

Yang, Yan; Copeland, Laurence

**Working Paper**

## The Effects of Sentiment on Market Return and Volatility and The Cross-Sectional Risk Premium of Sentiment- affected Volatility

Cardiff Economics Working Papers, No. E2014/12

**Provided in Cooperation with:**

Cardiff Business School, Cardiff University

*Suggested Citation:* Yang, Yan; Copeland, Laurence (2014) : The Effects of Sentiment on Market Return and Volatility and The Cross-Sectional Risk Premium of Sentiment-affected Volatility, Cardiff Economics Working Papers, No. E2014/12, Cardiff University, Cardiff Business School, Cardiff

This Version is available at:

<https://hdl.handle.net/10419/109058>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Cardiff Economics Working Papers



Working Paper No. E2014/12

## The Effects of Sentiment on Market Return and Volatility and The Cross-Sectional Risk Premium of Sentiment-affected Volatility

*Yan Yang and Laurence Copeland*

July 2014

Cardiff Business School  
Aberconway Building  
Colum Drive  
Cardiff CF10 3EU  
United Kingdom  
t: +44 (0)29 2087 4000  
f: +44 (0)29 2087 4419  
[business.cardiff.ac.uk](http://business.cardiff.ac.uk)

This working paper is produced for discussion purpose only. These working papers are expected to be published in due course, in revised form, and should not be quoted or cited without the author's written permission.

Cardiff Economics Working Papers are available online from:

[econpapers.repec.org/paper/cdfwpaper/](http://econpapers.repec.org/paper/cdfwpaper/) and

[business.cardiff.ac.uk/research/academic-sections/economics/working-papers](http://business.cardiff.ac.uk/research/academic-sections/economics/working-papers)

Enquiries: [EconWP@cardiff.ac.uk](mailto:EconWP@cardiff.ac.uk)

# The Effects of Sentiment on Market Return and Volatility and The Cross-Sectional Risk Premium of Sentiment-affected Volatility

Yan Yang and Laurence Copeland

April 2014

Cardiff Business School

Cardiff University

## **Abstract**

We construct investor sentiment of UK stock market using the procedure of principal component analysis. Using sentiment-augmented EGARCH component model, we analyse the impacts of sentiment on market excess return, the permanent component of market volatility and the transitory component of market volatility. Bullish sentiment leads to higher market excess return while bearish sentiment leads to lower excess return. Sentiment-augmented EGARCH component model compares favourably to the original EGARCH component model which does not take investor sentiment into account. Furthermore, we test the cross-sectional risk premia of the permanent and transitory components of sentiment-affected volatility in the framework of ICAPM.

**Key Terms:** investor sentiment; principal component analysis; EGARCH component model; ICAPM; cross-sectional risk premium

**JEL Codes:** G12; G15

## **Corresponding Author**

**Yan Yang**

**Cardiff Business School**

**Aberconway Building**

**Colum Drive**

**Cardiff CF10 3EU**

**Email: [yangy16@cardiff.ac.uk](mailto:yangy16@cardiff.ac.uk); [grace\\_yang@hotmail.co.uk](mailto:grace_yang@hotmail.co.uk)**

## 1. Introduction

A long-running debate in financial economics concerns the role and possible effect of investor sentiment on asset prices. There are various ways to define investor sentiment. It is the feeling or tone of a market, or its crowd psychology, as revealed through the activity and price movement of the securities traded in that market. Market sentiment is also called "investor sentiment" and is not always based on fundamentals. Baker and Wurgler (2006) explain it as the propensity to speculate or the overall optimism or pessimism about an asset. Baker and Wurgler (2007) further define investor sentiment broadly, as a belief, usually influenced by emotion, about future cash flows and investment risks that is not justified by the facts at hand.

Traditional asset pricing theory suggests that rational arbitrage necessarily forces prices closer to fundamentals and leaves no role for investor sentiment. The capital asset pricing model (CAPM) theoretically argues that systematic risk is measured by the exposure to the market portfolio. Prior literature has shown, however, that the standard CAPM cannot explain the returns on stocks with certain firm characteristics or price histories such as the size effect, value effect and momentum effect which have been termed as asset-pricing anomalies in literature. In the attempt to capture the dimensions of risk other than exposure to the market risk, Fama and French (1992, 1993) further include size and value factors and Pastor and Stambaugh (2003) consider a liquidity factor.

The existing literature document that investor sentiment exhibits certain degree of predictability of time-series stock returns. Fisher and Statman (2003) reveal the level of investor sentiment in one month is negatively related to the stock returns over the next month and the next 6 or 12 months. Meanwhile, there's a positive relationship between the monthly changes in investor sentiment and contemporaneous market excess returns. Brown and Cliff (2004, 2005) suggest that their measures of sentiment co-move with the market in the long run. They find that returns over future multiyear horizons are negatively associated with investor sentiment. Lee et al. (2002) demonstrate that excess returns are contemporaneously positively related to shifts in sentiment.

Baker and Wurgler (2006, 2007) show that sentiment affects the cross-section of stock returns. They give us an excellent illustration of the theoretical effects of sentiment on the cross-section. The two main channels through which sentiment can affect pricing are investor sentiment and arbitrage. Under the first channel, sentimental demand shocks vary across stocks while arbitrage limits are constant. Investor sentiment might be interpreted as the propensity to speculate. Sentiment drives the relative demand for stocks that are more vulnerable to speculation, whose valuations are subjective and difficult to determine. For instance, small, young, extreme growth, unprofitable, and non-dividend paying stocks, should be more difficult to price. Therefore, opaque stocks are more vulnerable to broad shifts in investor sentiment. On the other hand, translucent stocks are less likely to be affected by fluctuations in the propensity to speculate.

Under the second channel, investor sentiment might be interpreted as optimism or pessimism about stocks in general. The effect of investor sentiment will be uniform while the arbitrage forces differ among stocks. A body of theoretical and empirical research shows that arbitrage is particularly costly and risky with certain stock types, for example, young, small, extreme growth, and unprofitable stocks.

The two channels appear to affect the same type of stocks. Put another way, the same stocks that are the hardest to arbitrage also tend to be the most difficult to value in practice. Hence, the two channels are likely to have overlapping effects and reinforce each other.

The behavioural finance literature shows that sentiment has impact on trading decisions. The influence of investors' future expectation can lead to the over- or under-pricing of stocks, and thus affect pricing models. Various studies provide supportive evidence that investor sentiment plays a critical role in determining stock price behaviour. Hence, the question now is no longer whether investor sentiment affects stock prices, but how to measure investor sentiment and quantify its effects.

In this chapter, our study focuses on three aspects. First, limited by the availability of data, we get five sentiment measures, consumer confidence, market turnover by volume, market turnover by value, the number of IPOs in each month, and the initial day return of IPOs within each month. We use the first principal component analysis (PCA) to construct a composite sentiment index. Second, we test the time-series relationship of market level excess returns and investor sentiment. We examine whether investor sentiment affects time-series market excess returns. Also, we examine whether market excess returns are indirectly affected by investor sentiment through the risk caused by sentiment in the form of volatility. Third, in the cross-sectional examination, we demonstrate whether the short- and long-run sentiment-affected volatilities obtained from the time-series regression are priced factors in the 25 Fama-French size and BTM sorted portfolios.

Inspired by the empirical framework of Lee et al. (2002), we augment investor sentiment to Adrian and Rosenberg's EGARCH component model by adding the sentiment to mean and variance equations. However, our model differs from their model in the following aspects. First, the main framework of our model is EGARCH component model proposed by Adrian and Rosenberg (2008), while Lee et al. (2002) utilize the GARCH-in-mean framework. Furthermore, Adrian and Rosenberg (2008) do not consider investor sentiment in their model. Second, the dummy variables of sentiment enters their model as intercept dummy which suggests sentiment affects the conditional volatility directly. However, our investigation suggests that the direct effects of sentiment on short- and long-run volatilities are statistically insignificant. Therefore, we adapt the approach by adding the dummy variables as slope dummies rather than intercept dummies. The empirical results turn out that sentiment has influences on short- and long-run volatilities through their impacts on previous short- and long-run volatilities. Third, both the level and change of sentiment are investigated in each model framework.

The analysis and results of this chapter contributes to the existing literature by investigating the extent to which the impact of investor sentiment on stock market volatility and returns. Furthermore, we examine the cross-sectional prices of risks of sentiment-affected volatilities. First, the results complement earlier work that shows sentiment help to explain the time-series of returns. Previous research has focused on the influence of investor sentiment on the mean of stock returns. Our study investigates the impact of investor sentiment on both the market excess returns and the volatility of returns. Second, most research utilises the U.S. data and to our best knowledge, there are very few empirical research on market sentiment concentrating on U.K. market. The existing studies related to U.K. market are mainly discussing the international sentiment and U.K. market is just one part of the European or global market. Third, we decompose the market volatility into transitory and permanent components. By applying investor sentiment to this model, we investigate the influences of sentiment on decomposed market volatility and examine effects on short- and long-run volatilities separately. Finally, we further examine whether the short- and long-run sentiment-affected volatility are priced factors in the cross-section.

The chapter is organized as follows. The next section displays a brief literature review on investor sentiment and a survey of proxies for sentiment proposed in the literature. Section 3 provides the methodology of the construction of investor sentiment and the empirical model of this chapter. Section 4 describes the summary statistics of sentiment measures and the construction of the composite index. Section 5 shows us the empirical analysis of the time-series and cross-sectional estimations. Section 6 provides robustness checks for both the model specification and measure of investor sentiment and the last section concludes.

## 2. Literature review

### 2.1 Measures of Investor Sentiment

#### 2.1.1 Economic variables as sentiment measures

The existing literature has established several different measures of investor sentiment. One approach is directly through economic variables. A number of studies use observable economic variables to measure levels of sentiment. Baker and Wurgler (2007) summarize some potential economic proxies for sentiment, including retail investor trades; mutual fund flows; trading volume; premia on dividend-paying stocks; closed-end fund discounts; option implied volatility; first day returns on initial public offerings (IPOs); volume of initial public offerings; new equity issues; and insider trading.

*Retail Investor Trades.* Barber et al. (2007) and Kumar and Lee (2006) find in micro-level trading data that trading of retail investors is highly correlated and persistent, which is consistent with systematic sentiment. Consequently, Kumar and Lee (2006) suggest constructing sentiment measures for retail investors based on their trading comovements.

*Mutual Fund Flows.* Brown et al. (2002) find evidence that daily mutual fund flows may be instruments for investor sentiment about the stock market and provide evidence that this sentiment factor is priced. Frazzini and Lamont (2006) find some affirmative evidence by using fund flows to proxy for sentiment for individual stocks. They find that strong inflows of stock within a mutual fund, predicts a relative low future return.

*Trading Volume.* Trading volume, or more generally liquidity, can be viewed as an investor sentiment index. Baker and Stein (2004) note that in the presence of short-sales constraints, which is actually the case in practice, irrational investors are more likely to trade, and thus add liquidity, when they are optimistic and betting on rising stocks rather than when they are pessimistic and betting on falling stocks. Higher turnover predict lower subsequent returns in both firm-level and aggregate data. Similarly, Scheinkman and Xiong (2003) claim that trading volume reveals underlying difference of opinions, which is accompanied by bubbles in asset price when short selling is difficult.

*Dividend Premium.* Baker and Wurgler (2004a, b) define the dividend premium as the difference between the average market-to-book-value ratios of dividend payers and nonpayers. When dividends are at a premium, firms are more likely to pay them, and are less so when they are discounted. In other words, on the margin, when the prevailing demand for the stock market dividend premium is high, the propensity to pay dividend increases whereas with a low demand, the propensity to pay dividends decreases.

*Closed-End Fund Discount.* The closed-end fund discount (or occasionally premium) is the difference between the net asset value of a fund's actual security holdings and the fund's market price. Many authors, including Lee et al. (1991) and Neal and Wheatley (1998) consider the closed-end fund discounts to measure individual investor sentiment. They have argued that if closed-end funds are disproportionately held by retail investors, the average discount on closed-end equity funds may be a sentiment index, with the discount increasing when retail investors are bearish. Both these two papers suggest closed-end fund discount predict the size premium.

*Option Implied Volatility.* The Market Volatility Index (VIX), which measures the implied volatility of options on the Standard and Poor's 100 stock index, is often referred as "investor fear gauge" by practitioners. Whaley et al. (2008) define VIX as a measure of investor's certainty (or uncertainty) regarding volatility. It's about fear of unknown such that the higher the VIX is, the greater the fear.

*IPO First-Day Returns and IPO Volume.* The IPO market is often viewed to be sensitive to sentiment. Specifically, high first day return on IPOs is considered as a measure of investor enthusiasm, and the low idiosyncratic return on IPOs is often interpreted as a symptom of market timing. The underlying

demand for IPOs is also said to be extremely sensitive to investor sentiment. Furthermore, Average first-day returns display peaks and troughs which are highly correlated with IPO volume.

