more or less equal to 1 transaction per quarter. After netting off the transaction cost, the market timing strategy will still earn the investor a monthly gain of 63 basis points, which is equivalent to 7.5% per year during the entire evaluation period.

The remaining four columns are based on the buying signal when the forecasted excess return is above zero. Because the trading signal is based on return forecasts rather than the lagged sentiment value, the forecasting technique does play a role here. The market timing strategy differs across the four forecasting techniques (eg. OLS, mARM, uARM, and MBJK). The trading strategy using OLS and MBJK forecasts will provide a monthly gain of 40 basis points per month, while that using mARM and uARM will offer a monthly gain of 50 basis points per month before considering transaction cost. The trading signals generated from the forecasts of mARM and uARM methods will incur 25 transactions, while the OLS and MBJK methods will incur 2 more transactions. Accordingly, the monthly return net of transaction is the highest for the mARM and uARM method, which is 20 basis points per month, equivalent to 2.4% per year. The net returns of the OLS and MBJK method are 7 basis points per month, equivalent to 0.84% per year. Still, the results of the market timing strategy based on the forecasted excess returns are uniformly better than the benchmark buy-and-hold strategy. In all cases, the trading frequency is more or less 1 transaction per quarter, a highly feasible trading frequency for practical implementation.

Summing up the results in **Table 7**, we find consistent evidence that investor sentiment is a strong momentum signal at the monthly frequency and its forecasting power can be translated into significant economic gains when implementing a simple modified market timing strategy based on sentiment-related trading signals.

[Insert Table 7 here]

4.2. Long-term Return Predictability

4.2.1 Long-horizon Predictive Regression

In this subsection, we shift our focus to the long-run forecasting power of investor sentiment. So far, the empirical literature has produced strong evidence that market sentiment is a contrarian return predictor in the long run ([Baker & Stein 2004](#); [Brown & Cliff 2005](#); [Baker et al. 2012](#)). This strand of literature usually adopts the following multi-horizon predictive regression

$$R_{t+1:t+h} = \alpha + \beta S_t + u_{t+1:t+h} \quad [4.11]$$

where $R_{t+1:t+h}$ is the excess return cumulated (or averaged) over the period $t + 1$ till $t + h$, S_t is the h -period lagged predictor (eg. sentiment proxy) measured at the end of period t . The null hypothesis for testing ($H_0: \beta = 0$) is that the predictor has no predictive ability.

For example, [Huang et al. \(2014\)](#) find that investor sentiment predicts *negatively* future market returns from 1 month up till 12 months. They attribute the empirical pattern of long-run sentiment effect to the classical behavioral argument of limit-of-arbitrage. In the long run, market gradually corrects mispricing as it cannot eliminate mispricing in a short period of time. Similarly, [Schmeling \(2009\)](#) provides international evidence across 18 advanced economies that investor sentiment is a contrarian predictor at 1, 6, 12, 24 month horizon.

Despite the growing evidence of long-run contrarian predictability of sentiment in the financial markets, the above estimation method and its associated results, however, are not without substantial debates ([Boudoukh et al. 2008](#)). To overcome the problematic issues of using overlapping returns as the dependent variable and the level of persistence of predictive variable, we opt for an alternative route to investigate the term structure of return predictability for investor sentiment. To be specific, we consider the following predictive regression for return horizons range from 1 month till 60 months.¹²

$$R_{t+h} = \alpha + \beta S_t + \psi' Z_t + u_{t+h} \quad [4.12]$$

where R_{t+h} is the market excess return in the period $t + h$, S_t is the h -period lagged predictor (sentiment proxy) measured at the end of period t . We also include a vector of conventional macro variables to control for the shifts in fundamentals. The additional control variables are

¹² We are highly indebted to an anonymous referee for suggesting this estimation model over the problematic long-run predictive model in Equation [4.11].

business cycle indicator (BCI), industrial production (IP), and aggregated PE ratio similar as in the subsection 4.1.2. Again, we test the null hypothesis ($H_0: \beta = 0$) that investor sentiment has no predictive ability by evaluating the slope coefficient.

Figure 2 plots the t -statistics from the predictive regressions across different time horizons (1 till 60 months), using the PCA sentiment index as the return predictor. For robustness, we report the estimation results under alternative estimation techniques (OLS, mARM, uARM, and MBJK). Investigate carefully the four plots reveals the following patterns. First, market sentiment is a strong momentum signal in the short run (eg. the first year). The t -statistics for the first 8 consecutive months are all positive, with 3 out of them are significant at 5% level. The momentum trend is the most pronounced at the 1st, 2nd, and 7th month lag. Second, immediately following the short-term momentum pattern, t -statistics turn to be negative and become more significant over the subsequent months (i.e. from the 9th month till the 21st month). The change in the sign of the slope coefficients indicates that sentiment reverts to be a contrarian predictor over the subsequent periods. Moreover, starting from the 9th month till the 60th month, we find that the majority of the t -statistics (36 out of 52) are negative, indicating market sentiment predicts negatively market returns in the long run. In particular, the t -statistics at the long end of the term structure (from the 47th month till the 60th month) are all negative, with 6 out of them are significant at 10% level, reinforcing the notion expressed in [Baker and Wurgler \(2007\)](#) that the contrarian predictability of sentiment is linked with market revisions over time.

In unreported robustness checks, we also test the multi-horizon predictive pattern of each of the three sentiment components ($TURN_t^O$, IIA_t^O , and PE_t^O) separately. Again we find very similar pattern that the sentiment proxy shifts from a short-term momentum predictor to a long-term contrarian predictor. In fact, for certain sentiment proxy (such as $TURN_t^O$), the short-term momentum patterns last even longer horizons than 8 consecutive months as for the PCA sentiment index.

To sum up, we find a very striking and robust term structure with different sentiment proxies and under alternative estimation methods. That is, investor sentiment shifts from a momentum predictor mainly within the first year to a contrarian predictor in the relatively longer horizons (approximately from the second year onwards), which is consistent with the short-term momentum hypothesis and the long-term contrarian hypotheses developed in Section 2.

4.2.2 Impact of Global Sentiment

In this subsection, we take a further step by investigating the (possible) contagious effect of global sentiment on China's local stock market. Prior literature suggests that sentiment spreads across markets as country-level sentiments are positively correlated with each other ([Baker et al. 2012](#)). Using a panel regression in which they regress the monthly country-level returns on the beginning-of-year global and local sentiment, [Baker et al. \(2012\)](#) find strong evidence that the global sentiment is a contrarian predictor of country-level returns.¹³ However, the time-series effect of the local sentiment becomes trivial after controlling for the spillover from global sentiment.

It should be noted that emerging markets typically have less openness to the global financial markets than developed markets. For example, only qualified foreign institutional investors (QFII) are able to invest in the Chinese A-share market and they are subjected to capital controls by the local regulator body. The restriction on the market participation of international investors weakens the link between local sentiment and global sentiment. Accordingly, our market-based sentiment index has a slightly negative correlation of -13.5% with the US sentiment index during our sample periods (see **Figure 3**). Therefore, prior evidence on global sentiment from the advanced economies might be misleading for the emerging markets. Whether and how global sentiment spills over to the Chinese local market remains an interesting, unexplored empirical question. To establish the link between China's stock returns and the global sentiment, we estimate the following augmented long-horizon predictive models.

$$R_{t+h} = \alpha + \beta_1 S_t^{Local} + \beta_2 S_t^{Global} + \psi' Z_t + u_{t+h} \quad [4.13]$$

where R_{t+h} is the monthly market excess return over the period $t+h$, S_t^{Local} is the (lagged) local sentiment measured at the end of period t . For our analysis, we use again the PCA sentiment index. For the (lagged) global sentiment index, S_t^{Global} , we use the [Baker and Wurgler \(2007\)](#) US sentiment index used as a valid proxy.¹⁴ Unfortunately, the US sentiment index ends at the end of 2010, which shortens the sample size for our analysis. Both sentiment indices are re-

¹³ In [Baker et al. \(2012\)](#) the global sentiment is formed by extracting the common component from the six country-level sentiment indices using the principle component method.

¹⁴ According to [Baker et al. \(2012\)](#), the US sentiment index contributes the most in the contagion of global sentiment. It should be noted that we do not perform a linear regression to generate a local sentiment index which is orthogonal to global sentiment index, simply because multicollinearity is not a big concern here as China's local sentiment has very little correlation with the US sentiment.

standardized to have zero mean and unit variance during the estimation period to facilitate the interpretation of the economic relevance. Again, we test whether the slope coefficients on these sentiment indices are statistically different from zero.

Figure 4 and **5** plot the t -statistics for the local and global sentiment indexes from the multi-horizon predictive regression using the OLS and mARM estimation methods, respectively. The empirical patterns are very similar under two estimation method. First, the multi-horizon predictive pattern for the local PCA sentiment index remains virtually intact as compared to the pattern in **figure 2**: It serves as a strong momentum predictor in the short run and gradually turns to a contrarian signal in the long run. Second, more interestingly, we find that global sentiment does have a substantial impact on China's stock market. However, the impact from global sentiment is not immediate, as none of the slope coefficients on S_t^{Global} are significant at the 5% level during the first two years. After approximately 2 years, global sentiment does seem to be a warning sign for the local stock market. The dramatic negative impact from the global sentiment is mostly concentrated between three- and four-year horizons. It should also be noted that the inclusion of the global sentiment also increase the explanatory power of the model as revealed by the increased R^2 values compared to their univariate counterparts (unreported). Overall, we confirm the findings in prior literature [Baker et al. \(2012\)](#) that sentiment spreads across markets and global sentiment is a contrarian predictor for country-level returns in the long run.

5. Cross-sectional Analysis

5.1. Size-sorted Portfolios

Although our main focus of the paper is on the aggregated market level, we also shed some light on the sentiment impact in the cross section during our sample period [Lee et al. \(1991\)](#) contend that investor sentiment is more of a size story: small firms are influenced more than large firms, which is later confirmed by other studies ([Baker & Wurgler 2006, 2007](#); [Baker et al. 2012](#); [Stambaugh et al. 2012](#)).

Given the strong market power of retail investors in China, it makes sense to use market capitalization as a valid proxy for sentiment sensitivity, and to test the forecasting power of

investor sentiment in the cross section. Accordingly, we sort stocks into large-, medium-, and small-cap portfolios based on empirical breakpoints of market capitalization at the 30th and 70th percentiles. That is, large-cap stocks represent sentiment immune stocks, while small-cap stocks represent sentiment prone stocks. We should point out that if sentiment does impact more on small firms than large ones, we would expect to find that the short-term momentum signal is more pronounced for small firms as those are the firms subject more to irrational sentiment and are more difficult to arbitrage away these sentiment-induced mispricing.

To establish the cross sectional effect of sentiment, we adopt the same predictive regression framework as in Equation [4.12], except that the dependent variables are replaced by the monthly excess return series of the three size-sorted portfolios and the return of a zero-cost portfolio, in which we go long the small-cap and short the big-cap portfolio. Again, we test the null hypothesis ($H_0: \beta = 0$) that investor sentiment has no predictive ability based on the significance on the slope coefficient.

