Investor sentiment and stock prices in the subprime mortgage crisis

Alexander Wolff

Bachelor thesis for Industrial Engineering and Management
Student number: 0201820
First Supervisor University of Twente: Ir. H. Kroon (deventerhenk@yahoo.com)
Second Supervisor University of Twente: Ass. Prof. B. Roorda (deventerhenk@yahoo.com)
Supervisor Luxembourg School of Finance: Prof. Dr. T. Lehnert (thorsten.lehnert@uni.lu)

Contents

Abstract

As evidence from existing literature contradicts classical finance suggesting that there is room for investor sentiment in stock prices, and there is evidence that indicates a possible change in this relationship in the subprime mortgage crisis (since 2007), this paper uses several methods to measure investor sentiment for both individual and institutional investors in the U.S., and then frees it from macroeconomic trends. The Granger-causality test is used to ensure that it is investor sentiment causing stock market price changes, and not vice versa. Changes in stock market prices are then regressed on changes in the investor sentiment data freed of macroeconomic trends. The resulting regression coefficients from before and during the crisis are compared to determine the impact the subprime mortgage crisis has had on the relationship between investor sentiment and stock market prices. Out of the five investor sentiment indexes used, only two exhibit significant causality in the desired direction (the Individual One-Year Confidence Index and the Individual Valuation Confidence Index, both from Yale). For the Individual One-Year Confidence Index, strong statistical evidence is found that the predictive power of investor sentiment over stock market prices has increased in the subprime mortgage crisis.

1. Introduction

Traditional finance relies on models in which the behavior of investors is considered rational. These rational investors value stocks based on available information on risk and return. When new information becomes available on a particular stock, they re-evaluate it and act accordingly, updating its market price to reflect the new information. This implies that there is no investment strategy that can generate average returns greater than warranted by the risk taken. However, much research has been conducted on this subject, and it has been shown that these models often do not fully explain stock prices. Another approach to this matter is behavioral finance. This approach argues that in order to describe financial phenomena models should be used in which investors are not always rational. This irrationality, referred to as investor sentiment, can have an impact on stock prices (Barberis, 2003).

Chung, Hung and Yeh (2012) have found evidence that the relationship between investor sentiment and stock returns is subject to change. According to their findings, this relationship only exists in times of economic growth, and not in times of recession. Brown and Cliff (2005) have found evidence that is somewhat conflicting with these findings; they have demonstrated that investor sentiment is always an important factor in predicting future market returns. Lemmon and Portniaguina (2006) have shown that consumer confidence plays an important role in the concept of investor sentiment. In their findings there is strong evidence for a relationship between consumer confidence subsequent stock returns, but only after 1977. They have found that before this time, this relationship was not present. All of the presented literature suggests that a relationship between investor sentiment and market returns exists, but the stability of this relationship is not agreed upon.

The subprime mortgage crisis of 2007 has lead to a strong recession in the U.S and elsewhere. Leading to many bankruptcies, home foreclosures and regulatory revisions, this crisis has had a definite impact on consumer confidence. The subprime mortgage crisis has had a substantial impact on the behavior of economic agents, causing them to question their current beliefs and attitudes (Perriman, Ramsaran-Fowdar, & Baguant, 2011). This can be interpreted as a significant change, the kind that changes the impact of investor sentiment on market returns (like the growth-recession impact cycle and the 1977 change in this relationship). To further investigate this, this paper poses the following problem statement: "**Has the subprime mortgage crisis had an influence on the relationship between investor sentiment and U.S. stock market prices?**"

2. Background and literature review

The classical Efficient Market Hypothesis states that the price and earnings of a particular stock or portfolio are equal to the optimal forecast of the future cash flows and investment risks based on all the information that is currently available (Shiller, Market Volatility, 1992). Investors behave in a rational manner, taking into account all relevant information that is available to them in the decision-making process. When new information becomes available about a particular stock, rational investors will swiftly respond, adjusting the stock price to match the new risk that is attached to it.

Considerable evidence has been found that this hypothesis is not sufficient to completely explain the behavior of security prices. Basu (1997) states that the efficient market hypothesis prevents investors from earning excess returns, and proceeds to find anomalies pertaining to security information not being fully

reflected in its prices. A market inefficiency seems to exist where lags and frictions in the price adjustment process prevent security prices from immediately absorbing publically available information in its price. Lamont and Thaler (2003) have shown that some stocks are unambiguously overpriced, but at the same time do not represent exploitable arbitrage opportunities. Both of these papers suggest that stock prices often tend to differ from what classical finance theories expect them to be.

Alan Greenspan, the former chairman of the Federal Reserve Board, used the term "irrational exuberance" in his 1996 speech to describe the market overvaluation at that time. Further into his speech, he posed a rhetorical question: "But how do we know when irrational exuberance has unduly escalated asset values, which then become subject to unexpected and prolonged contractions as they have in Japan over the past decade?" The stock markets reacted strongly to this speech, falling 2-4% worldwide. Since then, this term has been commonly used to describe irrational speculative fervor (Shiller, Irrational Exuberance, 2000).