*Equity Issues over Total New Issues.* Baker and Wurgler (2000) find that high values of the equity share predict low stock market returns, and suggest that this pattern reflects firms shifting successfully between equity and debt to reduce the overall cost of capital. The authors argue that this pattern need not imply that individual firms or their managers can predict prices on the market as a whole. Rather, correlated mispricings across firms may lead to correlated managerial actions, which may then forecast correlated corrections of mispricings, that is, forecast market returns.

*Insider Trading.* Seyhun (1998) presents evidence on the ability of insider trading activity to predict stock return and reap significant profits. Corporate executives, board members and large shareholders have better information about the true value of their firms than outside investors. Thus, legalities aside, their personal portfolio decisions may also reveal their views about the mispricing of their firms. If sentiment leads to correlated mispricings across firms, insider trading patterns may contain a systematic sentiment component.

There are a few other economic variables that have been employed as proxies for sentiment in recent literature. Brown and cliff (2004) and Wang et al.(2006) outline and examine a number of sentiment indicators, such as ARMS index, put-call trading volume and open interest ratios, the percentage change in margin borrowing, the percentage change in short interest, and the ratio of odd-lot sales to purchases.

*ARMS Index.* The ARMS index on day  $t$  is equal to the number of advancing issues scaled by the trading volume (shares) of advancing issues divided by the number of declining issues scaled by the trading volume (shares) of declining issues. ARMS can be interpreted as the ratio of volume per declining issue to the volume in each advancing issue. If the index is greater than one, more trading is taking place in declining issues, whilst if it is less than one, more volume in advancing stocks outpaces the volume in each declining stocks. Its creator, Richard Arms (1989), argue that if the average volume in declining stocks far outweighs the average volume in rising stocks then the market is oversold and that this should be treated as bullish sign. Likewise, he argue that if the average volume in rising stocks far outweighs the average volume in falling stocks then the market is overbought and that this should be treated as a bearish sign.

*Put-Call Trading Volume.* The put-call trading volume ratio is a measure of market participants' sentiment derived from options and equals the trading volume of put options divided by the trading volume of call options. The ratio of CBOE equity put to call trading volume is widely viewed as a bearish indicator in US market. When market participants are bearish, they buy put options either to hedge their spot positions or to speculate bearishly. Therefore, when the trading volume of put options becomes large relative to the trading volume of call options, the sentiment goes up, and vice versa.

*Put-Call Open Interest ratios.* Wang et al. (2006) further introduce the approach of using the open interest of options instead of trading volume to calculate the put-call ratio. The ratio can be calculated on a daily basis using the day or on a weekly basis using the open interest of options at the end of the week. Wang et al. (2006) claim that this might be a preferred measure of sentiment as it may be argued that the open interest of options is the final picture of sentiment at the end of the day or the week and is therefore likely to have better predictive power for volatility in subsequent periods.

*Percentage Changes in Margin Borrowing.* This measure is frequently cited as a bullish indicator since it represents investors using borrowed money to invest.

*Percentage Changes in Short Interest.* The argument is made that the specialists are well-informed and relatively savvy investors, so when their short-selling becomes relatively large, the market is likely to decline. Hence, the percentage change in short interest is usually viewed as a bearish indicator.

*The Ratio of Odd-Lot Sales to Purchases.* Fosback (1993) suggests this ratio be a bearish measure.

### **2.1.2 Survey data as sentiment measures**

Another strand of recent research expands the direct measures of investor sentiment to consider aggregate market views regarding sentiment across investor types, including both institutional and individual investors. Brown and Cliff (2004) assume that the survey data conducted by the American Association of Individual Investors (AAII) and Investors Intelligence (II) are reasonable proxies for the true sentiment. They demonstrate that surveys measuring investor sentiment are related to other popular measures of investor sentiment and recent stock market returns. Brown and Cliff (2005) use survey data from Investors Intelligence (II) as a contrarian indicator. Lemmon and Portniaguina (2006) explore the time-series relationship between investor sentiment and the small-stock premium using consumer confidence conducted in the United States as a measure of investor optimism. One of the survey data is collected by the Conference Board [the Index of Consumer Confidence (CBIND)] and the other is independently conducted by the University of Michigan Survey Research Centre [the Index of Consumer Sentiment (ICS)]. Schmeling (2009) examines whether consumer confidence - as a proxy for individual investor sentiment-affects expected stock returns internationally in 18 industrialized countries.

### **2.1.3 Composite sentiment index**

The last measure is to construct a composite proxy index from the available economic variables. Prior research presents a number of proxies for sentiment to use as time-series conditioning variables. However, there are no definitive or uncontroversial measures. Brown and Cliff (2004) indicate that the survey data alone are most likely incomplete measures of sentiment. Conceptually, it is appealing to extract the common component (s) of the available economic series which might represent a cleaner measure of investor sentiment. In order to exploit as much information as possible, they combine the various sentiment measures, indirect and direct ones, and use two well-established methods to extract common features of the data: the Kalman filter and principal component analysis (PCA). Likewise, Baker and Wurgler (2006, 2007) argue that data availability narrows the list of sentiment measures considerably. They suggest a composite index of sentiment which is based on the common variation in the available underlying proxies for sentiment. They propose the principal component methodology, like Brown and Cliff (2004), to define a sentiment index, which captures the common component in the underlying economic variables.

## **2.2 Empirical Studies of Investor Sentiment**

Behavioural finance argues that the arbitrage will be limited in some senses, and investors might be affected by psychology biases, noise, or sentiment. As Baker and Wurgler (2007) summarize, researchers in behavioural finance have therefore been working to modify the standard model with an alternative model built on two basic assumptions.

The first assumption, put forward by DeLong et al. (1990), is that investors are subject to sentiment. Investor sentiment, defined broadly, is a belief about future cash flows and investment risks that is not justified by the facts or economic theory. The remarkable work of DeLong et al. (1990) models the influence of noise trading on equilibrium prices, in which noise traders act in concert on non-fundamental signals. The simultaneous actions introduce a systematic risk that is priced. In their model, the deviations in price from fundamental value induced by changes in investor sentiment are unpredictable. Arbitrageurs betting against mispricing run the risk that investor sentiment becomes more extreme and prices vary even further away from fundamental values, at least in short run. The possibility of loss and the arbitrageurs' risk aversion reduce the size of positions they are willing to take. Consequently, arbitrage fails to completely eliminate mispricing and investor sentiment affects asset prices in equilibrium. The model of DeLong et al. (1990) predicts that the direction and magnitude of changes in noise trader sentiment are relevant in asset pricing.

The second assumption, emphasized by Shleifer and Vishny (1997), is that betting against sentimental investors is costly and risky and hence there are limits to arbitrage. Rational arbitrageurs are not as aggressive in bringing prices to fundamentals as the standard model would suggest.



A pioneer and well-known set of studies of sentiment and aggregate stock returns appears in the mid-1980s. In these researches, the role of sentiment was left implicit and the statistical evidence was not very strong. Nowadays, the systematic role of investor sentiment has been suggested by many empirical and theoretical studies. One set of studies focuses on demonstrating how sentiment predicts future returns in stock market.

Neal and Wheatley (1998) utilize three popular measures of investor sentiment: closed-end fund discount, net mutual fund redemptions and the ratio of odd-lot sales to purchases as sentiment measures. They exploit the forecast power of these three measures and show that the first two measures forecast the size premium, but little evidence that the odd-lot ratio predicts returns. Fisher and Statman (2000) report a negative relationship between investor sentiment and future stock returns. Baker and Wurgler (2000) use the share of equity issues in total new issues, that is equity and debt issues to proxy for investor sentiment. They demonstrate that this measure predicts significantly negative market returns which cannot be explained by efficient market hypothesis. Brown and Cliff (2005) use a direct survey measure of investor sentiment to forecast market returns over the following 1-3 years. The estimation of coefficient on investor sentiment is significantly positive which suggests the market is overvalued during periods of optimism. They further show that sentiment is positively related to changes in market valuations, in the error correction version of the cointegrating regression. Corredor et al. (2013) analyse the forecast performance of investor sentiment in four European stock markets: France, Germany, Spain and the UK. They claim that sentiment has a significant effect on returns, though there is dispersion in intensity across countries.

The second set of studies exploits the possibility of a causal relationship between market returns and investor sentiment or changes in investor sentiment. The Granger causality tests of Brown and Cliff (2004) failed to reject the null hypothesis of no predictability in returns to sentiment for small and large stocks. On the other hand, changes in investor sentiment appear significantly negatively impact on subsequent market returns of small but not of large stocks. By estimating bivariate VAR models, Wang et al. (2006) also take a look at the causality between sentiment and market returns in both directions. They confirm the result of Brown and Cliff (2004) that sentiment is not causal variables of market returns. On the contrary, sentiment is granger caused by market returns. Schmeling (2009) reports that there is a two-way causality between the sentiment measures and stock returns.

Recent researches shed more light on the cross-sectional effects of investor sentiment. Brown and Cliff (2005) use the 25 Fama and French portfolios, together with 5 portfolios sorted from univariate size sorts, 5 portfolios from book-to-market sorts and the overall market portfolio. They show that for large firms or low book-to-market firms, sentiment is a significant predictor of future returns at the 1-, 2-, and 3-year horizon. Baker and Wurgler (2006) examine how investor sentiment impacts the cross-section of stock returns. They form equal-weighted deciles portfolios based on several firm characteristics, and look for patterns in the average returns across deciles conditioning on the beginning-of-period level of sentiment. They demonstrate that the subsequent returns are relatively low for small stocks, young stocks, high volatility stocks, distressed stocks, unprofitable stocks, stocks with no dividend payment, and stocks experiencing extreme growth, when sentiment measures are high, and vice versa. Berger and Turtle (2012) report that investor sentiment sensitivities increase directly with the opacity of firms in the cross-section. They display an inverse relation between *ex ante* investor sentiment and the marginal performance of opaque stocks. The performance of translucent stocks, on the contrary, exhibits relatively little variability across levels of sentiment.

### **3. Methodology**

#### **3.1 Principal Component Analysis (PCA)**

Principal component analysis (PCA) is a statistical procedure that uses orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly correlated variables called principal components. The transformation is defined in such a way that the first

principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to the preceding components. PCA is sensitive to the relative scaling of the original variables.

### 3.2 EGARCH Component Model

Nelson (1991) suggests the EGARCH model to improve the performance of GARCH model. Engle and Lee (1999) introduce a component GARCH model where the conditional variance is decomposed into transitory and permanent components. Many studies find that two-component volatility model is superior to one-component specification in explaining equity market volatility and the log-normal model of EGARCH performs better than square-root or affine volatility specifications. Appealed to the merits of the component GARCH and the EGARCH models, Adrian and Rosenberg (2008) incorporate the features of these two models and specify the dynamics of the market return in excess of risk-free rate  $R_t^M$  and the conditional volatility  $\sqrt{h_t}$  as:

$$\text{Market return: } R_{t+1}^M = \mu_t^M + \sqrt{h_t} z_{t+1} \quad (1a)$$

$$\text{Market volatility: } \ln \sqrt{h_t} = l_t + s_t \quad (1b)$$

$$\text{Short-run component: } s_{t+1} = \theta_4 s_t + \theta_5 z_{t+1} + \theta_6 (|z_{t+1}| - \sqrt{2/\pi}) \quad (1c)$$

$$\text{Long-run component: } l_{t+1} = \theta_7 + \theta_8 l_t + \theta_9 z_{t+1} + \theta_{10} (|z_{t+1}| - \sqrt{2/\pi}) \quad (1d)$$

In equation (1a),  $z_t$  is a normal i.i.d. error term with zero expectation and unit variance, and  $\mu_t^M$  is the one-period expected excess return. The log-volatility in equation (1b) is the sum of two components  $l_t$  and  $s_t$ . Each component is an AR(1) processes with its own rate of mean reversion. Without loss of generality, let  $l_t$  be the slowly mean-reverting, long-run component and  $s_t$  be the quickly mean-reverting, short-run component ( $\theta_4 < \theta_8$ ). The unconditional mean of  $s_t$  is normalized to zero.

The terms  $|z_{t+1}| - \sqrt{2/\pi}$  in equations (1c) and (1d) are the shocks to the volatility components. Their expected values are equal to zero, given the normality of  $z_t$ . For these error terms, equal sized positive or negative innovations results in the same volatility change, although the magnitude can be different for the short- and long-run components ( $\theta_6$  and  $\theta_{10}$ ). The asymmetric effect of returns on volatility is allowed by including the market innovation in equations (1c) and (1d) with corresponding coefficients  $\theta_5$  and  $\theta_9$ .