Figure 5 and **6** reports the t -statistics of the slope coefficient from the predictive regressions across different time horizons (1 till 60 months), using the OLS and mARM estimation methods, respectively. Summarizing the two graphs, we observe a number of empirical patterns. First, it is reassured that (local) investor sentiment is a reliable momentum signal in the short run across the three size-sorted portfolios. Second, the short-term momentum effect seem to matter more for the small sized firms than for the large ones, as there are more significantly positive t -statistics within the first year. Besides, the slope coefficients on the size portfolios are monotonically increasing in magnitude from large-cap to small-cap stocks, while the R^2 values also rise monotonically (untabulated). Third, while market sentiment serves as a contrarian predictor in the long run, the negative impact gets weaker from large-cap to small-cap stocks. In other words, the strong positive time-series predictability is more pronounced and lasts longer horizons for small-cap stocks than large cap ones. Finally and the most strikingly, we find the t -statistics on the slope coefficients for the regression using the zero-cost portfolio returns (as the dependent variable) are mostly positive during the 60-month forecasting horizons. That is, high sentiment implies high size premium over future investment horizons. This striking cross-sectional difference goes counter to the existing findings of contrarian predictability regarding sentiment effect from the developed markets. The bottom line, however, is that we confirm that sentiment is a size story.

That is, it reinforces our conjecture that the momentum signal extracted from market sentiment is more pronounced for small firms than for large firms in China. Considering the strong short-sales constraints in China, it seems reasonable to argue that small firms subject more to irrational sentimental beliefs and it is more difficult to arbitrage away these sentiment-induced mispricing, compared to their large-cap counterparts.

Overall, we can conclude that the sentiment impact is stronger for small-sized stocks than for large-sized stocks, indicating that the price of small stocks exhibits possibly stronger run-up induced by retail investors over relatively long investment horizons.

5.2. Cross-sectional Effect of Global Sentiment

Following the cross-sectional analysis in the prior subsection, we check whether the impact of the global sentiment might also differ in the cross section. We adopt the same predictive regression framework with both local and global sentiment as in Equation [4.13], except that the dependent variables are now replaced by the monthly excess return series of the three size-sorted portfolios and the return of a zero-cost portfolio, in which we go long the small-cap and short the big-cap stocks. The sample period is restricted from July 1997 to December 2010 due to the limits of the data on global sentiment.

Figure 7 and **8** visualize the t -statistics of the slope coefficient on the global sentiment, using the OLS and mARM estimation methods, respectively. Based on the two graphs, a number of empirical patterns can be summarized. First, comparing the plots on the three size sorted portfolios, we find again that the global sentiment does not have an immediate impact, but influence stock returns negatively in the long run across all the size groups. In particular, the long-run contrarian predictability is concentrated between three- and four-year horizons, confirming the global sentiment spillover to the local markets as documented in ([Baker et al. 2012](#)).

Second, we evaluate whether there are any cross-sectional differences by checking the impact of the global sentiment on the zero-cost portfolio which we go long the small-cap stocks and short the large-cap ones. We find some (weak) evidence that the global sentiment acts as a warning sign for the size premium in the long run as the t -statistics are mostly negative from the 13th month onwards. However, none of the t -statistics are significant at the conventional 5% level.

Overall, our finding reinforced the role of global sentiment, which spread across markets and predicts negatively stock returns in the cross section.

6. Discussion and Concluding Remarks

This paper contributes to the growing literature on investor sentiment in the financial markets. Our focus on the Chinese A-share market offers complementary evidence of the sentiment effects from an emerging market perspective. For example, the dominant presence of retail investors, stringent short-sales constraints, heavy capital controls and the less openness to the international investors. All of these features of an emerging market provide an important piece of “out-of-sample” evidence to gauge the forecast power of investor sentiment at the market level and in the cross section.

In contrast to the prevailing view that (local) investor sentiment is a contrarian predictor of market returns at nearly *all* horizons, we find compelling evidence that (local) investor sentiment in China is a reliable momentum signal at the monthly frequency: A one-standard-deviation increase in sentiment translates into a (bias-adjusted) 1.50% rise in market excess returns per month, an economically significant result for all market participants. More strikingly, we document an interesting pattern that investor sentiment shifts from a momentum predictor in the short term (mainly within one year) to a long-run contrarian predictor in subsequent periods. Cross-sectional evidence reveals two things. First, it reinforces the conventional view that investor sentiment is more of a small firm effect. Second, it suggests that the strong positive time-series predictability is more pronounced and lasts longer horizons for small-cap stocks than large cap ones. In comparison, the impact of the global sentiment is less immediate, but does predict returns negatively over longer horizons and in the cross section, confirming that the sentiment spillovers from the global market.

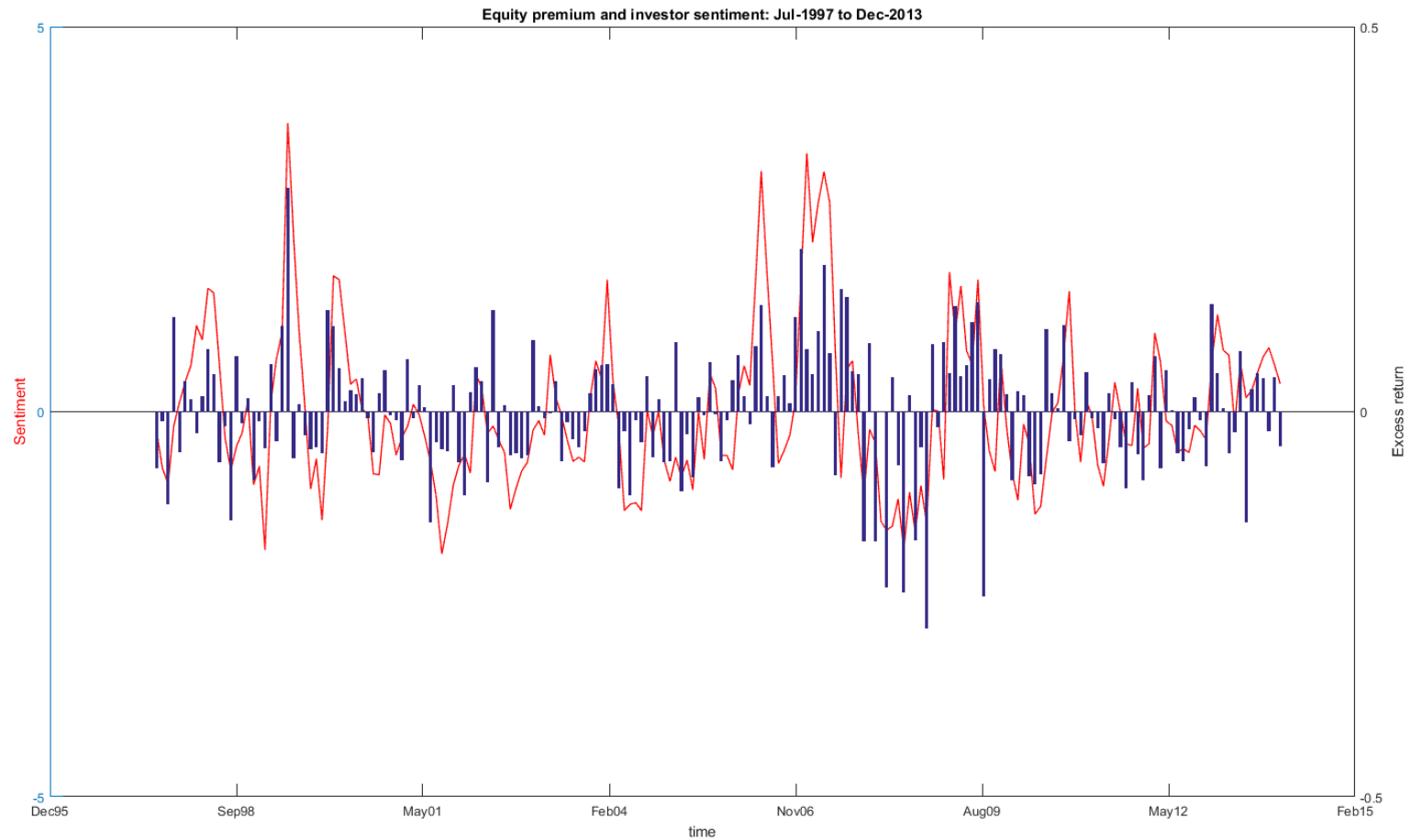
The dramatic shift of investor sentiment from a short-term momentum predictor to a long-run warning sign implies two things. First, (local) investor sentiment gets the future market direction correct on average over the short investor horizon. Second, despite that the systematic biased beliefs of the uninformed investors can persist for an unexpectedly long period of time, it eventually leads to subsequent price reversals shown by the long-run contrarian predictability. In

a broad perspective, the documented term structure of the sentiment effect and its associated pricing impact can help address a number of seemingly anomalous phenomena in China. One such example is the long-lasting price discrepancies between A-shares and B-shares in China ([Mei et al. 2009](#)), or the same Chinese firm which crossed-listed in the Chinese A-share markets and in the Hong Kong Stock Exchange or in NYSE/Nasdaq in the US ([Chan et al. 2008](#)). Another example is the IPO frenzy in China's primary market and its associated extreme IPO underpricing in the secondary market ([Tian 2011](#)). Therefore, the policy implication of our study is to further improve the short-selling mechanisms in China. The short-sales mechanism should be introduced to not only large cap stocks, but, more importantly, to the small cap stocks as well.

Acknowledgements

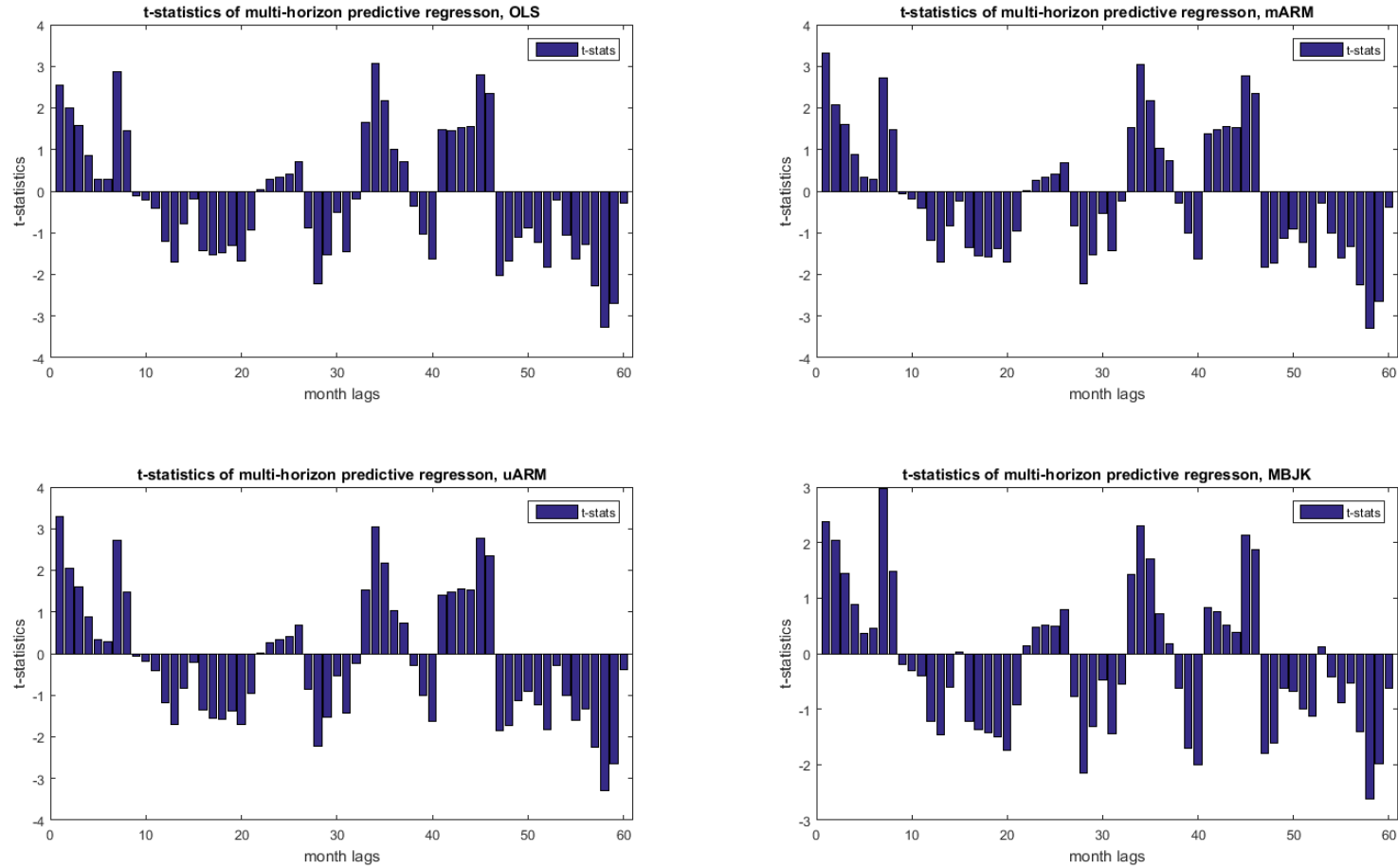
The authors are grateful to Péter Erdős, Gerdie Everaert, Michael Frömmel, Xuezhong He, Kai Li, Alexander Mende, Lei Shi, Joris Wauters and seminar participants at the 2015 annual conference of the European Economics and Finance Society (Brussels), the 2015 joint conference on Institutional Investors and Emerging Market Finance (Ghent), the 2016 annual conference of the Royal Economics Society (Brighton), the 2016 Spring Meeting of Young Economists (Lisbon), the 2016 workshop on Asia and the Pacific Economies (Suzhou), the IÉSEG School of Management (Lille), University of Technology Sydney (Sydney), and ShanghaiTech University (Shanghai) for their detailed comments and suggestions. Han gratefully acknowledges Mengna Zhu and the research department of China International Capital Corporation Limited for sharing (part of) the data and the Fonds Wetenschappelijk Onderzoek (FWO Research Fund – Flanders) for financial support. Li acknowledges the support of National Natural Science Foundation of China (No. 71571197). All remaining errors are ours.