Barberis (2003) has explored the field of behavioral finance, and has found five indications that support the view that the Efficient Market Hypothesis is not the sole driver of stock prices. The first is that earlier papers drawing this conclusion were dismissed as erroneous. Since then, numerous authors have conducted similar research leading to the same conclusion, establishing consensus on the matter. Second, in the past it was the opinion of financial economists that the forces of arbitrage enforced the Efficient Market Hypothesis to be true. It is now understood that arbitrage has limits, creating opportunities for mispricing. Third, the concept of bounded rationality has been introduced by cognitive scientists. This model contains the notion that human behavior is influenced by more factors than pure rationality. Similar to the developments in psychology, behavioral finance has seen an influx of papers working with economic agents that are less than fully rational. Finally, recent work has tried to document and understand how both individual and institutional investors make their portfolio choices.

Brown and Cliff (2005) found evidence that a direct survey measure of investor sentiment (the categorizing of market newsletters into bullish, bearish and neutral) has the ability to predict market returns over the next few years and explain discrepancies between the stock prices expected by classical financial theory and actual stock prices. They explain these discrepancies as follows: *"Overly optimistic (pessimistic) investors drive prices above (below) fundamental values and these pricing errors tend to revert over a multi-year horizon. This pattern is consistent with the predictions of many behavioral models that prices under react in the short run and overreact in the long run."*

This pattern will be used to define investor sentiment. In measuring this, existing literature offers a few methods. Baker and Wurgler (2005) used market proxies for matters such as first-day returns of initial public offerings of companies as indications of investor sentiment. They have combined six of these proxies to develop a sentiment index, and have shown that this index has clearly discernable effects on the stock markets. These proxies are: trading volume as measured by NYSE turnover, dividend premium, the closed-end fund discount, the number and first-day returns on initial public offerings (IPOs) and the equity share in new issues. These proxies were then regressed on a number of U.S. macroeconomic variables to free them from any macroeconomic trends: growth in industrial production, real growth in durable, nondurable and services consumption, growth in employment, and the National Bureau of Economic Research (NBER) recession indicator. The residuals from these regressions have been used as measures of investor sentiment, and were averaged into an index to test for return predictability in stock markets.

Lemmon and Portniaguina (2006) made use of consumer confidence surveys (the University of Michigan survey of consumer sentiment and the Conference Board survey of consumer confidence) and regressed them on a set of U.S. macroeconomic variables: consumption-to-wealth ratio, closed-end fund discount, growth in personal consumption, inflation rate, default spread, dividend yield, real gross domestic product, the growth in labor income, and the unemployment rate. The results of these regressions were interpreted as measures of investor sentiment free of macroeconomic trends, and were compared to subsequent stock returns. They have found evidence that these measures of investor exhibit forecasting powers for the returns on small stocks and for future economic activity. They have also found that their measure of investor sentiment is not strongly related to those of Baker and Wurgler (2005), indicating that they either measure different aspects of investor sentiment, or one (or both) of them fail to capture the essence of this matter. Finally, they have found that before 1977, the relationship between investor sentiment and subsequent stock returns is nonexistent.

3. Methodology

To measure investor sentiment, this paper will use elements from both Baker and Wurgler (2005) and Lemmon and Portniaguina (2006). Investor sentiment data freed from macroeconomic trends will be collected. This data will then be broken into two periods of time: before and during the crisis (as the effects of the crisis are still ongoing). They will then be regressed on subsequent stock returns to determine its predictive power in these two sub periods.

The data from Baker and Wurgler Index is readily available, already freed from macroeconomic trends. No further action is required for this data in the first step. In its most recent form, this data is available at a monthly level from January 1977 until December 2010.