The market model defined by equations (1a) – (1d) converges to a continuous-time, two-factor stochastic volatility process. An advantage of this specification is that it can be estimated in discrete time via maximum likelihood. The daily log-likelihood function is:

$$f_t(\theta; s_t, l_t | R_t^M) = -\frac{1}{2} \ln(2\pi) - (s_{t-1} + l_{t-1}) - \frac{(R_t^M - \theta_1 - \theta_2 s_{t-1} - \theta_3 l_{t-1})^2}{2h_{t-1}},$$

where  $t=1, \dots, T$  is the daily time index,  $T$  is the total number of daily observations, and  $R_t^M$  is the daily market excess return.

### 3.3 Sentiment-Augmented EGARCH Component Model

Delong et al. (1990) present a notable paper that, in a simple overlapping-generation model of an asset market, irrational noise traders with erroneous stochastic beliefs affect prices and actually enjoy higher expected returns. In this model, there are two types of investors: rational investors and noise traders. Each period, rational investors and noise traders trade assets based on their respective beliefs of expected return. There are two crucial assumptions of this model. First, the authors assume that the investment horizons of rational investors are short, so that they care about the interim resale prices of

the assets they hold, not just the present values of future dividends. Second, this model assumes that noise traders' sentiment is stochastic and cannot be perfectly predicted by rational investors.

The optimism or pessimism of noise trader creates a risk in the price of the asset that causes transitory divergences between price and intrinsic value, even in the absence of fundamental risk. Rational investors run the risk that sentiment will become more extreme and prices deviate further away from fundamentals. The risk aversion and pressure in the fund of rational investors limits their willingness of taking extremely volatile positions to bring the prices back to intrinsic values. Therefore, noise trading limits the effectiveness of arbitrage and rational arbitrage fails to eliminate mispricing. Furthermore, sentiment induces trading that occurs contemporaneously across many assets in the markets. This introduces additional variability in returns which is a non-diversifiable systematic risk that is priced in equilibrium.

In Delong et al. (1990), the impact of noise trading on the returns of risk assets is a result of interaction of four effects. They summarize that hold more effect and create space effect tend to increase noise traders' relative expected returns. The Friedman effect and price pressure effect tend to lower noise traders' relative expected returns. In particular, the hold more and price pressure effect affect mean returns directly, since they are related to the *direction* of shifts in noise trader sentiment. Meanwhile, Friedman effect and create space effect is related to the *magnitude* of the shifts in noise trader sentiment. Hence, the influence is indirectly on mean returns through changes in noise traders' misperceptions of the asset's risk.

In accordance with the research of Delong et al. (1990), Lee et al. (2002) propose a sentiment-augmented GARCH-in-mean model to capture the four effects of noise trading. Contemporaneous shifts in investor sentiment are included in the mean equation and lagged shifts in the magnitude of investor sentiment are embodied in the conditional volatility equation. Their model takes the following form:

$$R_t - R_{ft} = \alpha_0 + \alpha_1 h_t + \alpha_2 Jan_t + \alpha_3 Oct_t + \alpha_4 \Delta S_t + \varepsilon_t$$

$$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \varepsilon_{t-1}^2 I_{t-1} + \beta_3 h_{t-1} + \beta_4 R_{ft} + \beta_5 (\Delta S_{t-1})^2 D_{t-1} + \beta_6 (\Delta S_{t-1})^2 (1 - D_{t-1})$$

where  $R_t$  is the weekly return on a market index,  $R_{ft}$  is the risk free rate, and  $\Delta S_t$  is a measure of noise trader risk. Lee et al. (2002) apply two alternative measures of the noise trader risk. One is the changes in Investor's Intelligence (II) sentiment index and the other one is the percentage changes in II sentiment index. Furthermore,  $\varepsilon_t \sim N(0, h_t)$  and  $I_{t-1}$  and  $D_{t-1}$  are dummy variables where (i)  $I_{t-1} = 0$  if  $\varepsilon_{t-1} \leq 0$  and  $I_{t-1} = 1$  if  $\varepsilon_{t-1} > 0$ ; and (ii)  $D_{t-1} = 0$  if  $\Delta S_{t-1} \leq 0$  and  $D_{t-1} = 1$  if  $\Delta S_{t-1} > 0$ .

Arik (2012) also examines the impact of sentiment on stock returns in the framework of GARCH process. The author augments his sentiment measures to the mean equation of the GARCH specification. Inspired by the model specifications of Lee et al. (2002) and Arik (2002), we intend to include investor sentiment in Adrian and Rosenberg's EGARCH component model. Levels of Investor sentiment or changes in investor sentiment are introduced into the mean and variance equations. Besides the base model of EGARCH component, we test three alternative sets of empirical models.

-Model 1: Benchmark model - EGARCH component model of Adrian and Rosenberg (2008), as in equations (1a) – (1d) above

-Model 2: Investor sentiment in mean and variance equation in accordance with GARCH-in-mean model of Lee et al. (2002)

$$R_{t+1}^M = \theta_1 + \theta_2 s_t + \theta_3 l_t + \theta_{11} SENT_{t+1} + \sqrt{h_t} z_{t+1}$$

$$s_{t+1} = \theta_4 s_t + \theta_5 z_{t+1} + \theta_6 (|z_{t+1}| - \sqrt{2/\pi}) + \theta_{12} SENT_t^2 D_t + \theta_{13} SENT_t^2 (1 - D_t)$$

$$l_{t+1} = \theta_7 + \theta_8 l_t + \theta_9 z_{t+1} + \theta_{10} (|z_{t+1}| - \sqrt{2/\pi}) + \theta_{14} SENT_t^2 D_t + \theta_{15} SENT_t^2 (1 - D_t)$$

where  $D_t = 0$  if  $S_t \leq 0$  and  $D_t = 1$  if  $S_t > 0$ . Since the mean of the sentiment index ( $SENT_t$ ) is close to zero, which will be shown later, the variance of the sentiment index can be approximated by  $SENT_t^2$ . Furthermore, we examine both the levels and changes in investor sentiment. Hence, in a parallel regression, all the levels of sentiment are replaced by the changes in sentiment ( $\Delta SENT_t$ ).  $D_t = 0$  if  $\Delta S_t \leq 0$  and  $D_t = 1$  if  $\Delta S_t > 0$ . Similarly, the levels together with changes in investor sentiment are applied respectively to Model 3 and 4. Coefficients of  $\theta_{12}$ ,  $\theta_{13}$ ,  $\theta_{14}$  and  $\theta_{15}$  describe the asymmetric effects of sentiment on short- and long-run volatilities.

It is worth pointing out that Lee et al. (2002) introduce the dummy variable  $I_{t-1}$  in their model. The intuition of the inclusion of  $I_{t-1}$  is to encompass the well-known volatility asymmetry or leverage effect in financial market. The argument is that investors form their expectations of conditional volatility which may perceive positive and negative shocks differently. If  $\beta_2$  is negative as expected, a negative shock is more likely to induce a larger upward revision of volatility than a positive shock of the same magnitude. However, this dummy is not included in the EGARCH component model since the parameters of  $\theta_5$  and  $\theta_9$  of the model already allow for the leverage effect.

-Model 3: Investment sentiment only in mean equation which is consistent with Arik (2012)

$$R_{t+1}^M = \theta_1 + \theta_2 s_t + \theta_3 l_t + \theta_{11} SENT_{t+1} + \sqrt{h_t} z_{t+1}$$

$$s_{t+1} = \theta_4 s_t + \theta_5 z_{t+1} + \theta_6 (|z_{t+1}| - \sqrt{2/\pi})$$

$$l_{t+1} = \theta_7 + \theta_8 l_t + \theta_9 z_{t+1} + \theta_{10} (|z_{t+1}| - \sqrt{2/\pi})$$

This model assumes that market sentiment affects the contemporaneous stock returns and has no direct effect on volatility components.

-Model 4: Investor sentiment in mean equation and variance equation which is different from the model of Lee et al. (2002)

$$R_{t+1}^M = \theta_1 + \theta_2 s_t + \theta_3 l_t + \theta_{11} SENT_{t+1} + \sqrt{h_t} z_{t+1}$$

$$s_{t+1} = \theta_4 s_t + \theta_5 z_{t+1} + \theta_6 (|z_{t+1}| - \sqrt{2/\pi}) + \theta_{16} s_t D_t + \theta_{17} s_t (1 - D_t) + \theta_{18} z_{t+1} D_{t+1} \\ + \theta_{19} z_{t+1} (1 - D_{t+1})$$

$$l_{t+1} = \theta_7 + \theta_8 l_t + \theta_9 z_{t+1} + \theta_{10} (|z_{t+1}| - \sqrt{2/\pi}) + \theta_{20} l_t D_t + \theta_{21} l_t (1 - D_t) + \theta_{22} z_{t+1} D_{t+1} \\ + \theta_{23} z_{t+1} (1 - D_{t+1})$$

In model 2 in accordance with the framework of Lee et al. (2002), sentiment enters the models as the intercept dummy and the sentiment directly affects the level of short- and long-run volatilities. In Model 4, sentiment dummies act as the slope dummies, and sentiment influences the short- and long-run volatilities through their impacts on market shocks ( $Z_{t+1}$ ) and lagged short- and long-run volatilities ( $s_t$  and  $l_t$ ). Through the dummy variables  $D_t$  and  $(1 - D_t)$ , both the direction and magnitude of investor sentiment can have an asymmetric impact on conditional variance and market returns.

The short- and long-run sentiment-affected volatility components are obtained from the time series regressions. After that, the Fama-Macbeth two-stage regressions are employed to investigate the cross-sectional pricing abilities of short- and long-run volatility.

### 3.4 ICAPM Framework.

Ang et al. (2006), Petkova (2006) and Da and Schaumburg (2011) establish systematic theoretical motivations of pricing market volatility and are presented as follows. When investment opportunities vary over time, the multifactor models of Merton (1973) and Ross (1976) show that risk premia are associated with the conditional covariance between asset returns and innovations in state variables that describe the time-variation of the investment opportunities. And hence, covariance with these innovations will therefore be priced. In Campbell's (1993, 1996) ICAPM framework, investors care about risks both from the market return and from changes in forecasts of future market returns. For an investor that is more risk-averse than log utility, assets that covary positively with good news about future market expected returns enjoy higher average returns. These assets reduce a consumer's ability to hedge against deterioration in investment opportunities and a compensation is required. The intuition from Campbell's model is that risk-averse investors want to hedge against variations in aggregate volatility because volatility positively affects future expected market returns, as Merton (1973).

$$E(R) \approx rra \text{cov}(R, R^M) + \lambda_2 \text{cov}(R, \Delta v) \quad (1)$$

However, as Ang et al. (2006) point out, in Campbell's setup, there is no direct role for fluctuations in market volatility to affect the expected returns of assets because Campbell's model is premised on homoscedastic consumption. Chen (2002) extends Campbell's model to a heteroskedastic environment and allows for time-varying covariances and stochastic market volatility. Chen shows that risk-averse investors tend to directly hedge against changes in future market volatility. In Chen (2002)'s model, an asset's expected return depends on risk from the market return, changes in forecasts of future market returns, and changes in forecasts of future market volatilities. For an investor more risk averse than log utility, Chen (2002) demonstrates that an asset shall have a lower expected return if its return positively covaries with a variable that forecasts higher future market volatilities. This effect arises because risk-averse investors reduce current consumption to increase precautionary savings in the presence of increased uncertainty about market returns.

Motivated by these multifactor models, we express market volatility risk explicitly in equation (1),

$$E_t(R_{t+1}) \approx rra \text{cov}(R_{t+1}, R_{t+1}^M) + \lambda_2 \text{cov}(R_{t+1}, \Delta v_{t+1}) + \sum_{k=1}^K \lambda_k \text{cov}(R_{t+1}, f_{k,t+1}) \quad (2)$$

Where  $f_k$  represent other factors other than aggregate volatility that induce changes in the investment opportunity set.