Figure 1. Market Excess Returns and Investor Sentiment of the Entire Sample: 1997 - 2013



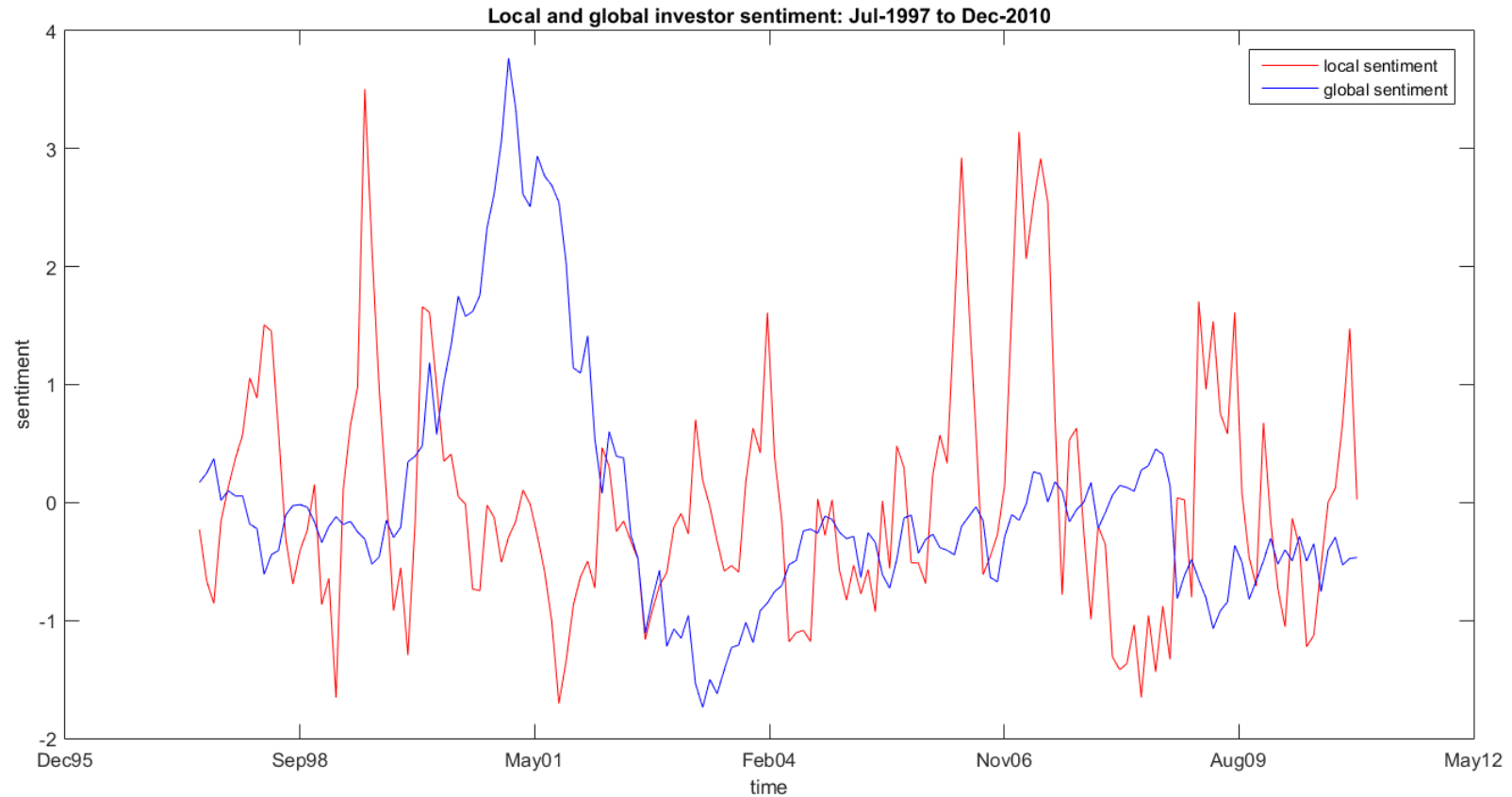
The solid line (the left y-axis) depicts the standardized investor sentiment index obtained from the principal component analysis, the PCA sentiment index. The solid bar (the right y-axis) depicts the (monthly) market excess return (in deciles) over the risk-free rate. The plot covers the entire sample period from July 1997 till December 2013.

Figure 2. Multi-horizon Predictability Pattern under Alternative Estimation Methods.



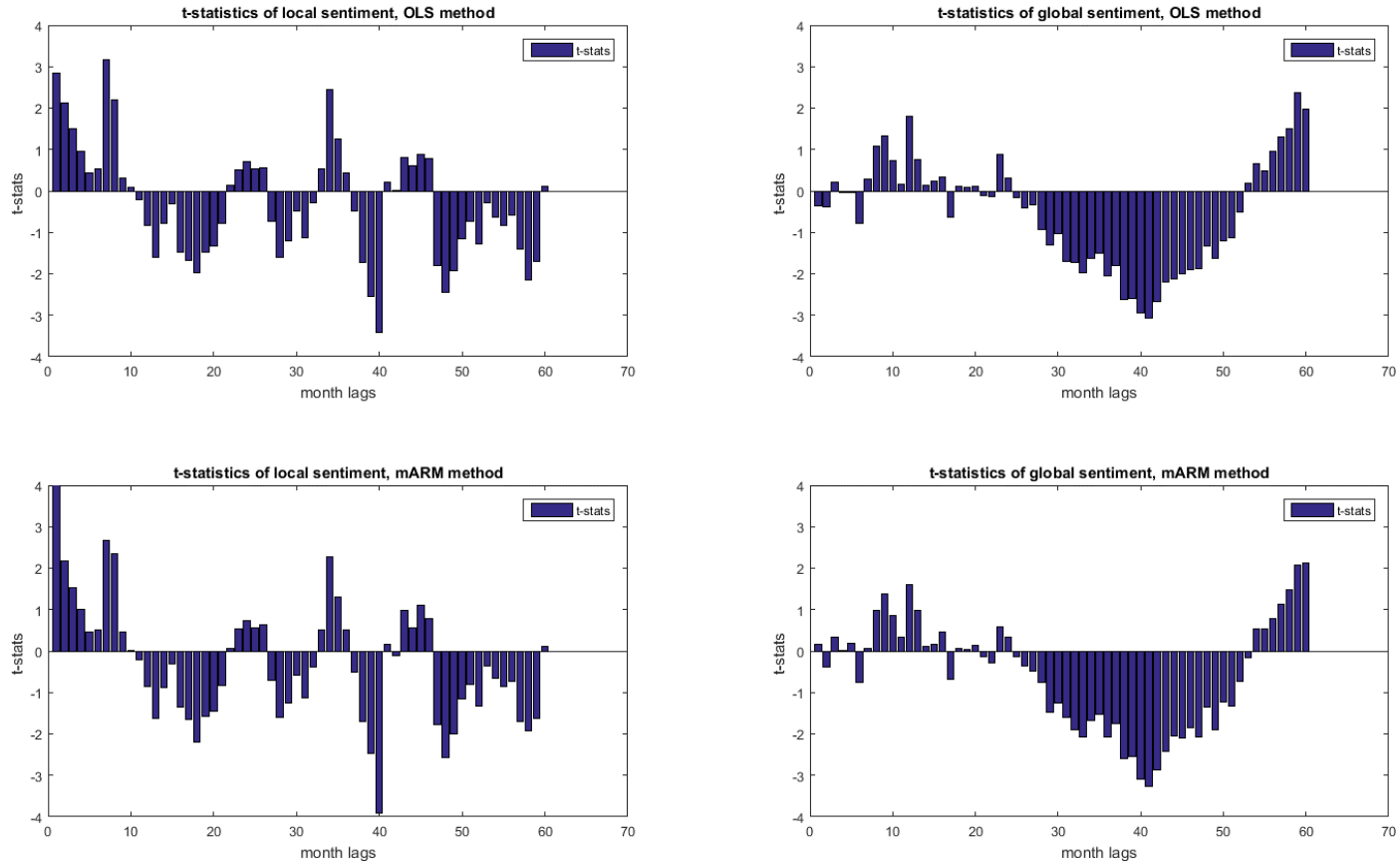
The figure depicts the t -statistics of the slope coefficients in the multi-horizon predictive regression under alternative estimation methods. Panel A to D represent the OLS, mARM, uARM, and MBJK methods, respectively. We regress the monthly excess return of the total market on the lagged sentiment index over various horizons. The horizons (month lags) vary from 1 month up till 60 months. The plot is obtained from the entire sample period from July 1997 till December 2013.