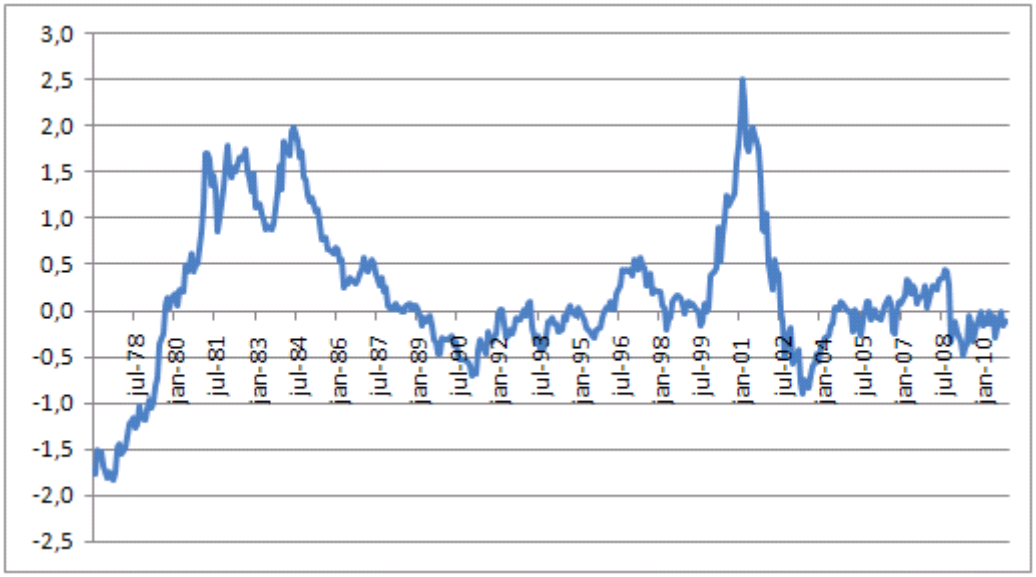

Figure 1: Baker & Wurgler Investor Sentiment Index (http://people.stern.nyu.edu/jwurgler/)

As a second source of data, this paper will adopt the method used by Lemmon and Portniaguina (2006) that regresses two surveys of consumer confidence on a set of macroeconomic variables, to generate investor sentiment data freed of any macroeconomic trends. Since it makes much more sense to use confidence

data related to investors rather than common consumers, this paper will use indexes based on investor confidence. Based on the recommendations of Zhang (2008), the One-Year Confidence Index (OYC) and the Valuation Confidence Index (VC) from the Yale School of Management's Stock Market Confidence Index are selected. These indexes have been derived from the responses to a single question that has been asked consistently since 1989 to a consistent sample of respondents, concerning stock prices. The question for the One-Year Confidence Index is as follows:

```
How much of a change in percentage terms do you expect in the following
(use + before your number to indicate an expected increase, a - to
indicate an expected decrease, leave blanks where you do not know):

                                        [FILL IN ONE NUMBER FOR EACH]

                          In 1     In 3     In 6     In 1     In 10
                          month    months   months   year     years

Dow Jones Industrial      _____%   _____%   _____%   _____%   _____%  Average
```

The One-Year Confidence Index is the percentage of respondents giving a number strictly greater than zero for "In 1 year". The Valuation Confidence Index uses the following question:

```
Stock prices in the United States, when compared with measures of true fundamental
value or sensible investment value, are:

                                [CIRCLE ONE NUMBER]

   1. Too low.      2. Too high.      3. About right.    4. Do not know.
```

The Valuation Confidence Index is the number of respondents who answer this question with 1) Too Low or 3) About Right as a percentage of those who choose 1, 2 or 3. The data for these indexes are available for both individual and institutional respondents, so an analysis will be conducted for both groups (with 1 for individual and 2 for institutional data, for example OYC1 will be used for Individual One-Year Confidence index). The data for these indexes are available at a monthly level from April 2001 until March 2012.

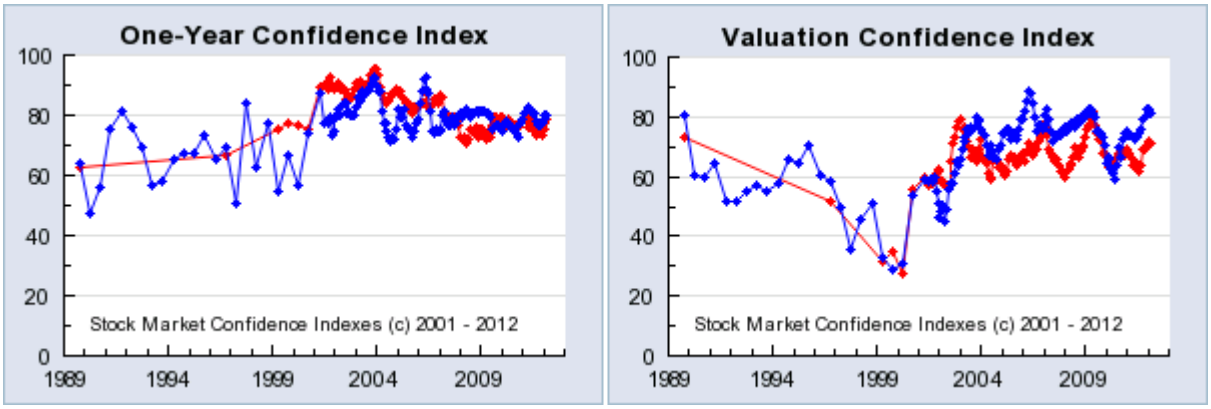

Figure 2: Yale Stock Market Confidence Indexes (http://icf.som.yale.edu/stock-market-confidence-indices-united-states) Institutional Investors shown in blue, Individual Investors shown in red.