Recent empirical study concentrates on how the volatility and other factors are priced in the cross-section of stock returns. For the convenience of empirical application, the above model can be written in terms of factor innovations. Suppose  $R_{t+1}^m - \gamma_{m,t}$  represents innovation in the market return,  $v_{t+1} - \gamma_{v,t}$  represents the innovation in the factor reflecting aggregate volatility risk, and innovations to the other factors are represented by  $f_{k,t+1} - \gamma_{k,t}$ . A true conditional multifactor representation of expected returns in the cross-section would take the following form:

$$R_{t+1}^i = a_t^i + \beta_{m,t}^i (R_{t+1}^M - \gamma_{m,t}) + \beta_{v,t}^i (v_{t+1} - \gamma_{v,t}) + \sum_{k=1}^K \beta_{k,t}^i (f_{k,t+1} - \gamma_{k,t}) \quad (3)$$

Where  $R_{t+1}^i$  is the excess return on stock  $i$ ,  $\beta_{m,t}^i$  is the loading on the excess market return,  $\beta_{v,t}^i$  is the asset's sensitivity to market volatility risk, and the  $\beta_{k,t}^i$  coefficients for  $k=1, \dots, K$  represent loading on other risk factors. In the full conditional setting in equation (3), factor loadings, conditional means of factors, and factor premia potentially vary over time. The conditional mean of the market and aggregate volatility are denoted by  $\gamma_{m,t}$  and  $\gamma_{v,t}$ , respectively, while the conditional means of the other factors are denoted by  $\gamma_{k,t}$ . In equilibrium, the conditional mean of stock  $i$  is given by

$$a^i = E(R^i) = \beta_m^i \lambda_m + \beta_v^i \lambda_v + \sum_{k=1}^K \beta_k^i \lambda_k \quad (4)$$

where  $\lambda_{m,t}$  is the price of risk of the market factor,  $\lambda_{v,t}$  is the price of aggregate volatility risk, and the  $\lambda_{k,t}$  are the prices of risk of the other factors. Note that only if a factor is traded is the conditional mean of a factor equal to its conditional price of risk, that is  $\lambda_m = \gamma_{m,t}$ ,  $\lambda_v = \gamma_{v,t}$  and  $\lambda_k = \gamma_{k,t}$ .

If the short- and long-run volatility components are also asset pricing factor, in the spirit of the ICAPM, the equilibrium pricing kernel thus depends on both short- and long-run volatility components as well as the excess market returns. Denote returns on asset  $i$  in excess of the risk free rate by  $R_t^i$ . The equilibrium expected return for asset  $i$  is :

$$E_t(R_{t+1}^i) = \lambda_1 \text{cov}(R_{t+1}^i, R_{t+1}^M) + \lambda_s \text{cov}(R_{t+1}^i, s_{t+1}) + \lambda_l \text{cov}(R_{t+1}^i, l_{t+1}) \quad (5)$$

where  $\lambda_1$  is the coefficient of relative risk aversion, and  $\lambda_s$  and  $\lambda_l$  are proportional to changes in the marginal utility of wealth due to changes in the state variable  $s_t$  and  $l_t$ .

Equation (5) shows that expected returns depend on three risk premia. The first risk premium arises from the covariance of the asset return with the excess market return, multiplied by relative risk aversion  $\lambda_1$ . This is the risk-return tradeoff in a static CAPM. The second and third risk premia depend on the covariance of the asset return with the innovations in the short- and long-run factors. These are scaled by the impact of changes in the volatility factors on marginal utility of wealth,  $\lambda_s$  and  $\lambda_l$ .

The two-stage Fama-MacBeth regression estimates the premium rewarded to particular risk factor exposure by the market, the short- and long-run volatilities in our case. This methodology provides a particularly robust way to test the theoretical model empirically. However, betas are estimated with error in the first-stage time-series regression, an errors-in-variables (EIV) problem is introduced in the second-stage cross-sectional regression. Shanken (1992) finds that the Fama-MacBeth two-pass procedure for computing standard errors fails to reflect measurement error in the betas and overstates the precision of the estimates of factor premium. However, Jagannathan and Wang (1998) argue that if the error terms are heteroskedastic, then the Fama-MacBeth procedure does not necessarily result in smaller standard errors of the risk premium estimated. Nevertheless, we use the correction procedure proposed by Jagannathan and Wang (1998) to account for the errors-in-variables problem. Hence, the  $p$  values reported in this paper are computed from the corresponding  $t$ -values which are adjusted to account for the first-step estimation error and potential heteroskedasticity and autocorrelation using the Newey-West (1987) correction with 12 lags.

## 4. Data

### 4.1 Data Summation and Description of Sentiment Proxies

In this paper, the EGARCH-component volatility model is estimated using monthly excess returns. FTSE All Share Index with its dividend yield is used as the proxy for the market return,  $r^M$ , and one month return on Treasury Bills for the risk free rate,  $r^f$ . The data covers from March 1987 to December 2012 which is in line with the data range of sentiment index and are collected from LSPD (London Stock Price Database) and datastream. For the cross-sectional price test of the ICAPM, we apply the Fama and French 25 portfolios sorted on size and book-to-market (B/M) equity. The portfolio returns and monthly factors are taken directly from Gregory et al. (2009) website.

Data availability narrows down the sentiment measures considerably. Existing literature suggests a variety of approaches of proxies for sentiment. However, there are no definitive or uncontroversial measures. Hence, we construct the composite index of sentiment which is based on the common variation in the available underline proxies for sentiment, in accordance with Baker and Wurgler (2004, 2005) and Brown and Cliff (2004). Baker et al. (2012) study UK stock market as part of their global market and the volatility premium, number and first-day returns of IPOs and turnover by value are employed to construct UK sentiment index. Corredor et al. (2013) use consumer confidence, turnover and volatility premium to measure UK sentiment as a part of European stock market. In our research, the individual proxies include share turnover by value on LSE, share turnover by volume on LSE, the number and average first-day returns on IPOs, and the consumer confidence. The first four variables are the same as those used in Baker and Wurgler index, and the aim of the last variable is to compensate the lack of closed-end fund discounts. The sentiment proxies are measured monthly from October 1986 to December 2012. However, the beginning 5 data are omitted due to data process procedure and hence the sample period starts from March 1987 eventually.

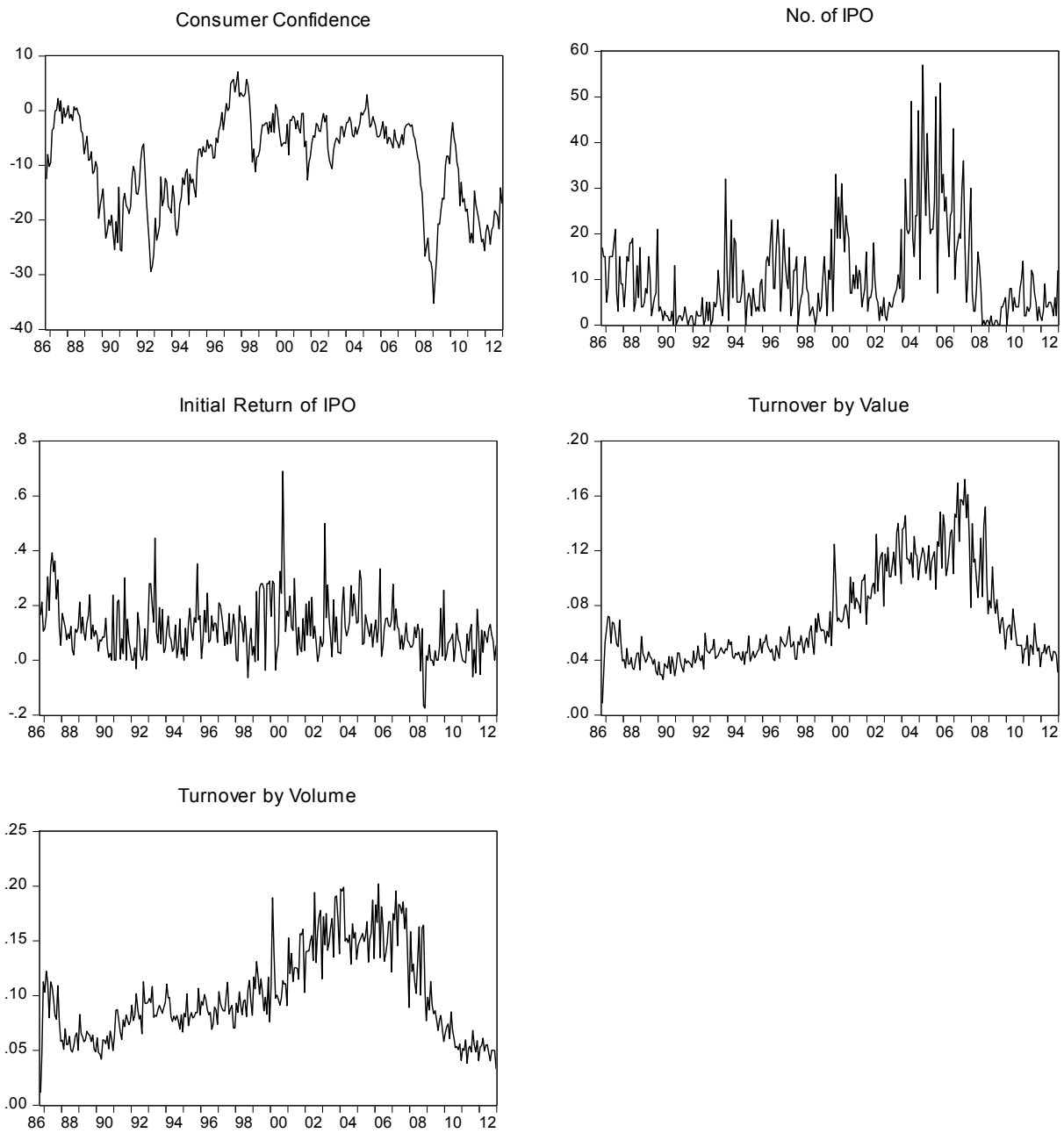
Market share turnover can be defined both by trading volume and trading values. Market turnover by value is the total sterling volume over the month divided by total capitalization of London Stock Exchange (LSE). Market turnover by volume is the number of total share traded on LSE over the month divided by the number of shares listed on the exchange. The daily trading volume, trading values, total capitalization of LSE and total share trade on LSE are aggregated within each month to get the monthly data respectively. The data are extracted from Datastream Global Equity Indices, which are calculated on a representative list of stocks for each market. The number of stocks for each market is determined by the size of the market and the sample covers a minimum of 75% - 80% of total market capitalisation.

Numbers of IPOs within each month, denoted by *NO\_IPO*, are taken from two sources. One is the New Issue and IPO Summary spreadsheet from the London Stock Market website which contains IPO summary since June 1995. The other source is the London Share Price Database (LSPD). The population of IPOs are identified using the LSPD "birth maker" and investment trust offerings are excluded since they are classified as financial institution offerings.

First-day return of IPOs is defined as the difference between initial trading price and offer price divided by offer price of the IPO stock. The offer prices are obtained from Thomson One Bank, LSE new issue and IPO Summary, together with the LSPD. The initial trading prices are collected from Datastream as the first day open price. The equal-weighted average first-day returns are then computed and denoted by *RE\_IPO*.

Consumer confidence, denoted by *CC*, is a business survey data reported by the European Commission, Economic and Financial Affairs. UK respondents express their economic or financial expectations over the next 12 months in the following areas: the general economic situation, unemployment rate, personal household financial position and personal savings.

Figure 1: Figures of the five raw sentiment measures.



## 4.2 Construction of the Level Sentiment Index

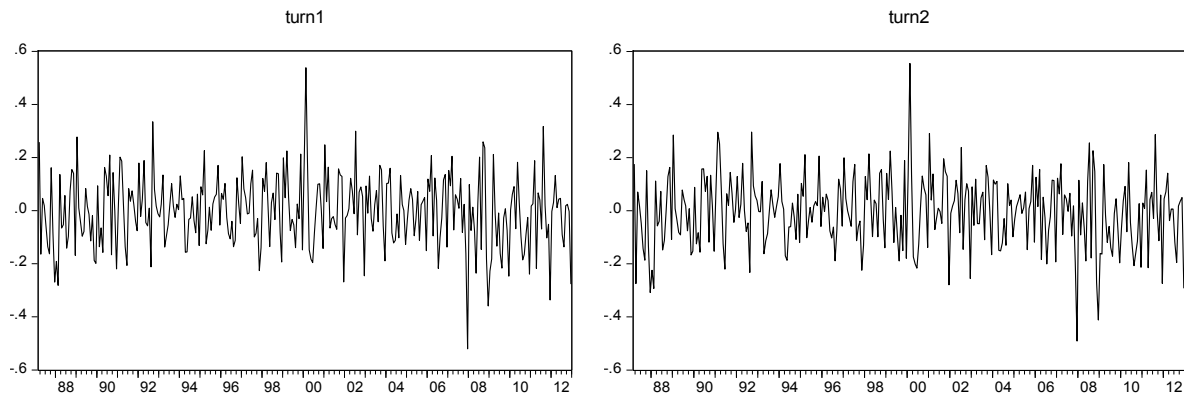
Table 1: Summary statistics of investor sentiment measures

	Mean	Std. Dev.	Skewness	Kurtosis	Unit root test	$I^{st}$ lag autocorrelation
CC	-9.314	8.575	-0.562	2.412	$I(0)^*$	0.941***
NO_IPO	10.263	10.025	1.734	6.689	$I(0)^{**}$	0.636***
RE_IPO	0.114	0.102	1.157	6.555	$I(0)^{***}$	0.322***
Turnover_value	0.070	0.034	0.976	2.877	$I(1)$	0.895***
Turnover_volume	0.101	0.041	0.648	2.428	$I(1)$	0.854***



The statistics of the five sentiment measures are presented in table 1. All these measures display a skewed and leptokurtic pattern and are rejected by the null hypothesis of normality. The unit root tests detect that there is a time trend in both turnover by value and by volume, so we use the log of turnovers and detrend them with an up-to-five-month moving average. The detrended turnovers by value and by volume are defined by *TURN1* and *TURN2*, respectively. After detrending, these two time series become  $I(0)$  process. The autocorrelation tests show that the five time series suffer from high autocorrelations, including *TURN1* (with the first lag correlation of 0.103) and *TURN2* (with the first lag correlation of 0.089). The log transformation is applied to the Number and first-day returns of IPOs and the transformed variables are denoted as *NIPO* and *RIPO*.