Figure 3. Local and Global Investor Sentiment during the Sample Period: 1997 - 2010



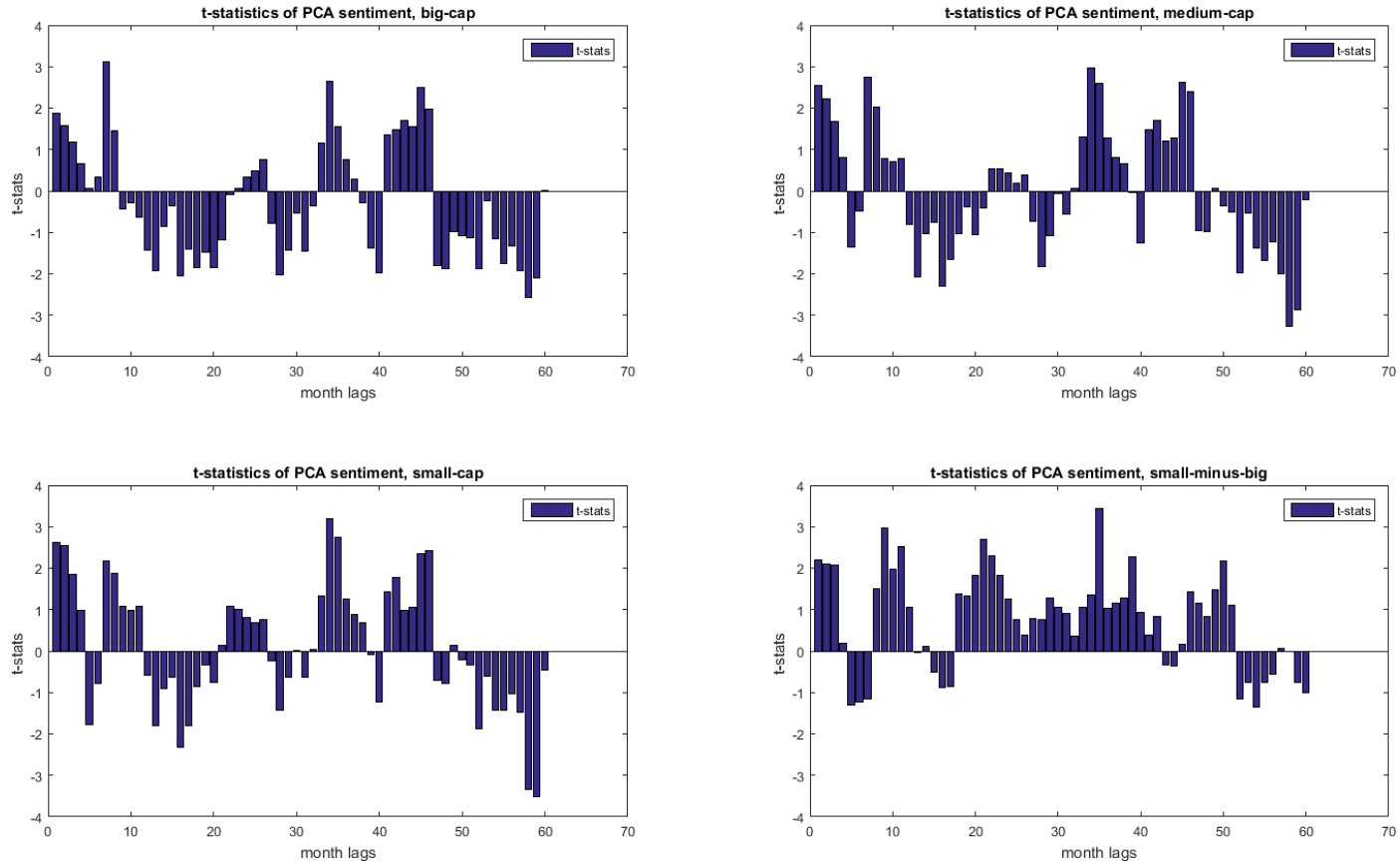
The red line depicts the local investor sentiment, proxied by the PCA sentiment index, while the blue line depicts the global sentiment, proxied by the US sentiment index by [Baker and Wurgler \(2006\)](#). Both time-series are re-standardized to have mean zero and unit variance during the common period from July 1997 till December 2010.

Figure 4. Multi-horizon Predictability Patterns of the Local and Global Sentiment, OLS and mARM method



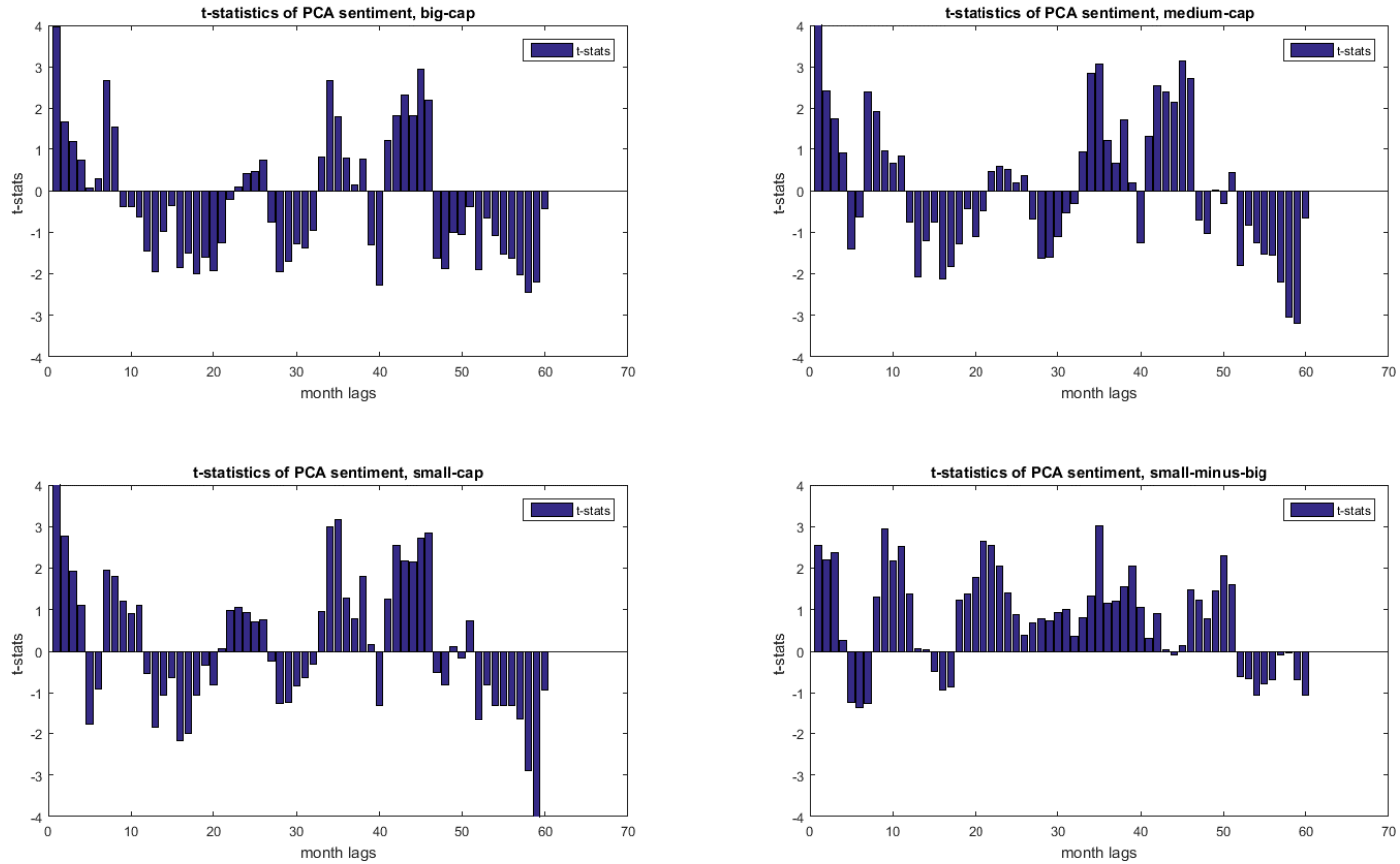
The figure depicts the t -statistics of the slope coefficients on the local and global sentiment indices, respectively. We use both OLS and mARM methods in the multi-horizon predictive regression. The horizons (month lags) vary from 1 month up till 60 months. The plot is obtained from the entire sample period from July 1997 till December 2013.

Figure 5. Multi-horizon Predictability Patterns in the Cross Section, OLS Method



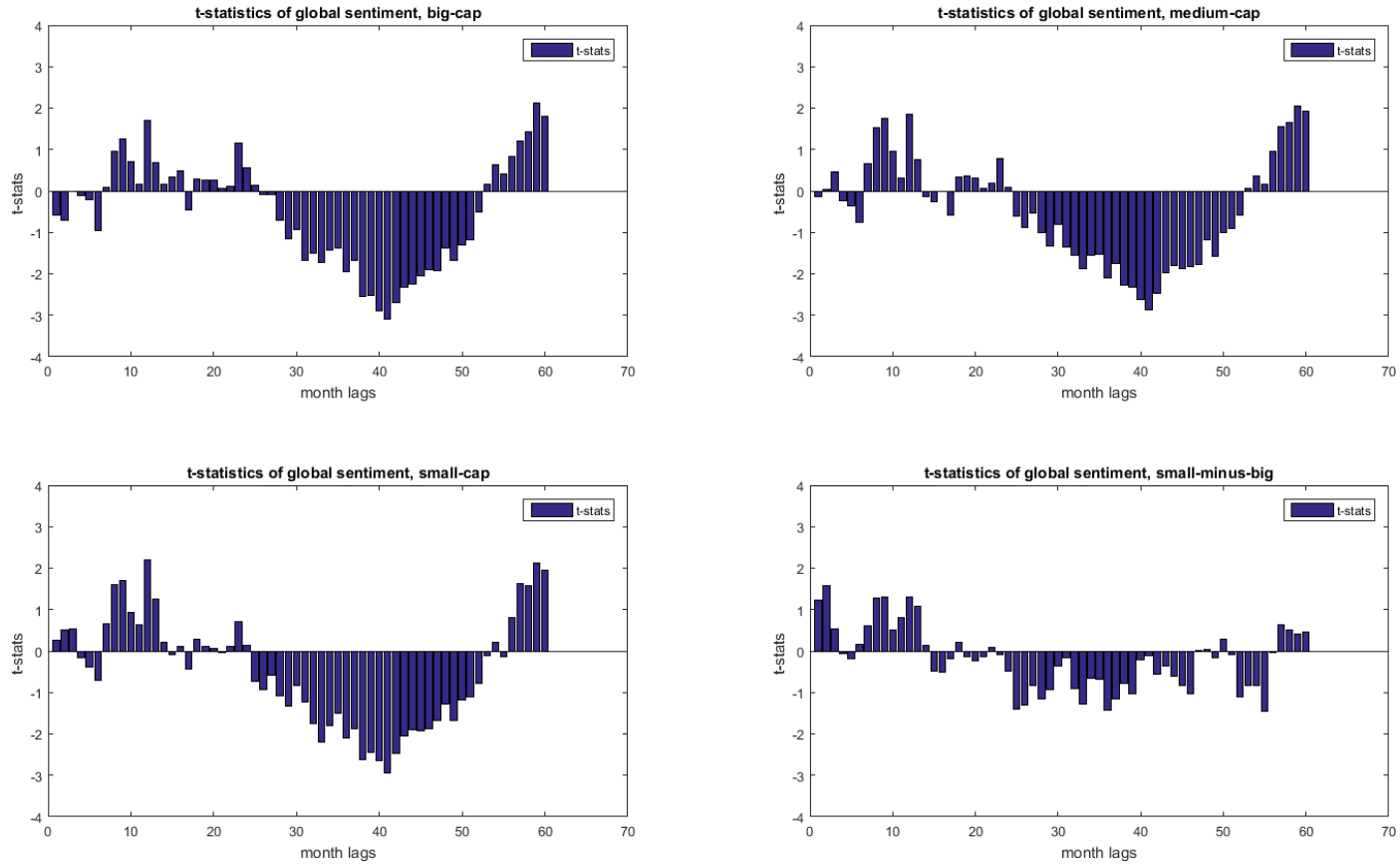
The figure depicts the t -statistics of the slope coefficient on the sentiment index in the multi-horizon predictive regression using OLS methods. We regress the monthly excess return of the big-cap, medium-cap, small-cap, and small-minus-big portfolio on the lagged sentiment index over various horizons. The horizons (month lags) vary from 1 month up till 60 months. The plot is obtained from the entire sample period from July 1997 till December 2013.