This paper will follow existing literature in selecting macroeconomic variables. Six suitable variables can be derived from the Lemmon and Portniaguina (2006) article: default spread (DEF), dividend yield (DIV), inflation (INF), the three-month Treasury bill rate (TBILL), the unemployment rate (URATE), and the continuously compounded return of consumption (CONS). As GDP data is only available quarterly, the continuously compounded return of industrial production (IP) will be used instead, as described in Cuche and Hess (2000). Also, from the Estrella and Mishkin (1988) article, a final macroeconomic variable can be derived: the NBER recession cycle (NBER). Like in the Lemmon and Portniaguina (2006) article, these variables will be used contemporaneously (t) with the confidence index data. Based on the recommendations of my supervisor and other professors at the LSF, these variables will be used with a one-period (t-1) time lag as well.

To compare both sets of data before and during the subprime mortgage crisis, they will be split in two parts. Around the end of June 2007, several events were unfolding that marked the beginning of the subprime mortgage crisis. June 23rd, Bear Sterns pledged $3.2 billion U.S. Dollars to aid of its ailing hedge funds. June 26th, SEC began investigating 12 collateralized debt obligation (CDO) issuers in the mortgage market. July 17th, the FED, OTS and FTC launched a new program to supervise subprime mortgage lenders. Because these events started occurring in late June, this paper will use July as the first month of the crisis period (Federal Reserve Bank of New York, 2011). To be able to make meaningful comparisons between the confidence indexes, only the months in which data are available for all indexes will be used. This amounts to April 2001-December 2010.

The following model will be used to determine the fraction of investor sentiment unaccounted for by macroeconomic variables, for both individual and institutional respondents, following the Lemmon and Portniaguina (2006) method by conducting an Ordinary Least Squares (OLS) regression (see Appendix A for data sources):

$$C_{it} = \alpha_i + \beta_1 DEF_t + \beta_2 DIV_t + \beta_3 INF_t + \beta_4 IP_t + \beta_5 CONS_t + \beta_6 TBILL_t + \beta_7 URATE_t + \beta_8 NBER_t + \beta_9 DEF_{t-1} + \beta_{10} DIV_{t-1} + \beta_{11} INF_{t-1} + \beta_{12} IP_{t-1} + \beta_{13} CONS_{t-1} + \beta_{14} TBILL_{t-1} + \beta_{15} URATE_{t-1} + \beta_{16} NBER_{t-1} + \varepsilon_{it}$$

where
C_{it} = Value of confidence index i after month t (i_1 = OYC Individual, i_2 = VC Individual, i_3 = OYC Institutional, i_4 = VC Institutional)
α_i = Intercept coefficient for index i
β_j = Parameters to be estimated (j = 1 … 14)
DEF_t = Default spread after month t, measured as the difference between the yields to maturity on Moody's BAA-rated and AAA-rated bonds
DIV_t = Dividend yield after month t
INF_t = U.S. inflation percentage after month t
IP_t = Continuously compounded return of industrial production after month t
$CONS_t$ = Continuously compounded return of personal consumption after month t
$TBILL_t$ = Three-month U.S. treasury bill rate
$URATE_t$ = U.S. unemployment rate after month t

$NBER_t$ = NBER recession cycle status after month t (0 = growth, 1 = recession)
ε_{it} = The fraction of investor sentiment unaccounted for by macroeconomic variables
After the investor sentiment indicators for both methods are obtained (using the residuals for the second method), the returns on the S&P-500, NYSE Composite and NASDAQ Composite stock prices will be regressed on the changes of these indicators (with a one-month time lag) to determine whether changes in investor sentiment are reflected in stock prices and whether the subprime mortgage crisis has affected this relationship. To determine the relationship between ε_{it} and stock prices, the following model will be used (OLS regression), with t split up into before and during/after the crisis:

$\Delta S_t = \alpha_{i1} + \lambda_{i1} \Delta\varepsilon_{it-1} + \eta_{i1t-1}$ (with t for April 2001-June 2007)
$\Delta S_t = \alpha_{i2} + \lambda_{i2} \Delta\varepsilon_{it-1} + \eta_{i2t-1}$ (with t for July 2007-December 2010)

where
ΔS_t = Change in stock price after month t
α_{ic} = Intercept coefficient (i_1 = OYC Individual, i_2 = OYC Institutional, i_3 = VC Individual, i_4 = VC Institutional, i_5 = BAKER) (c = 1, 2)
λ_{ic} = Parameters to be estimated
$\Delta\varepsilon_{it}$ = Change in investor sentiment index i after month t
η_{ict} = Regression error term

As it is not unlikely that stock market returns and investor sentiment have a causal relationship that flows in both directions, the Granger causality will be applied before applying the model. The hypotheses are formulated to determine whether the investor sentiment-stock market returns relationship has changed with the subprime mortgage crisis. The null hypotheses state that this relationship is unchanged. Formally:

$\lambda_{i1} = \lambda_{i2}$ (i_1 = OYC Individual, i_2 = OYC Institutional, i_3 = VC Individual, i_4 = VC Institutional, i_5 = BAKER)

4. Data analysis

Tables 1 and 2 summarize the descriptive statistics and correlations for the confidence index data and the macroeconomic variables, with the sample broken down into the two subperiods April 2001 - June 2007 (pre-crisis) and July 2007 - December 2010 (crisis).