Figure 2: Detrended market turnover by value (*TURN1*) and turnover by volume (*TURN2*).



Furthermore, sentiment measures reflect economic fundamentals to some extent and hence are likely to contain a sentiment component as well as idiosyncratic components that are not related to sentiment. To control and remove the information about rational factors that our sentiment measures may contain, we orthogonalize each proxy to six available macro series, following Brown and Cliff (2005) and Baker and Wurgler (2006, 2007, 2012). The macro variables suggested by Brown and Cliff (2005) include 1-month Treasury bill return, difference in monthly returns on 3-month and 1-month T-bills, term spread, default spread, dividend yield and rate of inflation. The control variables chosen by Baker and Wurgler(2006, 2007, 2012) are consumption growth, industrial production growth, employment growth, the short-term rate, and the term premium. According to the data availability, we take 1-month Treasury bill return which is the short-term return, the difference in monthly return on 3- and 1-month T-bills, the term spread as measured by the spread in yields on the 10-year Gilt and the 3-month T-bills, inflation rate, industrial production growth, and consumption growth as our control variables.

Table 2 shows the correlations of each of the raw proxies and its own variable orthogonalized by the above mentioned macro variables. It turns out that the macro series explain comparatively little of the variation in the sentiment measures, except for consumer confidence. The correlation between the raw and orthogonalized proxies is 88.67% on average across the five measures. The macro control variables that contain contemporaneous and forward-looking information about economic fundamentals are largely unrelated to the investor sentiment proxies. However, Baker and Wurgler (2012) emphasize that it is impossible to rule out that there might be an as-yet undiscovered risk factor driving all of the various relationships between the expected returns and the sentiment measures.

Table 2: Correlations of the raw proxies and their orthogonalized results

	Raw and orthogonalized <i>CC</i>	Raw and orthogonalized <i>TURN1</i>	Raw and orthogonalized <i>TURN2</i>	Raw and orthogonalized <i>NIPO</i>	Raw and orthogonalized <i>RIPO</i>
Correlation	0.6628	0.9807	0.9772	0.8645	0.9483

Furthermore, Baker and Wurgler (2006) point out that the sentiment measures might exhibit lead-lag relationships. Some economic variables may reflect a given shift in sentiment earlier than others, and hence we need to determine the relative timing of the variables in forming a composite index. Baker and Wurgler (2006) assert that proxies that are based directly on investor demand or investor behaviour can be expected to one period ahead of proxies that involve firm supply responses. Consequently, turnover and return of IPOs might be one period before the IPO volume. Perhaps sentiment is partly behind the high initial-day returns, and high sentiment attracts more IPO volume with a lag. Similarly, high sentiment triggers more trading volume, and leads to higher turnover, both in volume and value.

To encompass the issue of relative timing of the variables, the six proxies together with their lags are included in the principal component analysis (PCA), which will give us a first-stage index with 10 loadings, one for each of the current and lagged measures. The correlation between the first-stage index and the current and lagged variables are computed, and each respective proxy's lead and lag, whichever gives higher correlation with the first-stage index, will be kept for the PCA construction of the final sentiment index. It's worth pointing out that the principal component analysis (PCA) is sensitive to the scaling of the variables. Hence, each orthogonalized variable should be normalized to have zero mean and unit variance before applying the PCA procedure.

The procedure gives a parsimonious index

$$SENT_t = 0.202 * CC_{t-1} + 0.670 * TURN1_{t-1} + 0.669 * TURN2_{t-1} + 0.210 * NIPO_t + 0.135 * RIPO_{t-1} \quad (6)$$

where each of the proxies has been orthogonalized by the mentioned macro variables. The fraction of variance explained by the first principal component is 52.13%, which suggests that this composite factor captures much of the common variation. Meanwhile, the correlation between the sentiment in equation (6) and the 10-term first-stage index is 95.87%. Hence, we may conclude that there is little information loss in dropping the five terms with other time subscripts. The composite sentiment index already has a zero mean and is then standardized to have unit variance.

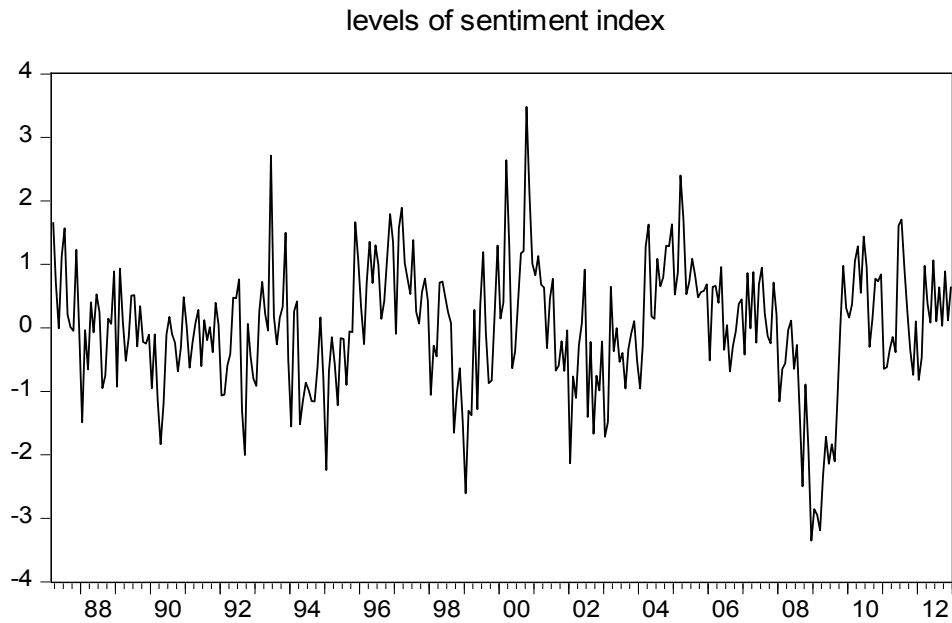
The SENTIMENT index has two appealing properties. First, as expected, all the five sentiment measures are positively associated with sentiment levels. Second, each individual proxy enters with the expected timing, such that price and investor behaviour variables (consumer confidence, market turnover, returns of IPOs) lead firm supply variables (IPOs volume).

The changes in the index of sentiment levels are obtained by taking the first-order difference.

Figure 3: Levels and changes in sentiment index

Panel A: Index of sentiment levels

$$SENT_t = 0.202 * CC_{t-1} + 0.670 * TURN1_{t-1} + 0.669 * TURN2_{t-1} + 0.210 * NIPO_t + 0.135 * RIPO_{t-1} \quad (6)$$



Panel B: Index of sentiment changes

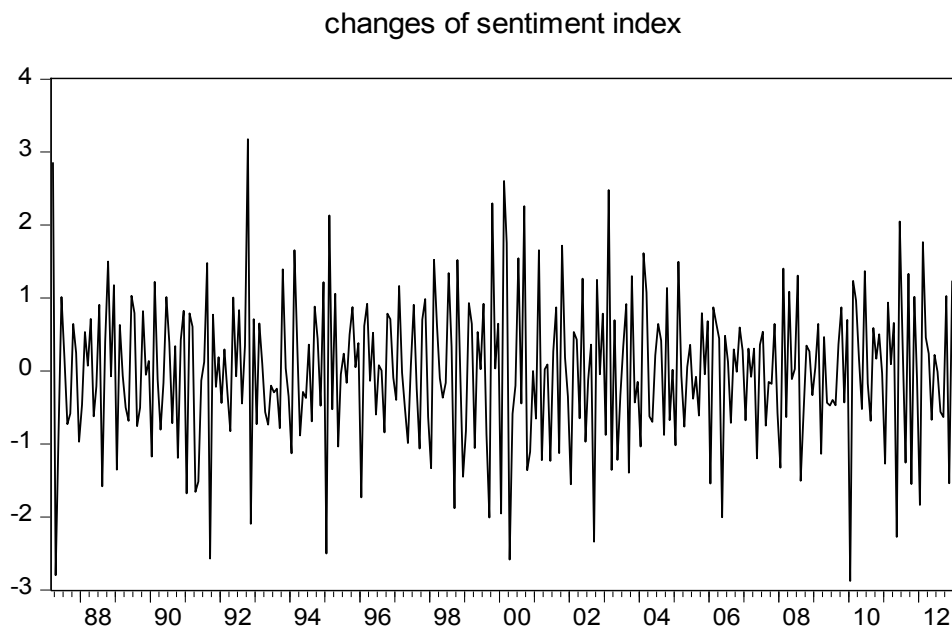


Figure 3 depicts the sentiment levels and changes from March 1987 to the end of 2012. There are two obvious patterns that coincide with the history of UK stock markets. First, sentiment peaks in 2000 which may result from end of internet bubble in early 2000. Second, sentiment crashes in 2008 and recovers thereafter which may depict the financial crisis from 2007 to 2010.

#### 4.4 Granger Causality Test of Market Returns and Sentiment Indices

Recognizing that sentiment itself is affected by recent market behaviour, we seek to determine the direction of any causal relationships between market return and sentiment. Results from simple bivariate (stock returns and investor sentiment) Granger-Causality tests are shown in Table 3. As can be inferred, the hypothesis that market return and sentiment are not causally related variables cannot be significantly rejected for lags of one and two. This suggests the time-series dependencies between our sentiment measures and market excess returns in short run are weak. In contrast, for lags of six or

twelve, market return is more likely to Granger cause the level of sentiment, and changes in sentiment are more likely to Granger cause excess returns.

Table 3: Granger Causality tests of market excess return and investor sentiment for 1-, 2-, 6 and 6-month lag

Null hypothesis		Lag 1	Lag 2	Lag 6	Lag 12
Market excess return does not Granger cause levels of sentiment	F-stat.	0.480	0.837	3.492	2.293
	<i>p-value</i>	<i>0.489</i>	<i>0.434</i>	<i>0.002</i>	<i>0.009</i>
Levels of sentiment does not Granger cause market excess return	F-stat.	1.092	1.264	1.552	2.182
	<i>p-value</i>	<i>0.297</i>	<i>0.284</i>	<i>0.161</i>	<i>0.013</i>
Market excess return does not Granger cause changes in sentiment	F-stat.	0.424	0.481	1.674	1.229
	<i>p-value</i>	<i>0.515</i>	<i>0.619</i>	<i>0.127</i>	<i>0.262</i>
Changes in sentiment does not Granger cause market excess return	F-stat.	0.012	0.009	3.397	1.189
	<i>p-value</i>	<i>0.915</i>	<i>0.991</i>	<i>0.003</i>	<i>0.036</i>

## 5. Empirical Results

### 5.1 Time-Series Estimations

Table 4: Estimation results of the four models in section 3.3

	Levels of Sentiment					Changes in Sentiment			
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 2	Model 3	Model 4	Model 5
$\theta_1$	3.234**	3.090***	3.535***	1.142***	0.423***	1.179***	2.280***	-0.371	1.171***
$\theta_2$	0.508**	0.139**	0.883***	0.919***	0.664***	0.401***	0.666***	0.780***	1.889***
$\theta_3$	-2.225**	-2.323***	-2.683***	-0.743***	-0.498***	-0.633***	-2.152***	-0.417**	-1.016***
$\theta_4$	0.815***	0.834***	0.898***	0.646***	0.706***	0.705***	0.893***	0.830***	0.717***
$\theta_5$	-0.149***	-0.105***	-0.133***	0.178***	-0.218***	-0.158**	-0.134***	-0.125***	-0.018***
$\theta_6$	0.144**	0.191***	0.096***	-0.173***	0.128***	0.276***	0.225***	0.182***	0.037***
$\theta_7$	0.033***	0.024***	0.036***	0.024***	0.001***	-0.007***	0.056***	0.030**	0.019***
$\theta_8$	0.975***	0.987***	0.975***	0.980***	0.989***	0.947***	0.962***	0.979***	0.988***
$\theta_9$	0.065*	0.057***	0.092***	-0.087***	0.034***	0.043	0.127***	-0.003***	0.037***
$\theta_{10}$	-0.065**	0.016	0.057***	0.314***	-0.037***	-0.178***	-0.003	-0.099***	0.015***
$\theta_{11}$		0.173**	0.426***	0.373***	1.046***				
$\theta_{11}$						1.576***	1.629***	1.182**	1.656***
$\theta_{12}$		-0.004				0.082***			
$\theta_{13}$		-0.006				-0.602			
$\theta_{14}$		0.008				-0.006			
$\theta_{15}$		0.008				-0.180***			
$\theta_{16}$				-0.011**	0.003***			-0.120***	0.087***
$\theta_{17}$				0.001***	-0.001***			0.034***	-0.123***
$\theta_{18}$				-0.133				-0.034	
$\theta_{19}$				-0.148				0.036	
$\theta_{20}$				0.337***	-0.345***			-0.065***	-0.659***
$\theta_{21}$				0.362***	0.140***			-0.220***	0.627***
$\theta_{22}$				0.188*				0.075	
$\theta_{23}$				0.159				-0.006	
L.L	-1158.96	-1175.73	-1171.42	-1167.99	-1164.23	-1175.73	-1167.50	-1172.72	-1154.42
AIC	7.542	7.683	7.629	7.648	7.608	7.682	7.628	7.688	7.545
SC	7.662	7.862	7.809	7.877	7.789	7.816	7.761	7.917	7.617

Note: L.L: Loglikelihood.