Figure 6. Multi-horizon Predictability Patterns in the Cross Section, mARM method



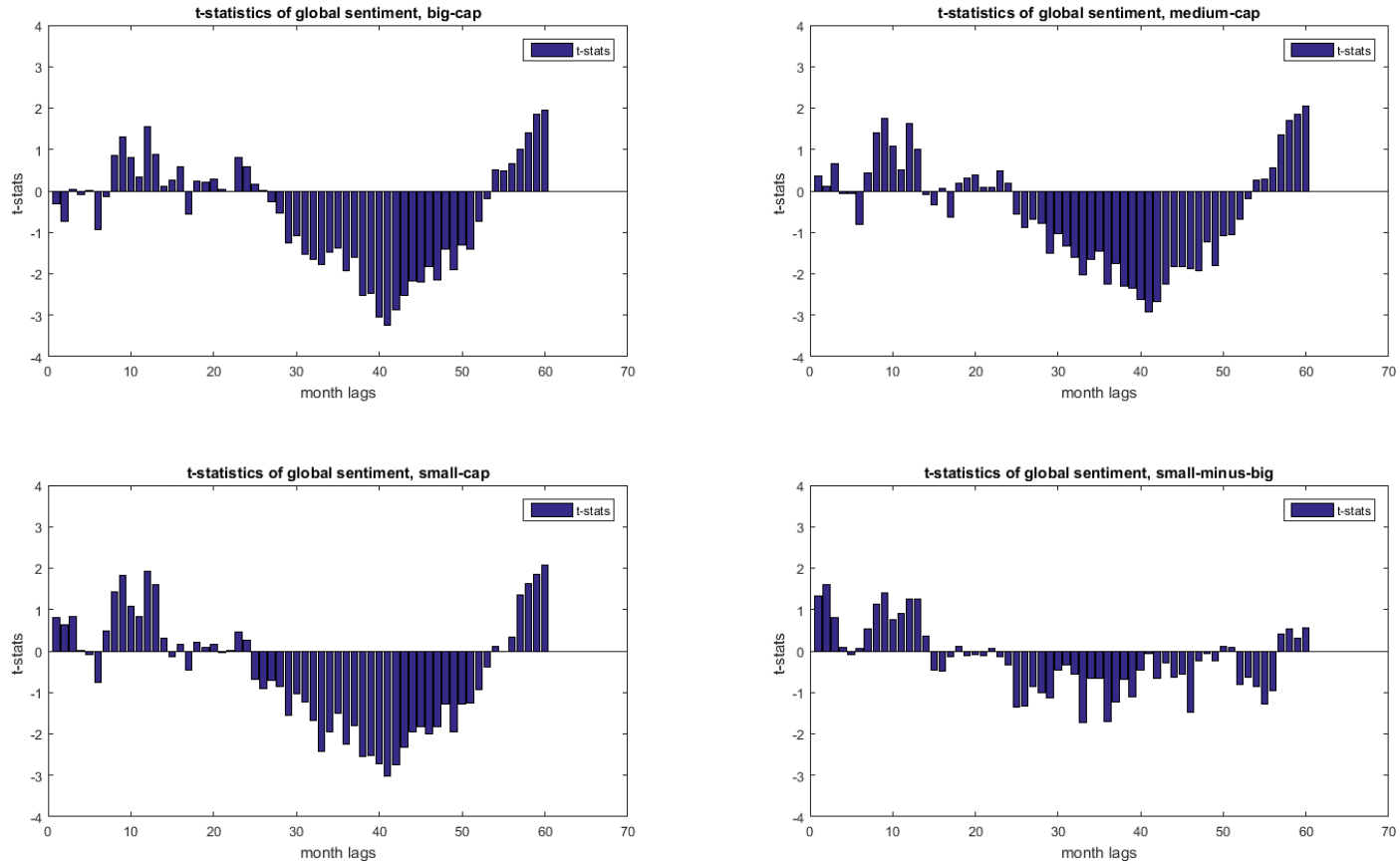
The figure depicts the t -statistics of the slope coefficient on the sentiment index in the multi-horizon predictive regression using mARM methods. We regress the monthly excess return of the big-cap, medium-cap, small-cap, and small-minus-big portfolio on the lagged sentiment index over various horizons. The horizons (month lags) vary from 1 month up till 60 months. The plot is obtained from the entire sample period from July 1997 till December 2013.

Figure 7. Multi-horizon Predictability Patterns of the Global Sentiment for the Size-sorted Portfolios, OLS method



The figure depicts the t -statistics of the slope coefficients on the global sentiment index in the multi-horizon predictive regression using OLS methods. We regress the monthly excess return of the big-cap, medium-cap, small-cap, and small-minus-big portfolio on the lagged local and global sentiment indexes over various horizons. The horizons (month lags) vary from 1 month up till 60 months. The plot is obtained from the entire sample period from July 1997 till December 2013.

Figure 8. Multi-horizon Predictability Patterns of the Global Sentiment for the Size-sorted Portfolios, mARM method



The figure depicts the t -statistics of the slope coefficients on the global sentiment index in the multi-horizon predictive regression using mARM methods. We regress the monthly excess return of the big-cap, medium-cap, small-cap, and small-minus-big portfolio on the lagged local and global sentiment indexes over various horizons. The horizons (month lags) vary from 1 month up till 60 months. The plot is obtained from the entire sample period from July 1997 till December 2013.

Table 1: Data and Data Sources

| Variables | Sources |
|---|--|
| Market capitalization (MV), Monthly return (Ret), Daily turnover (Turn), One-year time deposit rate (RF) | Thomson Reuters Datastream |
| Business cycle indicator (BCI), Industrial production (IP), Number of new investor accounts (NA), Money Supply (M2), 30-day National Interbank Offer Rate (IR), Exchange Rate (CNY/USD) | CEIC database |
| Market-wide PE ratio (PE) | CICC database |
| US sentiment index | Baker and Wurgler (2007) |

Table 2: Descriptive Statistics

In this table, we present the summary statistics for the monthly excess return of the market portfolio (Ret_t), sentiment index constructed by principal component analysis (S_t^{PCA}), the three composite sentiment proxies (all orthogonalized to macro effects): Turnover ratio ($TURN_t^O$), the number of newly opened investor accounts in Shanghai Stock Exchange ($SHIIA_t^O$), and the value-weighted PE ratio (PE_t^O). The business cycle indicator (BCI_t), the growth of the industrial production (IP_t), and the market-wide PE ratio (PE_t). The upper part of the table reports the arithmetic mean, median, standard deviation, skewness, kurtosis, minimum, maximum, 25% and 75% quantile points, and first order autocorrelation coefficient, while the lower part the cross-correlation coefficients. The sample period is from July 1997 till December 2013.

| | Obs. | Mean | Median | St.Dev. | Min | Max | Skew | Kurt | 25% | 75% | Rho |
|-------------|---------|-------------|------------|-------------|----------|---------|---------|--------|--------|------|------|
| Ret_t (%) | 198 | 0.16 | 0.24 | 8.07 | -28.11 | 29.03 | -0.18 | 4.39 | -5.19 | 4.91 | 0.04 |
| S_t^{PCA} | 198 | 0.00 | -0.18 | 1 | -1.85 | 3.75 | 1.06 | 4.58 | -0.65 | 0.55 | 0.70 |
| $TURN_t^O$ | 198 | 0.00 | -0.15 | 1 | -2.77 | 3.19 | 0.77 | 4.37 | -0.68 | 0.56 | 0.56 |
| $SHIIA_t^O$ | 198 | 0.00 | -0.13 | 1 | -2.94 | 3.35 | 0.91 | 4.78 | -0.63 | 0.46 | 0.63 |
| PE_t^O | 198 | 0.00 | 0.01 | 1 | -3.12 | 2.55 | -0.35 | 3.93 | -0.51 | 0.59 | 0.83 |
| CCI_t | 198 | 0.00 | 0.26 | 1 | -2.39 | 1.45 | -0.52 | 2.23 | -0.81 | 0.79 | 0.92 |
| BCI_t | 198 | 0.00 | 0.15 | 1 | -2.42 | 1.76 | -0.44 | 2.52 | -0.61 | 0.74 | 0.97 |
| IP_t | 198 | 0.00 | -0.03 | 1 | -2.89 | 2.66 | -0.10 | 2.69 | -0.81 | 0.80 | 0.58 |
| PE_t | 198 | 0.00 | 0.12 | 1 | -1.93 | 1.69 | -0.28 | 1.85 | -0.94 | 0.81 | 0.98 |
| | Ret_t | S_t^{PCA} | $TURN_t^O$ | $SHIIA_t^O$ | PE_t^O | CCI_t | BCI_t | IP_t | PE_t | | |
| Ret_t | 1 | 0.55 | 0.43 | 0.35 | 0.59 | 0.02 | -0.05 | -0.03 | 0.14 | | |
| S_t^{PCA} | 0.55 | 1 | 0.80 | 0.90 | 0.72 | -0.10 | -0.12 | -0.01 | 0.07 | | |
| $TURN_t^O$ | 0.43 | 0.80 | 1 | 0.63 | 0.29 | -0.15 | -0.12 | -0.03 | -0.17 | | |
| $SHIIA_t^O$ | 0.35 | 0.90 | 0.63 | 1 | 0.52 | -0.04 | -0.09 | 0.00 | 0.09 | | |
| PE_t^O | 0.59 | 0.72 | 0.29 | 0.52 | 1 | -0.07 | -0.08 | 0.02 | 0.26 | | |
| CCI_t | 0.02 | -0.10 | -0.15 | -0.04 | -0.07 | 1 | 0.16 | 0.28 | 0.66 | | |
| BCI_t | -0.05 | -0.12 | -0.12 | -0.09 | -0.08 | 0.16 | 1 | 0.67 | -0.09 | | |
| IP_t | -0.03 | -0.01 | -0.03 | 0.00 | 0.02 | 0.28 | 0.67 | 1 | 0.09 | | |
| PE_t | 0.14 | 0.07 | -0.17 | 0.09 | 0.26 | 0.66 | -0.09 | 0.09 | 1 | | |

Table 3. Single-factor Predictive Regression

The table reports the estimation results of the predictive regression with single regressor. Panel A considers the PCA sentiment index for the entire sample period and two subsample periods. Panel B considers the three individual sentiment composites (all orthogonalized to macro effects): Turnover ratio ($TURN_t^O$), the number of newly opened investor accounts in Shanghai Stock Exchange ($SHIIA_t^O$), and the value-weighted PE ratio (PE_t^O). Panel C considers two alternative sentiment indexes. The first one is the PCA sentiment index excluding PE_t^O as a component (denoted as S_t^{ExPE}), while the second one uses consumer confidence index (CCI_t), a survey-based index. For each predictor variable, we report the slope coefficients obtained from both the OLS method (first line) and the mARM method (second line). *, **, and *** stand for significance at 10%, 5% and 1% level, based on Newey-West t -statistics.