The confidence indexes have similar values for individual and institutional respondents before the crisis, and lie even closer together during the crisis. The means of OYC1 and OYC 2 have slightly decreased in the crisis (86.76 to 75.95 and 80.82 to 78.63), while the means of VC1 and VC2 have slightly increased (66.31 to 68.13 and 70.49 to 74.09). For all of the confidence indexes, the standard deviations have decreased. For OYC2 and VC2, they have almost been cut in half (5.71 to 2.53 and 10.24 to 5.94). Also, the standard deviations of INFL and URATE have increased significantly (0.86 to 2.06 and 0.58 to 2.20).

Table 1: Descriptive statistics

	Pre-crisis					Crisis				
	NOBS	Mean	Maximum	Minimum	Std. Dev.	NOBS	Mean	Maximum	Minimum	Std. Dev.
CONS	75	0.43	2.74	-1.32	0.49	44	0.19	1.20	-1.46	0.51
DEF	75	0.96	1.41	0.62	0.20	44	1.51	3.38	0.82	0.72
DIV	75	1.67	1.93	1.24	0.16	44	2.26	3.60	1.74	0.51
INFL	75	2.63	4.70	1.10	0.86	44	2.13	5.60	-2.10	2.06
IP	75	0.13	1.24	-2.00	0.55	44	-0.18	1.56	-4.21	1.09
NBER	75	0.11	1.00	0.00	0.31	44	0.41	1.00	0.00	0.50
TBILL	75	2.65	5.01	0.84	1.48	44	1.05	4.82	0.01	1.41
URATE	75	5.26	6.30	4.20	0.58	44	7.48	10.00	4.40	2.20
OYC1	75	86.76	95.62	78.30	3.89	44	75.95	80.85	70.75	2.56
OYC2	75	80.82	92.59	71.11	5.71	44	78.63	82.88	72.86	2.53
VC1	75	66.31	78.92	56.57	5.76	44	68.13	81.82	59.58	5.16
VC2	75	70.49	88.76	44.70	10.24	44	74.06	82.81	59.38	5.94

CONS, continuously compounded return of personal consumption; DEF, Default spread measured as the difference between the yields to maturity on Moody's BAA-rated and AAA-rated bonds; DIV, dividend yield; INFL, U.S. inflation percentage; IP, continuously compounded return of industrial production; NBER, the National Bureau of Economic Research recession cycle status; TBILL, three-month U.S. treasury bill rate; URATE, U.S. unemployment rate; OYC1, Individual One-Year Confidence Index; OYC2, Institutional One-Year Confidence Index; VC1, Individual Valuation Confidence Index; VC2, Institutional Valuation Confidence Index. The pre-crisis data are from April 2001 - June 2007, the crisis data are from July 2007 - December 2010.

Table 2: Correlations between confidence indexes and macroeconomic variables

Pre-crisis	CONS	DEF	DIV	INFL	IP	NBER	TBILL	URATE	OYC1	OYC2	VC1	VC2
CONS	1.00											
DEF	-0.06	1.00										
DIV	0.04	-0.01	1.00									
INFL	0.02	-0.51	0.25	1.00								
IP	0.03	0.04	0.15	-0.08	1.00							
NBER	-0.06	-0.22	-0.67	0.05	-0.38	1.00						
TBILL	-0.05	-0.32	0.19	0.62	-0.04	0.16	1.00					
URATE	0.03	0.39	0.14	-0.52	0.16	-0.50	-0.89	1.00				
OYC1	-0.02	0.29	-0.42	-0.26	-0.09	0.28	-0.61	0.46	1.00			
OYC2	0.00	0.36	-0.01	-0.10	0.09	-0.15	-0.26	0.40	0.48	1.00		
VC1	0.00	0.25	0.76	0.18	0.03	-0.41	0.13	0.13	-0.04	0.17	1.00	
VC2	0.09	-0.36	0.70	0.56	0.17	-0.50	0.33	-0.02	-0.28	0.30	0.50	1.00

Crisis	CONS	DEF	DIV	INFL	IP	NBER	TBILL	URATE	OYC1	OYC2	VC1	VC2
CONS	1.00											
DEF	-0.59	1.00										
DIV	-0.53	0.96	1.00									
INFL	-0.24	0.60	0.54	1.00								
IP	0.43	-0.57	-0.50	-0.61	1.00							
NBER	-0.47	0.71	0.66	0.67	-0.67	1.00						
TBILL	0.09	-0.32	-0.36	0.23	-0.05	0.02	1.00					
URATE	0.23	-0.17	-0.09	-0.60	0.43	-0.55	-0.80	1.00				
OYC1	0.27	-0.40	-0.33	-0.48	0.45	-0.70	-0.24	0.60	1.00			
OYC2	-0.32	0.39	0.28	0.61	-0.52	0.60	0.22	-0.57	-0.56	1.00		
VC1	-0.35	0.85	0.89	0.52	-0.47	0.54	-0.50	0.11	-0.10	0.14	1.00	
VC2	-0.36	0.65	0.60	0.90	-0.61	0.71	0.22	-0.61	-0.53	0.65	0.53	1.00