This table shows the time-series regression results of EGARCH component and sentiment-augmented EGARCH component model, where sentiment is augmented in various ways, with levels and changes of sentiment respectively. \*/\*\*/\*\* denotes significance at 10% / 5% / 1% level.

Model 1: the benchmark model

$$R_{t+1}^M = \theta_1 + \theta_2 s_t + \theta_3 l_t + \sqrt{h_t} z_{t+1} \quad (1a)$$

$$\ln \sqrt{h_t} = l_t + s_t \quad (1b)$$

$$s_{t+1} = \theta_4 s_t + \theta_5 z_{t+1} + \theta_6 (|z_{t+1}| - \sqrt{2/\pi}) \quad (1c)$$

$$l_{t+1} = \theta_7 + \theta_8 l_t + \theta_9 z_{t+1} + \theta_{10} (|z_{t+1}| - \sqrt{2/\pi}) \quad (1d)$$

Model 2:

$$R_{t+1}^M = \theta_1 + \theta_2 s_t + \theta_3 l_t + \theta_{11} SENT_{t+1} + \sqrt{h_t} z_{t+1}$$

$$s_{t+1} = \theta_4 s_t + \theta_5 z_{t+1} + \theta_6 (|z_{t+1}| - \sqrt{2/\pi}) + \theta_{12} SENT_t^2 D_t + \theta_{13} SENT_t^2 (1 - D_t)$$

$$l_{t+1} = \theta_7 + \theta_8 l_t + \theta_9 z_{t+1} + \theta_{10} (|z_{t+1}| - \sqrt{2/\pi}) + \theta_{14} SENT_t^2 D_t + \theta_{15} SENT_t^2 (1 - D_t)$$

Model 3:

$$R_{t+1}^M = \theta_1 + \theta_2 s_t + \theta_3 l_t + \theta_{11} SENT_{t+1} + \sqrt{h_t} z_{t+1}$$

$$s_{t+1} = \theta_4 s_t + \theta_5 z_{t+1} + \theta_6 (|z_{t+1}| - \sqrt{2/\pi})$$

$$l_{t+1} = \theta_7 + \theta_8 l_t + \theta_9 z_{t+1} + \theta_{10} (|z_{t+1}| - \sqrt{2/\pi})$$

Model 4:

$$\begin{aligned}
R_{t+1}^M &= \theta_1 + \theta_2 s_t + \theta_3 l_t + \theta_{11} SENT_{t+1} + \sqrt{h_t} z_{t+1} \\
s_{t+1} &= \theta_4 s_t + \theta_5 z_{t+1} + \theta_6 \left( |z_{t+1}| - \sqrt{2/\pi} \right) + \theta_{16} s_t D_t + \theta_{17} s_t (1 - D_t) + \theta_{18} Z_{t+1} D_{t+1} + \theta_{19} Z_{t+1} (1 - D_{t+1}) \\
l_{t+1} &= \theta_7 + \theta_8 l_t + \theta_9 z_{t+1} + \theta_{10} \left( |z_{t+1}| - \sqrt{2/\pi} \right) + \theta_{20} l_t D_t + \theta_{21} l_t (1 - D_t) + \theta_{22} Z_{t+1} D_{t+1} + \theta_{23} Z_{t+1} (1 - D_{t+1})
\end{aligned}$$

Model 5:

$$\begin{aligned}
R_{t+1}^M &= \theta_1 + \theta_2 s_t + \theta_3 l_t + \theta_{11} SENT_{t+1} + \sqrt{h_t} z_{t+1} \\
s_{t+1} &= \theta_4 s_t + \theta_5 z_{t+1} + \theta_6 \left( |z_{t+1}| - \sqrt{2/\pi} \right) + \theta_{16} s_t D_t + \theta_{17} s_t (1 - D_t) \\
l_{t+1} &= \theta_7 + \theta_8 l_t + \theta_9 z_{t+1} + \theta_{10} \left( |z_{t+1}| - \sqrt{2/\pi} \right) + \theta_{20} l_t D_t + \theta_{21} l_t (1 - D_t)
\end{aligned}$$

The estimation results of the four models are reported in table 4. For each of the models including investor sentiment, we estimate both the levels and changes of sentiment respectively. The major findings are summarized below.

First, across the three models with investor sentiment, almost all of the estimated coefficients in the base models are significant, for both levels and changes of sentiment indices. The only one estimate that is not significant is  $\theta_1$ , which is trivial in the model. The estimates of the short- and long-run volatilities are opposite for all the cases, which supports the fact that existing studies have difficulty detecting a time-series relationship between aggregate risk and expected returns.

Second, the levels of sentiment are significantly positively related to market returns while the changes are significantly negatively related to market excess returns. This suggests that sentiment is an important factor in explaining equity excess returns. On the contrary, in Model 2 where sentiment enters directly into the conditional volatility, we find that estimates of both levels and changes of sentiment are insignificant and the magnitudes are very small. Furthermore, in Model 4 where sentiment dummies influences transitory and permanent volatilities through their impacts on market shocks and lagged short- and long-components of volatilities. However, the estimates of sentiment dummy affecting market shocks are insignificant throughout the levels and changes of sentiment. The hypothesis that sentiment or changes of sentiment has no effect on market shocks in the EGARCH component model cannot be rejected statistically.

Empirical results of Model 2 and 4 suggest further refinement of our specifications of empirical models. In model 3, after dropping the sentiment elements in the volatility equations, all the estimates are significant and the log likelihood and information criteria are improved. Therefore, Model 3 can be treated as a remedy with respect to Model 4. We propose Model 5 which drops the effects of sentiment dummy on return shocks to rectify Model 4.

$$R_{t+1}^M = \theta_1 + \theta_2 s_t + \theta_3 l_t + \theta_{11} SENT_{t+1} + \sqrt{h_t} z_{t+1}$$

$$s_{t+1} = \theta_4 s_t + \theta_5 z_{t+1} + \theta_6 \left( |z_{t+1}| - \sqrt{2/\pi} \right) + \theta_{16} s_t D_t + \theta_{17} s_t (1 - D_t)$$

$$l_{t+1} = \theta_7 + \theta_8 l_t + \theta_9 z_{t+1} + \theta_{10} \left( |z_{t+1}| - \sqrt{2/\pi} \right) + \theta_{20} l_t D_t + \theta_{21} l_t (1 - D_t)$$

After the refinement, the estimates are all significant. Again, the shift in sentiment has a significant positive impact on market excess returns while the shift in changes in sentiment has a significant negative impact on excess return. The increase in log-likelihood value and decrease of information criteria attest to improvement in the goodness of fit for either the level or the changes of sentiment index. The estimation results are shown in Table 4.

As pointed out by DeLong et al. (1990), when investor sentiment is bullish, the trading of noise traders creates price pressure that leads to a purchase price higher than fundamental value and thereby lowers expected returns. On the other hand, when noise traders are bullish, they increase their demand for the risky assets which amplifies the level of market risk, which is known as hold-more effect. Hence, they thereby expect to enjoy a higher return. The overall effect of sentiment on stock returns depends on which effect dominates. Therefore, the significant positive estimate of level sentiment (1.046) implies that the hold-more effect dominates the price-pressure effect. The hold-more effect tends to dominate the price-pressure effect and leads to an increase in market excess return when investors are more bullish.

Third, from estimations of Model 5, we also find that both the levels and changes of sentiment have significant and asymmetric effects on short- or long-run volatility that in turn influence the future short- or long-run volatility respectively. When noise traders are bullish, sentiment has a significantly positive, though little, impact on short-run volatility which further increases future short-run volatility, whereas sentiment is negatively related to long-run volatility and decreases future long-run volatility.

Combined with the signs of coefficients on short- and long-run volatility in mean equation, the overall outcome of direct sentiment, variations of short- and long-run volatilities triggered by sentiment is an increase in market excess return, when investors are optimism, the market excess return. On the contrary, when sentiment is bearish, levels of sentiment have a negative impact on short-run volatility and positive impact on long-run volatility. The overall effect of sentiment, variations of short- and long-run volatilities triggered by sentiment is a decrease in market excess return.

Specifically, given a positive sentiment, a 1% increase in sentiment results in 0.3% increase in future short-run volatility and -34.5% decrease in future long-run volatility. The overall effect of 1% increase in sentiment leads to 1.285% increase in market excess return. When sentiment is bearish, the overall effects of 1% decrease in sentiment leads to 1.116% lower excess return. The patterns of effects of changes in sentiment on excess return are similar to levels of sentiment. A 1% upward shift after a positive change in sentiment results in a 2.49% increase in market excess return. A 1% percent downward shift after a negative change in sentiment results in a 2.53% decrease in market excess return.

Finally, Model 3 is preferred to Model 2, and Model 5 is preferred to Model 4. However, Model 3 and 5 are not more favourable compared to the benchmark model since the log-likelihood value and information criteria are not improved.

## 5.2 Cross-Sectional Regressions of ICAPM

In section 5.1, the sentiment affected transitory and permanent volatilities are obtained from the time-series estimation of sentiment-augmented Adrian and Rosenberg's EGARCH component model. Ang et al. (2006) set up a standard two-factor pricing kernel with the market return and stochastic volatility as factors in the framework of ICAPM. They show that market volatility is a significant cross-sectional asset pricing factor. Adrian and Rosenberg (2008) present that the short- and long-run volatility components have negative, highly significant prices of risk which is robust across sets of portfolios, sub-periods, and volatility model specifications. In this section, we run the Fama-Macbeth two-stage regression to determine whether these two volatilities at different horizons after the introduction of investor sentiment remain significantly priced across assets. Meanwhile, we examine whether the sentiment-augmented EGARCH component model outperforms the benchmark model which does not take investor sentiment into account.

For the cross-sectional pricing tests, the innovations of short- and long-run volatilities are acquired by subtracting the short- and long-run component from the value expected one month earlier.

$$sres_m = s_m - E_{m-1} [s_m] \quad (7)$$

$$lres_m = l_m - E_{m-1} [l_m] \quad (8)$$

In this section, we only analyse the refined models where investor sentiment affects the mean equation only, Model 3 and impacts both mean and conditional volatility, Model 5. The estimated prices of risks, including the transitory volatility, the permanent volatility and the aggregate volatility, are presented in table 5. Panel A reports the results of Model 3 and Panel B reports the results of Model 5. Both panels are displayed with the statistics of the benchmark model, Model 1. The overall performance of the model specifications, the adjusted cross-sectional R-square is provided to describe how well the model fits the data.

Column (i) of Panel A exhibits that by applying monthly data, the short- and long-run components of volatility have significant positive prices. The price of aggregate volatility is also significantly negative as shown in column (ii). The short- and long-run sentiment (changed sentiment)-affected volatility of Model 3 are significantly negative pricing factors in the cross-section as shown in column (iii) (column (v)). Their respective aggregate volatilities are also significantly negatively priced across portfolios. The conclusions hold true for Model 5 as reported in Panel B.



In terms of pricing performance, there are two main inferences. First, the EGARH component volatility decomposition model compares favourably with the aggregate volatility model proposed by Ang et al. (2006). Second, after the introduction of investor sentiment, the goodness of fit of Model 3 and Model 5 are greatly enhanced. Finally, level sentiment-affected specifications outperform the specifications with changed sentiment-affected volatilities, for both Model 3 and Model 5.

Table 5: Summary statistics of the cross-sectional Fama-MacBeth regression for the 25 size and book-to-market sorted portfolios.