| | <i>Beta</i> | <i>t</i> -stat | <i>p</i> -value | Empirical <i>p</i> -value | R^2 |
|---|-------------|----------------|-----------------|---------------------------|--------|
| Panel A: PCA sentiment index, entire sample | | | | | |
| OLS | 0.0138** | 2.55 | 0.0115 | 0.0150 | 2.95% |
| mARM | 0.0150*** | 3.32 | 0.0011 | 0.0135 | 38.98% |
| Subsample 1: July 1997 – December 2005 | | | | | |
| OLS | 0.0089** | 2.42 | 0.0174 | 0.0375 | 1.73% |
| mARM | 0.0112*** | 3.25 | 0.0016 | 0.0980 | 47.52% |
| Subsample 2: January 2006 – December 2013 | | | | | |
| OLS | 0.0169* | 1.90 | 0.0608 | 0.0480 | 3.44% |
| mARM | 0.0194** | 2.52 | 0.0133 | 0.0320 | 34.57% |
| Panel B: Individual composites of sentiment index, entire sample | | | | | |
| $TURN_t^O$ | | | | | |
| OLS | 0.0085** | 2.18 | 0.0303 | 0.0295 | 1.10% |
| mARM | 0.0091** | 2.39 | 0.0178 | 0.0790 | 20.59% |
| $SHIIA_t^O$ | | | | | |
| OLS | 0.0105 | 1.61 | 0.1100 | 0.0800 | 1.70% |
| mARM | 0.0111* | 1.87 | 0.0636 | 0.0480 | 13.99% |
| PE_t^O | | | | | |
| OLS | 0.0154*** | 3.15 | 0.0019 | 0.0050 | 3.64% |
| mARM | 0.0177*** | 5.69 | 0.0000 | 0.0065 | 64.40% |
| Panel C: Alternative sentiment index, entire sample | | | | | |
| S_t^{ExPE} | | | | | |
| OLS | 0.0105** | 2.19 | 0.0295 | 0.0340 | 1.70% |
| mARM | 0.0112** | 2.42 | 0.0165 | 0.0525 | 21.54% |
| CCI_t | | | | | |
| OLS | -0.0003 | -0.05 | 0.9615 | 0.9900 | 0.00% |
| mARM | 0.0003 | 0.04 | 0.9651 | 0.8900 | 0.35% |

Table 4. Multiple-factor Predictive Regression

The table reports the output of the multi-variate predictive regressions under alternative model specifications. Specification 1 till 3 are bivariate predictive regressions which estimate the sentiment effect while controlling for the effect of business cycle (BCI_t), the growth of the industrial production (IP_t), and the market-wide PE ratio (PE_t). Specification 4 is the kitchen sink model which includes all the three control variables simultaneously. For brevity, we report only the slope coefficients estimated from the mARM method. *, **, and *** stand for significance at 10%, 5% and 1% level, based on Newey-West t -statistics.

| <i>mARM</i> | <i>Beta</i> | <i>t</i> -stat | <i>p</i> -value | Empirical <i>p</i> -value | R^2 |
|-------------------------------|-------------|----------------|-----------------|---------------------------|--------|
| <u><i>Specification 1</i></u> | | | | | |
| S_t^{PCA} | 0.0144*** | 3.67 | 0.0003 | 0.0115 | 40.11% |
| BCI_t | 0.0000 | 0.01 | 0.9883 | 0.4890 | |
| <u><i>Specification 2</i></u> | | | | | |
| S_t^{PCA} | 0.0152*** | 3.23 | 0.0015 | 0.0115 | 40.80% |
| IP_t | 0.0006 | 0.16 | 0.8763 | 0.4875 | |
| <u><i>Specification 3</i></u> | | | | | |
| S_t^{PCA} | 0.0134*** | 4.30 | 0.0000 | 0.0235 | 76.49% |
| PE_t | 0.0041** | 2.12 | 0.0349 | 0.0310 | |
| <u><i>Specification 4</i></u> | | | | | |
| S_t^{PCA} | 0.0139*** | 4.95 | 0.0000 | 0.0230 | 77.16% |
| BCI_t | -0.0018 | -1.06 | 0.2922 | 0.8180 | |
| IP_t | -0.0057* | -1.80 | 0.0736 | 0.3565 | |
| PE_t | 0.0099*** | 3.01 | 0.0030 | 0.1385 | |

Table 5. Out-of-sample Forecast Performance

The table reports the out-of-sample performance of various predictors in predicting subsequent market returns. The out-of-sample forecasts are formed using fitted values of the slope coefficients from OLS, mARM, uARM, and MBJK estimation methods, respectively. Panel A considers the case of univariate predictive regression with either the PCA sentiment index or the naïve (equal-weighted) sentiment index as the predicting variable. Panel B considers the bivariate case with the PCA sentiment as one of the predictors, and another control variable as Business cycle indicator (BCI), growth of industrial production (IP), and PE ratio (PE), respectively. All of the sentiment variables and model parameters are estimated recursively based on information prior to the forecast period. R_{OS}^2 is the Campbell and Thompson (2008) out-of-sample R^2 , MSFE is the mean squared forecast error, DM-test is the Diebold and Mariano (1995) t-statistic, and CW-test is the Clark and West (2007) MSFE-adjusted statistic. *, **, and *** stand for significance at 10%, 5% and 1% level. The out-of-sample period ranges from January 2007 to December 2013.

| Method (Variables) | MSFE | R_{OS}^2 | DM-test | CW-test |
|--|----------------|--------------|-------------|--------------|
| Panel A: Univariate Predictive Regression | | | | |
| OLS (S_{t-1}^{PCA}) | 0.00849 | 3.45% | 0.79 | 1.33* |
| mARM (S_{t-1}^{PCA}) | 0.00845 | 3.89% | 0.79 | 1.39* |
| uARM (S_{t-1}^{PCA}) | 0.00846 | 3.82% | 0.79 | 1.38* |
| MBJK (S_{t-1}^{PCA}) | 0.00851 | 3.25% | 0.70 | 1.29* |
| OLS (S_{t-1}^{Naive}) | 0.00848 | 3.54% | 0.78 | 1.35* |
| mARM (S_{t-1}^{Naive}) | 0.00845 | 3.96% | 0.78 | 1.41* |
| uARM (S_{t-1}^{Naive}) | 0.00845 | 3.90% | 0.78 | 1.40* |
| MBJK (S_{t-1}^{Naive}) | 0.00850 | 3.34% | 0.70 | 1.31* |
| Panel B: Bivariate Predictive Regression | | | | |
| OLS (S_{t-1}^{PCA}, BCI_{t-1}) | 0.00850 | 3.40% | 0.77 | 1.39* |
| mARM (S_{t-1}^{PCA}, BCI_{t-1}) | 0.00863 | 1.84% | 0.37 | 1.23 |
| uARM (S_{t-1}^{PCA}, BCI_{t-1}) | 0.00860 | 2.25% | 0.50 | 1.11 |
| MBJK (S_{t-1}^{PCA}, BCI_{t-1}) | 0.00855 | 2.83% | 0.60 | 1.28* |
| OLS (S_{t-1}^{PCA}, IP_{t-1}) | 0.00856 | 2.67% | 0.62 | 1.21 |
| mARM (S_{t-1}^{PCA}, IP_{t-1}) | 0.00852 | 3.12% | 0.64 | 1.29* |
| mARM (S_{t-1}^{PCA}, IP_{t-1}) | 0.00862 | 1.99% | 0.42 | 1.08 |
| MBJK (S_{t-1}^{PCA}, IP_{t-1}) | 0.00856 | 2.69% | 0.57 | 1.22 |
| OLS (S_{t-1}^{PCA}, PE_{t-1}) | 0.00868 | 1.28% | 0.27 | 0.97 |
| mARM (S_{t-1}^{PCA}, PE_{t-1}) | 0.00856 | 2.63% | 0.51 | 1.37* |
| mARM (S_{t-1}^{PCA}, PE_{t-1}) | 0.00863 | 1.84% | 0.60 | 1.06 |
| MBJK (S_{t-1}^{PCA}, PE_{t-1}) | 0.00868 | 1.34% | 0.29 | 0.93 |

Table 6. Asset allocation results: univariate case

This table reports CEQ gains and Sharpe ratio for a mean-variance investor, who dynamically allocates between equities and the risk-free asset based on the forecasts from predictable variable(s) instead of historical average. The investor is assumed to have a risk aversion coefficient of 0.25, 0.5, and 1, respectively. The portfolio weights (with boundary restrictions) are determined recursively by the forecast on equity returns, formed using OLS, mARM, uARM, and MBJK estimation methods. We evaluate whether the CEQ gain is significantly larger than zero based on asymptotic distributions according to [DeMiguel *et al.* \(2009\)](#). We also perform the one-sided JKM-test on whether the testing portfolio's Sharpe ratio is significantly larger than that of the benchmark portfolio using historical average. *, **, and *** stand for significance at 10%, 5% and 1% level. The out-of-sample period is from January 2007 to December 2013.

| Method (Variables) | $\gamma = 0.25$ | | $\gamma = 0.5$ | | $\gamma = 1.0$ | |
|---------------------------------------|-----------------|--------------|----------------|--------------|----------------|--------------|
| | CEQ gain | Sharpe ratio | CEQ gain | Sharpe ratio | CEQ gain | Sharpe ratio |
| Panel A: PCA sentiment index | | | | | | |
| OLS (S_{t-1}^{PCA}) | 0.0147* | 0.08* | 0.0118* | 0.10* | 0.0106* | 0.13* |
| mARM (S_{t-1}^{PCA}) | 0.0161** | 0.10** | 0.0136* | 0.13* | 0.0115* | 0.14* |
| uARM (S_{t-1}^{PCA}) | 0.0157** | 0.09** | 0.0134* | 0.12* | 0.0114* | 0.14* |
| MBJK (S_{t-1}^{PCA}) | 0.0146* | 0.08* | 0.0092 | 0.07 | 0.0089 | 0.11 |
| Panel B: Naïve sentiment index | | | | | | |
| OLS ($S_{t-1}^{Naïve}$) | 0.0149** | 0.08** | 0.0112* | 0.10* | 0.0103* | 0.12* |
| mARM ($S_{t-1}^{Naïve}$) | 0.0157** | 0.09** | 0.0128* | 0.12* | 0.0112* | 0.14* |
| uARM ($S_{t-1}^{Naïve}$) | 0.0155** | 0.09** | 0.0126* | 0.11* | 0.0111* | 0.13* |
| MBJK ($S_{t-1}^{Naïve}$) | 0.0147* | 0.08* | 0.0094 | 0.08 | 0.0085 | 0.10 |
| Panel C: BCI | | | | | | |
| OLS (BCI_{t-1}) | 0.0129** | 0.06** | 0.0079** | 0.06** | 0.0043** | 0.05*** |
| mARM (BCI_{t-1}) | 0.0108** | 0.04** | 0.0057** | 0.04** | 0.0028 | 0.04** |
| uARM (BCI_{t-1}) | 0.0106** | 0.04** | 0.0054** | 0.03** | 0.0031 | 0.04** |
| MBJK (BCI_{t-1}) | 0.0143*** | 0.07** | 0.0106*** | 0.09*** | 0.0066** | 0.08*** |
| Panel D: IP | | | | | | |
| OLS (IP_{t-1}) | -0.0015 | -0.08 | -0.0026 | -0.05 | -0.0026 | -0.03 |
| mARM (IP_{t-1}) | -0.0016 | -0.08 | -0.0025 | -0.05 | -0.0026 | -0.03 |
| uARM (IP_{t-1}) | -0.0016 | -0.08 | -0.0025 | -0.05 | -0.0026 | -0.03 |
| MBJK (IP_{t-1}) | 0.0016 | -0.03 | 0.0000 | -0.02 | -0.0007 | 0.00 |
| Panel E: PE | | | | | | |
| OLS (PE_{t-1}) | -0.0025 | -0.06 | -0.0091 | -0.10 | -0.0094 | -0.14 |
| mARM (PE_{t-1}) | 0.0134* | 0.07** | 0.0083* | 0.07* | 0.0044** | 0.05* |
| uARM (PE_{t-1}) | 0.0035 | -0.03 | 0.0011 | -0.01 | 0.0007 | 0.00 |
| MBJK (PE_{t-1}) | -0.0019 | -0.09 | -0.0055 | -0.12 | -0.0050 | -0.14 |