CONS, continuously compounded return of personal consumption; DEF, Default spread measured as the difference between the yields to maturity on Moody's BAA-rated and AAA-rated bonds; DIV, dividend yield; INFL, U.S. inflation percentage; IP, continuously compounded return of industrial production; NBER, the National Bureau of Economic Research recession cycle status; TBILL, three-month U.S. treasury bill rate; URATE, U.S. unemployment rate; OYC1, Individual One-Year Confidence Index; OYC2, Institutional One-Year Confidence Index; VC1, Individual Valuation Confidence Index; VC2, Institutional Valuation Confidence Index. The pre-crisis data are from April 2001 - June 2007, the crisis data are from July 2007 - December 2010.

The positive correlations between VC1 & VC2 and DIV are strong both before (0.76 for individual and 0.70 for institutional) and during (0.89 for individual and 0.60 for institutional) the crisis. A similar correlation exists between these indexes and DEF although it only exists post-crisis, with 0.85 for VC1 and 0.65 for VC2. Pre-crisis, these values are weaker and even negative, with 0.25 for VC1 and -0.36 for VC2. For INFL this is also true, with 0.18 and 0.56 pre-crisis and 0.52 and 0.90 during the crisis. A remarkable observation is the correlations between all confidence indexes and NBER. Their signs were switched in the crisis, and the correlations are much stronger (pre-crisis: 0.28, -0.15, -0.41, -0.50; crisis: -0.70, 0.60, 0.54, 0.71). Also, DEF and DIV have changed from a nonexistent correlation of -0.01 to a very strong correlation of 0.96.

Table 3 summarizes the adjusted R^2 values after regressing the confidence indexes on macroeconomic variables, before and after the crisis. A fair portion of the changes in these confidence indexes is explained by macroeconomic trends for the One-Year Confidence indexes. A larger portion of these changes is accounted for by macroeconomic trends for the Valuation Confidence indexes, for which the adjusted R^2 seem to have increased after the crisis.

Table 3: Adjusted R^2 values

	Before	During
OYC1	0.67	0.54
OYC2	0.42	0.50
VC1	0.73	0.89
VC2	0.77	0.86

OYC1, Individual One-Year Confidence Index; OYC2, Institutional One-Year Confidence Index; VC1, Individual Valuation Confidence Index; VC2, Institutional Valuation Confidence Index. All data are adjusted R^2 values from the first regression model.

The residuals from these regressions form the part of investor confidence that is not explained by macroeconomic variables, and will be used as investor sentiment data along with the data from BAKER (the data sets of residuals are named OYC1_Before, OYC1_During, etc).

As mentioned earlier, it is possible that stock market returns and investor sentiment both have a causal effect on each other. To further investigate this, the pairwise Granger causality test has been applied on all pairs of stock indexes and confidence index residuals (both before and after the crisis). Based on the recommendations of Lieuw and Khim (2004), the amount of time lags to use has been determined by using the Akaike Information Criterion (AIC), resulting in 11 time lags. The results are shown in Table 4.

Table 4: Results of the Granger-causality test

	Confidence index to stock price			Stock price to confidence index		
	SP500	NYSE	NASDAQ	SP500	NYSE	NASDAQ
OYC1_Before	0.7862	0.8327	0.8530	0.7365	0.8707	0.4856
OYC1_During	0.007***	0.0012***	0.0137**	0.6260	0.3359	0.3377
OYC2_Before	0.8395	0.8593	0.8929	0.8300	0.8889	0.5013
OYC2_During	0.4115	0.3594	0.4044	0.3467	0.1855	0.4852
VC1_Before	0.6878	0.8253	0.3097	0.8793	0.9342	0.7705
VC1_During	0.0418**	0.0124**	0.0327**	0.7314	0.5391	0.8210
VC2_Before	0.9113	0.9800	0.8359	0.8023	0.8283	0.6480
VC2_During	0.2434	0.2063	0.3713	0.6363	0.6005	0.2447
BAKER_Before	0.6221	0.8784	0.7985	0.1712	0.3242	0.1174
BAKER_During	0.1196	0.0882*	0.2145	0.0239**	0.0055***	0.0694*

OYC1, Individual One-Year Confidence Index; OYC2, Institutional One-Year Confidence Index; VC1, Individual Valuation Confidence Index; VC2, Institutional Valuation Confidence Index; BAKER, Baker & Wurgler Index. <Index>_Before are the results for the April 2001 - June 2007 period, <Index>_During are the results for the July 2007 - December 2010 period. *, ** and *** denote respectively 10%, 5% and 1% significance.