Panel A: The cross-sectional regression of Model 3, using levels and changes of sentiment respectively

This table reports the two-stage cross-sectional regression results for the 25 size and B/M sorted portfolios under ICAPM model with different state variables. Specifically, column (i) uses market excess return and innovations of short- and long-run volatilities from Model 1 of Table 4 as state variables. Column (ii) uses market excess return and aggregate volatility innovation from Model 1 of Table 4 as state variables. Columns (iii) to (vi) of Model 3 use market excess return and innovations of short- and long-run volatilities from Model 3 of Table 4. Columns (iii) to (vi) of Model 5 use market excess return and innovations of short- and long-run volatilities from Model 5 of Table 4. The *t*-ratios are calculated using Jagannathan and Wang (1998) and Newey and West (1987) procedures to account for the estimation errors in first-stage estimation and correct for the possible heteroskedasticity and autocorrelation.

		Model 1		Model 3			
		(i)	(ii)	Level of sentiment		Change in Sentiment	
		s, l	variance	(iii)	(iv)	(v)	(vi)
		s, l	variance	s, l	variance	s, l	variance
Market excess return	Coef.	-0.432***	-0.853***	-0.456***	-0.848***	-0.392***	-0.577***
	t-stat.	-3.266	-4.302	-3.552	-3.981	-2.871	-3.142
Short-run volatility	Coef.	-0.289**		-0.278***		-0.237*	
	t-stat.	-3.567		-5.355		-1.723	
Long-run volatility	Coef.	-0.303***		-0.300**		-0.285***	
	t-stat.	-3.812		-2.732		-3.299	
Market variance	Coef.		-1.789*		-3.862**		-3.964***
	t-stat.		-1.847		-2.161		-4.151
Adjusted R-squared		0.392	0.326	0.552	0.488	0.545	0.455

Panel B: The cross-sectional regression of Model 5, using levels and changes of sentiment respectively

		Model 1		Model 5			
		(i)	(ii)	Level of Sentiment		Change in Sentiment	
		s, l	variance	(iii)	(iv)	(v)	(vi)
		s, l	variance	s, l	variance	s, l	variance
Market excess return	Coef.	-0.432***	-0.853***	-0.262*	-0.298**	-0.927***	-1.267***
	t-stat.	-3.266	-4.302	-2.045	-2.353	-3.171	-3.455
Short-run volatility	Coef.	-0.289***		-0.579***		-0.337**	
	t-stat.	-3.567		-4.698		-2.264	
Long-run volatility	Coef.	-0.303***		-0.457***		-0.390***	
	t-stat.	-3.812		-5.238		-4.040	
Market variance	Coef.		-1.789*		-4.718**		-3.883**
	t-stat.		-1.847		-2.161		-2.536
Adjusted R-squared		0.392	0.382	0.555	0.522	0.537	0.520

\*Significant at 10% level.

\*\* Significant at 5% level.

\*\*\* Significant at 1% level.

## 6. Robustness Analysis

### 6.1 Robustness Analysis with Alternative Sample Period

Table 6: Prices of risks over different subsamples

This table reports the two-stage cross-sectional regression results for the 25 size and B/M sorted portfolios using different sample periods. The state variables used are market excess return and innovations of short- and long-run volatilities from Model 3 and Model 5 of Table 4, respectively. The *t*-ratios are calculated using Jagannathan and Wang (1998) and Newey and West (1987) procedures to account for the estimation errors in first-stage estimation and correct for the possible heteroskedasticity and autocorrelation.

Panel A: summary statistics of Fama-MacBeth regressions of Model 3 using different sample periods, for levels and changes of sentiment respectively

		Level of Sentiment			Change in Sentiment		
		(i)	(ii)	(iii)	(i)	(ii)	(iii)
Excess market return	Coef.	-0.456*	-1.005**	-1.118***	-0.392**	-1.817***	-1.25**
	t-stat.	-2.011	-2.636	-3.322	-2.667	-3.523	-2.611
Short-run volatility	Coef.	-0.417***	-0.378***	-0.391**	-0.237*	-0.347***	-0.358**
	t-stat.	-5.355	-2.868	-2.116	-1.723	-3.577	-2.533
Long-run volatility	Coef.	-0.450**	-0.463*	-0.415*	-0.285***	-0.404**	-0.305*
	t-stat.	-2.732	-1.918	-1.944	-3.299	-2.481	-1.771
Adjusted R-squared		0.552	0.637	0.488	0.545	0.516	0.455

Panel B: summary statistics of Fama-MacBeth regressions of Model 5 using different sample periods, for levels and changes of sentiment respectively

		Level of Sentiment			Change in Sentiment		
		(i)	(ii)	(iii)	(i)	(ii)	(iii)
Excess market return	Coef.	-0.262***	-1.625***	-1.697*	-0.927***	-1.930***	-1.386*
	t-stat.	-2.878	-2.978	-1.756	-2.853	-3.918	-1.965
Short-run volatility	Coef.	-0.579***	-0.473***	-0.346**	-0.337**	-0.297*	-0.491*
	t-stat.	-4.698	-3.126	-2.140	-2.264	-1.867	1.773
Long-run volatility	Coef.	-0.457***	-0.470**	-0.461*	-0.390***	-0.325*	-0.396***
	t-stat.	-5.238	-2.266	-1.188	-4.040	-1.930	-4.988
Adjusted R-square		0.555	0.567	0.491	0.537	0.548	0.446

Column (i): sample period from 1987m03 to 2012m12, which is the full sample period;

Column (ii): sample period from 1987m03 to 2007m06, which is before crisis period;

Column (iii): sample period from 1987m03 to 2012m12, excluding period of 2007m07 to 2010m06, which is clarified as crisis period.

\*Significant at 10% level.

\*\* Significant at 5% level.

\*\*\* Significant at 1% level.

In this section, the cross-sectional pricing results over different sample periods are examined and depicted in Table 6. We find that the prices of risk for the two volatility components are significantly negative across sample periods, regardless the model specification, for both levels and changes of investor sentiment. The magnitudes of the prices of risk for the two volatility components are fairly similar across sample periods, which suggest our results are robust to model specification and sample selections.

The estimations of sample period before crisis shown in column (ii) achieve the highest adjusted cross-sectional  $R^2$ . We can infer that the sentiment-augmented EGARCH component model fits the sample period better before crisis period, whereas the estimations of sample excluding crisis period acquire the lowest goodness of fit.

## 6.2 Robustness Analyses of Different Measures of Investor Sentiment

This section demonstrates a robustness test on the construction of the sentiment index. A host of robustness checks is undertaken to examine if our results are driven by some admittedly arbitrary choice. We consider an alternative proxy for investor sentiment, the consumer confidence, as a direct investor sentiment measure. Furthermore, in the framework of principal component analysis, the put-call trading volume and open interest ratios of derivative market are included to construct a new composite sentiment index, in comparison with the sentiment analysed in the previous sections.

### 6.2.1 Consumer confidence as an alternative to proxy for investor sentiment

The consumer confidence is orthogonalized to the control variables and the residuals are taken as the level sentiment measures ( $SENT^{CC}$ ). The residuals from the changed consumer confidence are treated as measures of changed sentiment ( $\Delta SENT^{CC}$ ). There's no Granger causality relationship between market excess return and consumer confidence in short duration, for example, within 2 months. The correlation between the  $SENT$  and  $SENT^{CC}$  is 0.578, while the correlation between the changed sentiment between  $\Delta SENT$  and  $\Delta SENT^{CC}$  is 0.231 which is surprisingly not very high. The graphs of these sentiment measures are depicted in Figure 4.

Figure 4: Graphs of sentiment from different measures, dating from March 1987 to December 2012.

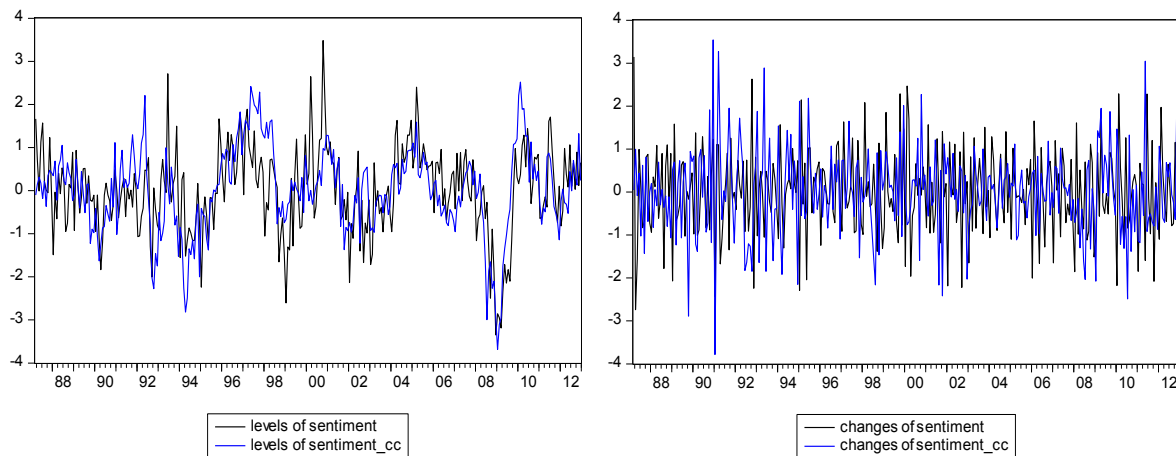


Table 7 presents the cross-sectional estimation results using short- and long-run volatilities from sentiment-augmented EGARCH component model where the consumer confidence is used directly as the measure of investor sentiment. The prices of short- and long-run components are significantly negative. The magnitudes of the estimates are smaller compared to those reported in Table 5.

Table 7: The Fama-MacBeth regressions of prices of short- and long-run sentiment-affected volatilities. Consumer confidence works as direct proxy for investor sentiment.

This table reports the two-stage cross-sectional regression results for the 25 size and B/M sorted using innovations of short- and long-run volatilities estimated from Model 3 and Model 5 employing consumer confidence as investor sentiment. The  $t$ -ratios are calculated using Jagannathan and Wang (1998) and Newey

and West (1987) procedures to account for the estimation errors in first-stage estimation and correct for the possible heteroskedasticity and autocorrelation.

		Level of Sentiment		Change in Sentiment	
		Model 3	Model 5	Model 3	Model 5
Market excess return	Coef.	-0.639**	-0.815**	-0.782**	-0.262*
	t-stat.	-2.455	-2.572	-2.263	-1.956
Short-run volatility	Coef.	-0.151***	-0.222***	-0.167**	-0.269**
	t-stat.	-3.302	-3.321	-2.374	-2.068
Long-run volatility	Coef.	-0.270***	-0.204*	-0.331**	-0.211*
	t_stat.	-3.219	-1.993	-2.105	-2.051
Adjusted R-square		0.447	0.516	0.491	0.555

\*Significant at 10% level.

\*\* Significant at 5% level.

\*\*\* Significant at 1% level.

## 6.2.2 First principal component analysis of sentiment index with the inclusion of data of FTSE 100 options

The put-call trading volume ratio, denoted as PCV, is widely recognized as a bearish indicator. It is a measure of market participants' sentiment derived from options which is equal to the trading volume of put options over the trading volume of call options. Investors tend to buy put options either to hedge their spot positions or to speculate, when they are bearish. When the trading volume of put options becomes large with respect to the trading volume of call option, the ratio goes up, and vice versa.

An alternative approach to calculate the put-call volume ratio is to use the open interest of options instead of trading volume. It may be argued that the open interest ratio is the final picture of sentiment at the end of day or on a monthly basis and therefore, it might be a preferred measure of sentiment index. The put-call open interest ratio is labelled as PCO.

We use the trading volume and open interests of FTSE 100 option (UKX) to calculate the put-call trading volume and open interest ratios. The data of trading volume and open interest are taken from Bloomberg. Unfortunately, the trading volume data start from 31/10/1994 while the complete data of open interests only originate from 31/12/1998. For convenient consideration, the new composite sentiment index is formed from December 1998 to December 2012. The procedure of forming this new sentiment and its changed values are the same as described in Section 4.2. The parsimonious new sentiment is presented in Equations (9).

$$SENT_t^{new} = 0.341 * CC_{t-1} + 0.542 * TURN1_{t-1} + 0.547 * TURN1_{t-1} + 0.3506 * NIPO_t + 0.231792 * RIPO_{t-1} + 0.192 * PCO_{t-1} + 0.279 * PCT_t \quad (9)$$

The first principal component of sentiment proxy explains 51.79% of the total variance. The changes in the index of sentiment levels are obtained by taking the first-order difference. There's no Granger causality relationship between the new sentiment index and market excess returns for one- and two-month lags.