Table 7. Trading Strategies

This table reports both the gross return and the return net of transaction costs for the market timing strategy. The returns are expressed in percentage points. The market timing strategy mimics the short-selling constraints in China, and proceeds as follows: At the end of each month, we will go long the market portfolio (for the next month) if the trading signal is above the threshold, and liquidate our positions otherwise (i.e. we will only go short if we previously held the market portfolio in the prior period). We report the performance of a simple buy-and-hold strategy, denoted as B&H; a timing strategy based on the buying signal when lagged sentiment is above zero, denoted as S_t^{PCA} ; and a timing strategy based on the buying signal when the predicted excess return generated by four different estimation techniques (OLS, mARM, uARM, and MBJK) is above zero. The out-of-sample evaluation period is from January 2007 to December 2013.

| | B&H | S_t^{PCA} | OLS (S_t^{PCA}) | mARM (S_t^{PCA}) | uARM (S_t^{PCA}) | MBJK (S_t^{PCA}) |
|---|-------|-------------|---------------------|----------------------|----------------------|----------------------|
| Panel A: trading signal based on lagged sentiment value is above zero | | | | | | |
| Return (pre-transaction cost) | -0.12 | 0.90 | 0.40 | 0.50 | 0.50 | 0.40 |
| Return (post-transaction cost) | -0.13 | 0.62 | 0.07 | 0.20 | 0.20 | 0.07 |
| # transactions | 1 | 23 | 27 | 25 | 25 | 27 |

Appendix A: Predictive Regression, Estimation Bias, and Monte Carlo Simulation

Assessing the return predictability of one economic variable is not only an economic issue but also an econometric one. Even for a univariate predictive regression, the standard OLS method produces biased coefficient (the well-known [Stambaugh \(1999\)](#) small sample bias), when the regressor is persistent and its innovations are correlated with those of the dependent variable. To illustrate this point, we follow the general setup in the econometrics literature by considering the following (single-regressor) predictive regression model.

$$y_t = \alpha + \beta x_{t-1} + u_t \quad [\text{A.1}]$$

$$x_t = \theta + \rho x_{t-1} + v_t \quad [\text{A.2}]$$

where y_t is the stock or portfolio returns (in excess of risk-free rate) at period t , x_{t-1} is the lagged predictor variable which follows an AR(1) process as specified in Eq. [A.2], and the innovation terms (u_t, v_t) follow a joint normal distribution with zero mean and a covariance matrix

$$\begin{pmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{pmatrix}. \quad [\text{A.3}]$$

The relationship between u_t and v_t can also be expressed in the following equation.

$$u_t = \xi v_t + e_t \quad [\text{A.4}]$$

where $\xi = \sigma_{uv}/\sigma_v^2$ and e_t is the i.i.d. error term.

Given the above setup, [Stambaugh \(1999\)](#) provides an analytical expression for the bias of the OLS estimator, $\hat{\beta}$, which is exact for any given sample size T . (The bias in $\hat{\beta}$ depends on both ξ and the estimation bias in the autoregressive estimator, $\hat{\rho}$.)

$$E(\hat{\beta}) - \beta = \xi[E(\hat{\rho}) - \rho] \quad [\text{A.5}]$$

Build on the work of [Marriott and Pope \(1954\)](#) on the bias in the estimation of the autocorrelation, [Stambaugh \(1999\)](#) further provides an approximation for the bias in $\hat{\beta}$.

$$E(\hat{\beta}) - \beta = -\xi \frac{(1 + 3\rho)}{T} + O\left(\frac{1}{T^2}\right) \quad [\text{A.6}]$$

As it stands, the magnitude of the bias is proportional to ξ and the (true) autoregressive coefficient, ρ , but is inversely proportional to sample size T .

One caveat to be noted is that unlike many other predictive variables (such as dividend yields), the innovations of sentiment has a positive correlation with those of stock returns. That is, a positive ξ in the case of stock return and investor sentiment in Eq. [A.4]. Given that the (true) autocorrelation coefficient ρ is bounded in $(0, 1)$ for a stationary autoregressive process, the $\hat{\beta}$

coefficient in Eq. [A.1] is bound to be downward biased (meaning that the OLS estimator is smaller than the true parameter, see Eq. [A.6]). The downward bias is more pronounced when the sample size is small. To sum up, empirical studies tend to be biased towards the finding of contrarian predictability of sentiment at short-horizons (weekly or monthly frequency), even if there exists no return predictability at all (or very small positive return predictability). To illustrate this point, we carried out simulations to provide additional insights on the potential adverse impact on the small-sample bias problem. Of course, the small-sample bias has made itself a hot topic in the econometric literature, as multiple bias-reduction techniques have been proposed in the literature ([Amihud & Hurvich 2004](#); [Amihud et al. 2009](#); [Chiquoine & Hjalmarsson 2009](#); [Zhu 2013](#)).

We provide below two simulation cases whose parameters are designed to mirror our sample statistics. Estimation results of standard OLS method and several bias-reduction techniques, including the univariate Augmented Regression Method (uARM) as in [Amihud and Hurvich \(2004\)](#), the multivariate Augmented Regression Method (mARM) of [Amihud et al. \(2009\)](#) and the Moving-block Jackknife (MBJK) method by [Zhu \(2013\)](#), are presented for comparison purpose.

We simulate stock market returns and investor sentiment data using Eq. [A.1] and [A.2], where the parameters α and θ are set as zero, the autoregressive root ρ is set as 0.6597, and the innovation terms (u_t, v_t) follow a joint normal distribution with zero mean and a covariance matrix $\begin{pmatrix} 0.0066 & 0.0372 \\ 0.0372 & 0.5883 \end{pmatrix}$. The sample size is fixed at 202. These parameter values follow closely our sample characteristics. To illustrate the adverse impact due to finite-sample bias, we set the β coefficient equal to 0 and 0.01 in case 1 and 2, respectively. **Table A.1** presents the simulation results.

As is demonstrated by the simulation evidence in **Table A.1**, the slope coefficient of the predictive regression using OLS method is downward biased in both simulation cases. Even if there is no real return predictability (as in case 1), the OLS method tend to find negative slope values which might lead scholars to (erroneously) conclude that investor sentiment is a contrarian predictors at short horizons (eg. weekly or monthly frequency). Even if sentiment is a momentum predictor (as in case 2), its economic relevance tends to be smaller as evaluated by OLS method due to the downward bias. In case 2, it seems that the OLS bias in the slope coefficient amounts to 10% of the true parameter, which is not negligible in real applications. **Table A.1** also offers a small scale horserace on the competing method of bias reduction. In both cases, the mARM

method tends to provide the best estimation results, as its bias is the smallest while the variation of the slope coefficient is also among the smallest. The uARM method performs slightly worse than mARM in case 1, as it leads to larger bias in absolute terms. Although the MBJK method reduces the estimation bias quite satisfactory, it, however, produces larger estimation uncertainty as indicated by its inflated standard errors, which is also admitted in [Zhu \(2013\)](#). Overall, our simulation evidence favors the use mARM in our empirical application of the article.

Table A.1. Finite-sample performance in univariate regressions

The table reports the estimated slope coefficient (β) in Eq. [A.1], denoted as *Coef.* and its standard error, denoted as *S.E.*, using the simulated standard deviations based on 1500 realizations for two simulation cases. The estimation methods are ordinary least square method (OLS), the multivariate Augmented Regression Method (mARM), the univariate Augmented Regression Method (uARM), and the Moving-block Jackknife method (MBJK) with the length of the moving-block subsamples set as 0.3, 0.5, and 0.7 (of the total sample length) respectively.

In both cases, stock market returns and investor sentiment data are simulated with Eq. [A.1] and [A.2], where $\alpha = \theta = 0$, $\rho = 0.6597$, and the innovation terms (u_t, v_t) follow a bivariate normal distribution with mean zero and a covariance matrix $\begin{pmatrix} 0.0066 & 0.0372 \\ 0.0372 & 0.5883 \end{pmatrix}$. The sample size is fixed at 202. The β coefficient equals to 0 and 0.01 in case 1 and 2, respectively. For each case, 1500 replications are performed.

| | Case 1: $\beta = 0$ | | Case 2: $\beta = 0.01$ | |
|--------|---------------------|-------------|------------------------|-------------|
| | <i>Coef.</i> | <i>S.E.</i> | <i>Coef.</i> | <i>S.E.</i> |
| OLS | -0.0012 | 0.0056 | 0.0091 | 0.0057 |
| mARM | -0.0002 | 0.0056 | 0.0100 | 0.0057 |
| uARM | -0.0003 | 0.0056 | 0.0100 | 0.0057 |
| MBJK_3 | -0.0003 | 0.0058 | 0.0101 | 0.0059 |
| MBJK_5 | -0.0004 | 0.0065 | 0.0101 | 0.0066 |
| MBJK_7 | -0.0005 | 0.0086 | 0.0102 | 0.0084 |

Table A.2. Statistics of Stock Exchanges**Panel A: Ranking of the major stock exchanges by the total value of market capitalization as of 31/12/2013**