None of the investor sentiment indexes seem to Granger-cause (or be caused by) stock market price changes before the crisis. There is strong evidence that the Individual One-Year Confidence Index and the Individual Valuation Confidence Index exhibit predictive power over stock price changes in the crisis period. The reverse does not seem to occur, except for the BAKER index for which stock prices Granger-cause changes in investor sentiment data (and less so in the other direction). In conclusion, the results from the Individual One-Year Confidence Index and the Individual Valuation Confidence Index in the second set of regressions will be useful to determine whether the relationship between investor sentiment and stock price changes has changed in the subprime mortgage crisis. For the BAKER index, however, this is not the case, as the Granger-causality indicates that it is the stock prices causing changes in this index. Possible results from this regression would imply a change in how changes in stock market prices cause changes in the BAKER index.

Each stock index has been regressed separately on every confidence index (OLS regression) both before and during the crisis. Statistical significance is assessed using White standard errors, correcting for possible heteroskedasticity (White, 1980). The resulting two coefficients of every confidence index have been subjected to a two-tailed Wald test, to test the hypothesis that the relationship between investor sentiment and stock market prices has been unchanged by the crisis. The results of these tests are shown in Table 5.

Table 5: Wald Test results

	SP500	NYSE	NASDAQ
OYC1	0.0356**	0.0791*	0.0079***
OYC2	0.3123	0.3536	0.3686
VC1	0.7803	0.8191	0.2153
VC2	0.2895	0.1862	0.4490
BAKER	0.6055	0.9200	0.1911

SP500, Standard & Poor's 500 Index; NYSE, New York Stock Exchange Composite Index; NASDAQ, NASDAQ Composite Index; OYC1, Individual One-Year Confidence Index; OYC2, Institutional One-Year Confidence Index; VC1, Individual Valuation Confidence Index; VC2, Institutional Valuation Confidence Index; BAKER, Baker & Wurgler Index. *, ** and *** denote respectively 10%, 5% and 1% significance.

The only results that are statistically significant are those of the Individual One-Year Confidence index. The coefficients between this index and all of the investigated stock indexes have significantly increased, thus the null hypothesis for this index can be rejected. For the other indexes, there is no significant statistical evidence that the coefficients have changes. The null hypotheses for these indexes are not rejected.

5. Conclusions

To free the consumer confidence data from Yale Confidence of macroeconomic trends, the raw data has been regressed on a set of macroeconomic variables in the first model. A fair portion of the trends in these confidence data sets could be attributed to macroeconomic variables. What could not be attributed to these variables has been used as a measure of investor sentiment.

To determine the impact of this investor sentiment on stock market prices, the second model regressed changes in stock market prices on changes in the residuals from the first regression (and the BAKER data) from the month before. To investigate the direction of the possible causal relationships, the pairwise Granger-Causality test has been applied. This test has revealed that the Individual One-Year Confidence Index (OYC1) and the Individual Valuation Confidence Index (VC1) Granger-cause the changes in stock prices in the post-crisis period. No significant results have been found for the pre-crisis period. For the institutional indexes OYC2 and VC2, no clear Granger-causality was established. For the Baker & Wurgler Index (BAKER), however, reverse Granger-causality was established. The changes in the Baker & Wurgler Index are the result of the changes in stock market prices, making this index unsuitable as a predictor of stock market changes.

In regressing the stock price changes on the changes of all of these indexes, the regression coefficients for before and after the crisis have been compared. The null hypotheses state that these coefficients would be the same before and after the crisis. This holds for all of the used indexes, except for the Individual One-Year Confidence Index (OYC1). For this index, the coefficient has significantly increased after the crisis. Based on the One-Year Confidence results, the predictive power of individual investor sentiment in regard to stock market prices seems to have increased in the subprime mortgage crisis. The attitude that individual investors have on future stock prices now has a stronger impact on stock market returns. Although the Valuation Confidence exhibits predictive power over stock price changes, the strength of this relationship

has remained unchanged in the crisis. As the Valuation Confidence Index is based on contemporary sentiment and the One-Year Confidence Index uses expected returns after a year, the conclusion can be drawn that since the subprime mortgage crisis, individual investors have been basing their investment decisions more on their beliefs about the future than before.

For institutional investors, no clear causality or changes have been found for both of these indexes. How they perceive the current and future stock prices does not seem to have an influence on the behavior of stock markets. This result can be interpreted in two ways: either they seem to be less sensitive to attitudes and beliefs than their individual counterparts in investing in the stock market, or their attitudes and beliefs are correct projections of future economic activity. Schmeling (2006) has found evidence that investor sentiment exists for both types of investors, but that institutional investors take expected individual sentiment into account whereas individual investors do not employ expected institutional sentiment in their predictions.

6. Discussion and future research

Upon reviewing this paper, my supervisor at LSF made two suggestions to make the evidence found in this paper more robust. The first suggestion was related to the correlations of the macroeconomic variables that were used. Some of these variables had strong (> 0.50) correlations with each other, which leaves room for error in the regression results. To investigate this matter, some of the variables (DEF, DIV, NBER, URATE) were eliminated and the regression equations were re-estimated. The results were similar to those of the original equations, but weaker: no significant changes in coefficients for SP500 and NYSE, and a change that is significant at 5% level for NASDAQ.