Table 8: Correlations of the three sentiment measures and changes in sentiment measures

	$SENT$	$SENT^{CC}$	$SENT_t^{new}$		$\Delta SENT$	$\Delta SENT^{CC}$	$\Delta SENT_t^{new}$
$SENT$	1.000			$\Delta SENT$	1.000		
$SENT^{CC}$	0.674	1.000		$\Delta SENT^{CC}$	0.293	1.000	
$SENT_t^{new}$	0.906	0.674	0.622	$\Delta SENT_t^{new}$	0.935	0.210	1.000

The correlation of levels of sentiment and changes in sentiment measured from three alternative approaches are reported in Table 8. Sentiment with and without the inclusion of option data are highly correlated with each other, for both levels and changes of sentiment. The graphs of levels and changes of sentiment measured from three different ways are exhibited in Figure 5.

Figure 5: Graphs of sentiment from different measures, dating from Dec 1998 to Dec 2012

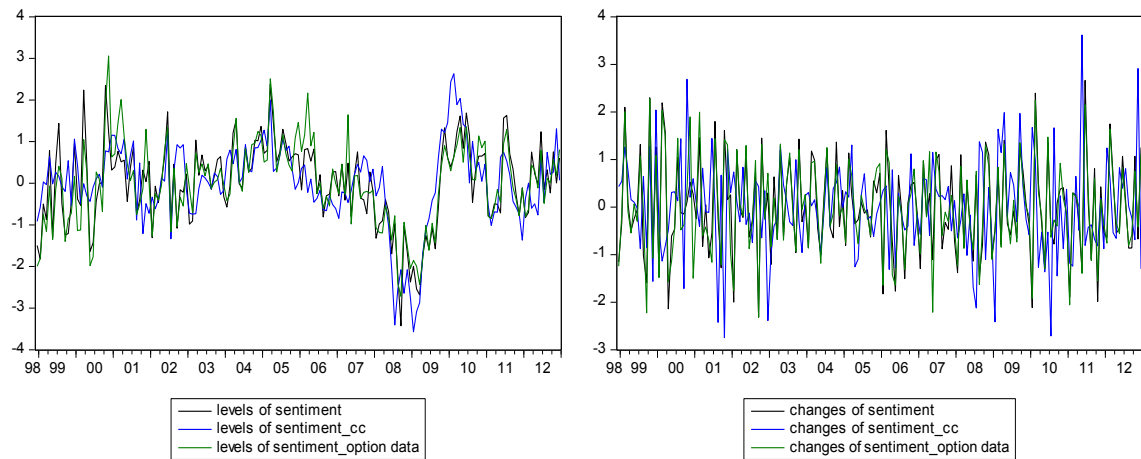


Table 9: The Fama-MacBeth regressions of prices of short- and long-run sentiment-affected volatilities. Levels of investor sentiment are measured in Equations (9) and changes in sentiment index are obtained by taking the first-order difference.

This table reports the two-stage cross-sectional regression results for the 25 size and B/M sorted using innovations of short- and long-run volatilities estimated from Model 3 and Model 5 employing the new composite sentiment index with option data. The *t*-ratios are calculated using Jagannathan and Wang (1998) and Newey and West (1987) procedures to account for the estimation errors in first-stage estimation and correct for the possible heteroskedasticity and autocorrelation.

		Level of Sentiment		Change in Sentiment	
		Model 3	Model 5	Model 3	Model 5
Market excess return	Coef.	-1.081**	-0.488*	-0.488*	-0.483**
	t-stat.	-2.195	-2.038	-2.038	-2.288
Short-run volatility	Coef.	-0.462*	-0.880***	-0.880***	-0.343***
	t-stat.	-2.037	-2.973	-2.973	-5.636
Long-run volatility	Coef.	-0.576*	-0.926***	-0.926***	-0.599***
	t-stat.	-1.835	-2.601	-2.601	-3.490
Adjusted R-squared		0.434	0.437	0.437	0.520

\*Significant at 10% level.

\*\* Significant at 5% level.

\*\*\* Significant at 1% level.

Table 9 shows the cross-sectional estimation results using short- and long-run volatilities from sentiment-augmented EGARCH component model where the put-call trading volume as well as open interest ratios of FTSE 100 options is taken into account for the first principal analysis. The prices of short- and long-run components are significantly negative. In contrast to the results displayed in table 7, the magnitudes of the estimates are higher compared to those shown in Table 5. However, the adjusted R-square is reduced among most cases.

In conclusion, the volatility components remain highly significant for different sample periods and the significance is robust to the choice of the measures of investor sentiment.

## 7. Conclusion

The classical finance theory leaves no role for investor sentiment in cross-section of stock prices, realized returns, or expected returns. This view has been challenged by researchers in behavioural finance. Empirical results suggest that investor sentiment has significant cross-sectional effects.

The model of Delong et al. (1990) predicts that the direction and magnitude of noise trading risk are relevant in asset pricing. Noise traders' belief, recognized as investor sentiment, affects asset returns and systematic risks in this model. Lee et al. (2002) propose a GARCH-in-mean specification to explicitly test the impact of noise trader risk on both the formation of expected return and conditional volatility. Inspired by the studies of Delong (1999), Lee et al. (2002) and Adrian and Rosenberg (2008), we use a sentiment augmented EGARCH component model to exploit the time-series relationship between sentiment and market return and market volatilities. We further test the cross-section prices of short- and long- run components of market volatility which are affected by investor sentiment.

By augmenting sentiment to the mean of Adrian and Rosenberg's EGARCH component model, or to both the mean and variance equations, market excess returns are significantly positively related to investor sentiment and are significantly negatively related to the changes in investor sentiment in the time-series estimations. In the cross-sectional estimations, significantly negative prices of short- and long-run components of volatility are detected. Our models which taking market sentiment into consideration outperform the pure EGARCH component model since they achieve higher explanatory

power indicated by larger adjusted R-square. Therefore, the incorporation of sentiment enhances the price ability of short- and long-run volatilities of EGARCH component model. The conclusion is robust to the choice of sample periods and alternative constructions of investor sentiment.

## Reference:

- Adrian, T., and J., Rosenberg. 2008. Stock Returns and Volatility: Pricing the Short-Run and Long-Run Components of Market Risk. *The Journal of Finance*, Vol 63(6), 2997-3030.
- Ang, A., R.J., Hodrick, Y.H., Xing, and X.Y., Zhang. 2006. The Cross-Section of Volatility and Expected Returns. *The Journal of Finance*, Vol LXI(1), 259-299.
- Arik, A. 2012. Modeling Market Sentiment and Conditional Distribution of Stock Index Returns under GARCH Process. *ProQuest*, UMI Dissertation Publishing.
- Arms, R.W. 1989. The Arms Index Trin: An Introduction to the Volume Analysis of Stock and Bond Markets Hardcover. Publisher: Irwin Professional Pub.
- Baker, M., and J., Stein. 2004. Market liquidity as a sentiment indicator. *Journal of Financial Markets*, 7, 271-299.
- Baker, M., and J., Wurgler. 2000. The equity share in new issues and aggregate stock returns. *Journal of Finance*, 55, 2219-2257.
- Baker, M., and J., Wurgler. 2004a. A catering theory of dividends. *Journal of Finance*, 59, 1125-1165.
- Baker, M., and J., Wurgler. 2004b. Appearing and disappearing dividends: The link to catering incentives. *Journal of Financial Economics*, 73, 271-288.
- Baker, M., and J., Wurgler. 2006. Investor sentiment and the cross-section of stock returns. *Journal of Finance*, 61, 1645-1680.
- Baker, M., and J., Wurgler, 2007. Investor Sentiment in the Stock Market. *Journal of Economics Perspectives*, 21(2), 129-151.
- Barber, B., T., Odean and N., Zhu. 2007. Systematic noise. *Working Paper*, University of California at Davis.
- Berger, D., and H.J., Turtle. 2012. Cross-sectional Performance and Investor Sentiment in a Multiple Risk Factor Model. *Journal of Banking and Finance*, 36, 1107-1121.
- Brown, G.W., and M.T., Cliff. 2004. Investor sentiment and the near-term stock market. *Journal of Empirical Finance*, 11, 1-27.
- Brown, G.W., and M.T., Cliff. 2005. Investor Sentiment and Asset Valuation. *Journal of Business*, 78(2), 405-40.
- Brown, S., W., Goetzmann, T., Hiraki, N., Shiraishi, M., Watanabe. 2002. Investor Sentiment in Japanese and U.S. Daily Mutual Fund Flows. *Working Papers*. Yale School of Management.
- Campbell, J.Y. 1993. Intertemporal Asset Pricing without Consumption Data. *American Economic Review*, 83, 487-512.
- Campbell, J.Y. 1996. Understanding Risk and Return. *Journal of Political Economy*, 104, 298-345.
- Chen J., 2002. Intertemporal, CAPM and the Cross-section of Stock Returns. *Working paper*, University of Southern California.
- Corredor, P., E., Ferrer, R., Santamaria. 2013. Investor Sentiment Effect in Stock Markets Stock Characteristics or Country-specific Factors. *International Review of Economics & Finance*, 27, 572-591.



- Da, Z., and E., Schaumburg. 2011. Relative Valuation and Analyst Target Price Forecasts. *Journal of Financial Markets*, Vol 14, 161-192.
- DeLong, J. B., A., Shleifer, L.H., Summers and R.J., Waldmann. 1990. Noise trader risk in financial markets. *Journal of Political Economy*, 98(4), 703-738.
- Engle, R.F., and G., Lee. 1999. A Permanent and Transitory Component Model of Stock Return Volatility. *Cointegration, Causality and Forecasting: A Festschrift in Honor of Clive W.J. Granger*, Oxford University Press, New York.
- Fama, E.F., and K.R., French. 1992. The Cross-Section of Expected Stock Returns. *The Journal of Finance*, Vol XLVII(2), 427-465.
- Fama, E.F., and K.R., French. 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, 33, 3-56.
- Fisher, K., and M., Statman. 2000. Investor Sentiment and Stock Returns. *Financial Analysts Journal*, 56(2), 16-23.
- Fisher, K., and M., Statman. 2003. Consumer Confidence and Stock Returns. *Journal of Portfolio Management*, 30(1), 115-128.
- Frazzini, A., and O., Lamont. 2006. Dumb Money: Mutual Fund Flows and the Cross-section of Stock Returns. *Working Paper*, 11526, National Bureau of Economic Research.
- Fosback, N.G. 1993. *Stock Market Logic: A Sophisticated Approach to Profit on Wall Street*. Dearborn Financial Publishing: Chicago, IL.
- Jagannathan, R., and Z.Y., Wang. 1998. An asymptotic theory for estimating beta-pricing models using cross-sectional regression, *Journal of Finance*, 53, 1285-1309.
- Kumar, A., and C., Lee. 2006. Retail Investor Sentiment and Return Comovement. *Journal of Finance*, 61(5), 2451-2486.
- Lee, C., A., Shleifer and R.H., Thaler. 1991. Investor Sentiment and the Closed-End Fund Puzzle. *Journal of Finance*, 46, 75-109.
- Lee, W.Y., C.X., Jiang and D.C., Indro. 2002. Stock Market Volatility, Excess Returns, and the Role of Investor Sentiment. *Journal of Banking and Finance*, 26, 2277-2299.
- Lemmon, M., and E., Portniaguina. 2006. Consumer Confidence and Asset Prices: Some Empirical Evidence. *Review of Financial Studies*, 19(4), 1499-1529.
- Neal, R., and S.M., Wheatley. 1998. Do Measures of Investor Sentiment Predict Returns?. *Journal of Financial & Quantitative Analysis*, 33(4), 523-548.
- Nelson, D.B. 1991. Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, Vol 59(2), 347-370.
- Newey, W.K., and K.D., West. 1987. A Simple, Positive-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55, 703-708.
- Pastor, L., and R.F., Stambaugh. 2003. Liquidity Risk and Expected Stock Returns. *The Journal of Political Economy*, 111(3), 642-685.

Petkova R. 2006. Do the Fama-French Factors Proxy for Innovations in Predictive Variables? *The Journal of Finance*, Vol 61(2), 581-612. Scheinkman, J., and W., Xiong. 2003. Overconfidence and speculative bubbles. *Journal of Political Economy*, 111, 1183-1219.

Schmeling, M. 2009. Investor Sentiment and Stock Return: Some International Evidence. *Journal of Empirical Finance*, 16, 394-408.

Shanken, J. 1992. On the Estimation of Beta-Pricing Models. *The Review of Financial Studies*, Vol 5(1), 1-33.

Shleifer, A., and R., Vishny. 1997. The limits of arbitrage. *Journal of Finance*, 52, 35-55.

Seyhun, H. N. 1998. Investment Intelligence from Insider Trading. *Cambridge, MA: MIT Press*, UK.

Wang, Y.H., A., Keswani, S.J., Taylor. 2006. The relationships between sentiment, returns and volatility. *Journal International Journal of Forecasting*, 22 (1), 109-123.

Whaley, R.E. 2008. Understanding VIX. *Working Paper*, Vanderbilt University.