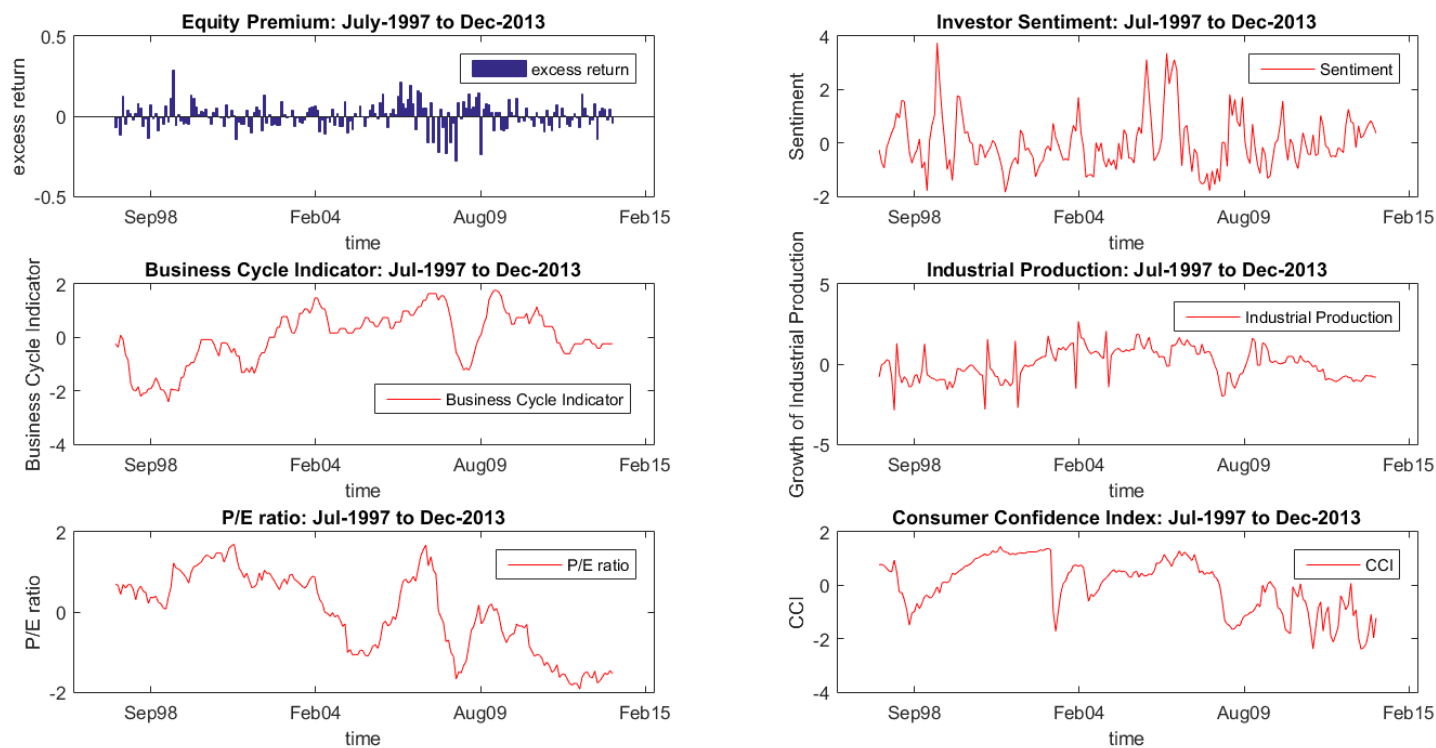
| Exchange Name | # of Stocks | | (1). Market Capitalization | (2). Trading Volume | (2)/(1) |
|---|-------------|---------|----------------------------|---------------------|---------|
| | Local | Foreign | (in USD million) | (in USD million) | |
| 1 NYSE | 1852 | 519 | 17,949,883.80 | 13,700,450.50 | 76.33% |
| 2 NASDAQ | 2328 | 309 | 6,084,969.70 | 9,584,742.20 | 157.52% |
| 3 Japan Exchange Group | 3408 | 11 | 4,543,169.10 | 6,304,927.50 | 138.78% |
| 4 London Stock Exchange | 2164 | 572 | 4,428,975.30 | 3,050,891.50 | 68.88% |
| 5 Euronext (Amsterdam, Brussel, Lisbon, Paris) | 935 | 127 | 3,583,899.70 | 1,661,878.30 | 46.37% |
| 6 HongKong Stock Exchange | 1553 | 90 | 3,100,777.20 | 1,323,373.30 | 42.68% |
| 7 Shanghai Stock Exchange | 953 | 0 | 2,496,989.90 | 3,731,128.90 | 149.43% |
| 8 TMX Group – Canada | 3810 | 76 | 2,113,821.80 | 1,371,477.70 | 64.88% |
| 9 Deutsche Börse | 639 | 81 | 1,936,106.30 | 1,334,544.90 | 68.93% |
| 10 SIX Swiss Exchange | 236 | 36 | 1,540,699.80 | 676,957.70 | 43.94% |
| 11 Shenzhen Stock Exchange | 1536 | 0 | 1,452,153.60 | 3,858,509.00 | 265.71% |
| 12 Australian Securities Exchange | 1951 | 104 | 1,365,958.10 | 881,555.60 | 64.54% |

Panel B: Breakdown of the ownership structure and total trading volume for the Shanghai Stock Exchange as of 31/12/2013

| | Market Capitalization | Trading Volume |
|-------------------------|-----------------------|----------------|
| Individual Investors | 21.78% | 82.24% |
| Legal Person | 63.64% | 2.46% |
| Financial Institutions: | 14.58% | 15.3% |
| - Investment Funds | 4.54% | 6.18% |

Source: Statistics of the major exchanges are from the yearly reports of the World Federation of Exchanges (<http://www.world-exchanges.org>); Statistics of the breakdown data of Shanghai Stock Exchange are from Annual Fact Book issued by Shanghai Stock Exchange. Note: Trading volume is measured as the dollar trading volume of the entire year. Financial Institutions include brokerage and security firms, investment funds (mutual funds), pension funds, asset under management, and QFII (qualified foreign institutional investor). Legal Person includes all other companies, corporations, enterprises, departments, juridical association and agents.

Figure. A.1. Time-series plots of excess market returns, investor sentiment, business cycle indicator and industrial production



The figure includes the bar chart of the (monthly) market excess return (measured in decimals) and the time-series plots of the investor sentiment index obtained from the principal component analysis, the business cycle indicator, the growth of industry production, the value-weighted PE ratio, and the consumer confidence index. All the economic and financial variables (except the excess returns) are standardized to have mean zero and unit variance. The time-series plot covers the entire sample period from July 1997 till December 2013.

List of Reference

- Amihud, Y., Hurvich, C.M., 2004. Predictive Regressions: A Reduced-Bias Estimation Method. *The Journal of Financial and Quantitative Analysis* 39, 813-841
- Amihud, Y., Hurvich, C.M., Wang, Y., 2009. Multiple-Predictor Regressions: Hypothesis Testing. *Review of Financial Studies* 22, 413-434
- Baker, M., Stein, J.C., 2004. Market liquidity as a sentiment indicator. *Journal of Financial Markets* 7, 271-299
- Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. *The Journal of Finance* 61, 1645-1680
- Baker, M., Wurgler, J., 2007. Investor sentiment in the stock market. *The Journal of Economic Perspectives* 21, 129-151
- Baker, M., Wurgler, J., Yuan, Y., 2012. Global, local, and contagious investor sentiment. *Journal of Financial Economics* 104, 272-287
- Banerjee, A.V., 1992. A Simple Model of Herd Behavior. *The Quarterly Journal of Economics* 107, 797-817
- Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. *Journal of Financial Economics* 49, 307-343
- Bodie, Z., Kane, A., Marcus, A.J., 2012. *Essentials of Investments*. McGraw-hill higher education
- Boudoukh, J., Richardson, M., Whitelaw, R.F., 2008. The Myth of Long-Horizon Predictability. *Review of Financial Studies* 21, 1577-1605
- Brown, G.W., Cliff, M.T., 2004. Investor sentiment and the near-term stock market. *Journal of Empirical Finance* 11, 1-27
- Brown, G.W., Cliff, M.T., 2005. Investor Sentiment and Asset Valuation*. *The Journal of Business* 78, 405-440
- Burdekin, R.C.K., Redfern, L., 2009. Sentiment effects on Chinese share prices and savings deposits: The post-2003 experience. *China Economic Review* 20, 246-261
- Campbell, J.Y., Thompson, S.B., 2008. Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? *Review of Financial Studies* 21, 1509-1531
- Chan, K., Menkveld, A.J., Yang, Z., 2008. Information Asymmetry and Asset Prices: Evidence from the China Foreign Share Discount. *The Journal of Finance* 63, 159-196
- Chen, H., Chong, T.T.L., She, Y., 2014. A principal component approach to measuring investor sentiment in China. *Quantitative Finance* 14, 573-579
- Chi, L., Zhuang, X., Song, D., 2012. Investor sentiment in the Chinese stock market: an empirical analysis. *Applied Economics Letters* 19, 345-348
- Chiquoine, B., Hjalmarsson, E., 2009. Jackknifing stock return predictions. *Journal of Empirical Finance* 16, 793-803
- Clark, T.E., West, K.D., 2007. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* 138, 291-311
- De Long, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990a. Noise trader risk in financial markets. *Journal of political economy* 98, 703-38
- De Long, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990b. Positive Feedback Investment Strategies and Destabilizing Rational Speculation. *The Journal of Finance* 45, 379-395

- DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies* 22, 1915-1953
- Diebold, F.X., Mariano, R.S., 1995. Comparing Predictive Accuracy. *Journal of Business & Economic Statistics* 13, 253-263
- Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3-56
- Fama, E.F., French, K.R., 2012. Size, value, and momentum in international stock returns. *Journal of Financial Economics* 105, 457-472
- Feng, L.E.I., Seasholes, M.S., 2004. Correlated trading and location. *The Journal of Finance* 59, 2117-2144
- Ferson, W.E., Sarkissian, S., Simin, T.T., 2003. Spurious Regressions in Financial Economics? *The Journal of Finance* 58, 1393-1414
- Fisher, K.L., Statman, M., 2000. Investor sentiment and stock returns. *Financial Analysts Journal* 56
- Frazzini, A., Lamont, O.A., 2008. Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of Financial Economics* 88, 299-322
- Frömmel, M., Han, X., 2014. Understanding the Controversy of Liquidity Beta. Available at SSRN
- Goyal, A., Welch, I., 2008. A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. *Review of Financial Studies* 21, 1455-1508
- Huang, D., Jiang, F., Tu, J., Zhou, G., 2014. Investor Sentiment Aligned: A Powerful Predictor of Stock Returns. *Review of Financial Studies*
- Indro, D.C., 2004. Does mutual fund flow reflect investor sentiment? *The Journal of Behavioral Finance* 5, 105-115
- Jobson, J.D., Korkie, B.M., 1981. Performance Hypothesis Testing with the Sharpe and Treynor Measures. *The Journal of Finance* 36, 889-908
- Kling, G., Gao, L., 2008. Chinese institutional investors' sentiment. *Journal of International Financial Markets, Institutions and Money* 18, 374-387
- Kumar, A., Lee, C.M.C., 2006. Retail investor sentiment and return comovements. *The Journal of Finance* 61, 2451-2486
- Lee, C.M.C., Shleifer, A., Thaler, R.H., 1991. Investor Sentiment and the Closed-End Fund Puzzle. *The Journal of Finance* 46, 75-109
- Lemmon, M., Portniaguina, E., 2006. Consumer Confidence and Asset Prices: Some Empirical Evidence. *Review of Financial Studies* 19, 1499-1529
- Lo, A., MacKinlay, A., 1990. Data-snooping biases in tests of financial asset pricing models. *Review of Financial Studies* 3, 431-467
- Marriott, F.H.C., Pope, J.A., 1954. Bias in the Estimation of Autocorrelations. *Biometrika* 41, 390-402
- Mei, J., Scheinkman, J.A., Xiong, W., 2009. Speculative trading and stock prices: Evidence from Chinese AB share premia. *Annals of Economics and Finance* 10, 225-255
- Mommel, C., 2003. Performance hypothesis testing with the Sharpe ratio. *Finance Letters* 1
- Nicholls, D.F., Pope, A.L., 1988. Bias in the Estimation Of Multivariate Autoregressions. *Australian Journal of Statistics* 30A, 296-309
- Schmeling, M., 2007. Institutional and individual sentiment: Smart money and noise trader risk? *International Journal of Forecasting* 23, 127-145
- Schmeling, M., 2009. Investor sentiment and stock returns: Some international evidence. *Journal of Empirical Finance* 16, 394-408

- Shleifer, A., Summers, L.H., 1990. The Noise Trader Approach to Finance. *The Journal of Economic Perspectives* 4, 19-33
- Stambaugh, R.F., 1999. Predictive regressions. *Journal of Financial Economics* 54, 375-421
- Stambaugh, R.F., Yu, J., Yuan, Y., 2012. The short of it: Investor sentiment and anomalies. *Journal of Financial Economics* 104, 288-302
- Tian, L., 2011. Regulatory underpricing: Determinants of Chinese extreme IPO returns. *Journal of Empirical Finance* 18, 78-90
- Verma, R., Soydemir, G., 2009. The impact of individual and institutional investor sentiment on the market price of risk. *The Quarterly Review of Economics and Finance* 49, 1129-1145
- Warther, V.A., 1995. Aggregate mutual fund flows and security returns. *Journal of Financial Economics* 39, 209-235
- Zhu, M., 2013. Jackknife for Bias Reduction in Predictive Regressions. *Journal of Financial Econometrics* 11, 193-220