The second suggestion was on sample sizes. 75 observations were used for the pre-crisis period, but only 44 observations were available for the crisis period. To make this equal, only the last 44 observations from the pre-crisis period were used. Re-estimating the equations with only this robustness check was not possible due to data availability. When both robustness checks were used simultaneously, the results were no longer significant. To determine whether this discrepancy was caused by biased results or lack of available data, these regressions should be re-estimated a few years from now. Another option would be to use different macroeconomic variables to replace the ones that have strong correlations with each other, but the variables that were used in this paper seem to cover most of the variables used in literature.

There is strong evidence that the concept of investor sentiment is one that should definitely be taken into consideration in regard to stock prices. As the previous literature is somewhat conflicting on the stability of the relationship between investor sentiment and stock prices, there is a lot of room for future research. Much of the existing literature on investor sentiment was based on either consumer confidence or institutional investor data. Based on the results of this paper, it would be wise to focus research efforts on the confidence of individual investors instead, as this makes much more sense than consumer confidence, and individual investors seem most prone to sentiments in their decision-making.

Appendix A: Reference list

Baker, M., & Wurgler, J. (2005). *Investor sentiment in the stock market.* Retrieved May 10, 2012, from http://www.nber.org/papers/w13189

Barberis, N. (2003). A survey of behavioral finance. In *Handbook of the Economics of Finance* (pp. 1053-1092). Elsevier Science.

Basu, S. (1977). Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: ATest of the Efficient Market Hypothesis. *The Journal of Finance, Vol. 32, No. 3* , 663-682.

Brown, G. W., & Cliff, M. T. (2005). Investor Sentiment and Asset Valuation. *Journal of Business, vol. 78, no. 2* .

Chung, S.-L., Hung, C.-H., & Yeh, C.-Y. (2012). When does investor sentiment predict stock returns? *Journal of Empirical Finance, Vol. 19* , 217-240.

Cuche, N. A., & Hess, M. K. (2000). Estimating Monthly GDP in a General Kalman Filter Framework: Evidence from Switzerland. *Economic & Financial Modelling* , Winter edition 153-193.

Estrella, A., & Mishkin, F. S. (1988). Predicting U.S. recessions: Financial variables as leading indicators. *The Review of Economics and Statistics, Vol. 80, No. 1* , 45-61.

Federal Reserve Bank of New York. (2011). *Crisis Timeline*. Retrieved May 14, 2012, from http://newyorkfed.org/research/global_economy/Crisis_Timeline.pdf

Lamont, O. A., & Thaler, R. H. (2003). Can the Market Add and Subtract? Mispricing in Tech Stock Carve-outs. *Journal of Political Economy, vol. 111, no. 2* .

Lemmon, M., & Portniaguina, E. (2006). Consumer Confidence and Asset Prices: Some Empirical Evidence. *The Review of Financial Studies, v19, n4* .

Liew, S., & Khim, V. (2004). Which Lag Length Selection Criteria Should We Employ? *Economics Bulletin, Vol. 3, No. 33* , 1-9.

Perriman, H., Ramsaran-Fowdar, R., & Baguant, P. (2011). *The impact of the global financial crisis on consumer behavior*. Retrieved May 15, 2012, from http://www.wbiconpro.com/06-Priya.pdf

Schmeling, M. (2006). *Institutional and Individual Sentiment: Smart Money and Noise Trader Risk.*

Shiller, R. J. (2000). *Irrational Exuberance.* Princeton: Princeton University Press.

Shiller, R. J. (1992). *Market Volatility.* Boston: MIT Press.

White, H. (1980). A heteroskedasticy-consistent covariance matrix estimator and a direct test for heteroskedasticy. *Econometrica, vol. 40, num. 4* , 817-838.

Zhang, C. (2008). *Defining, Modeling, and Measuring Investor Sentiment.* University of California, Berkeley.

Appendix B: Research data sources

Data	Data source	Literature source
Yale Confidence Indexes	Yale	(Zhang, 2008)
Baker and Wurgler Index	NYU Stern	(Baker & Wurgler, 2005)
Default spread	Moody's	(Lemmon & Portniaguina, 2006)
Dividend yield	Standard & Poor's 500	(Lemmon & Portniaguina, 2006)
Inflation	US. Bureau of Labor Statistics	(Lemmon & Portniaguina, 2006)
Industrial Production	Thomson Reuters Datastream	(Cuche & Hess, 2000)
Consumption	Thomson Reuters Datastream	(Lemmon & Portniaguina, 2006)
Treasury bill rate	Thomson Reuters Datastream	(Lemmon & Portniaguina, 2006)
Unemployment rate	U.S. Bureau of Labor Statistics	(Lemmon & Portniaguina, 2006)
NBER recession indicator	U.S. National Bureau of Economic Research	(Estrella & Mishkin, 1988)
SP500 stock index	Standard & Poor's 500	
NYSE Composite stock index	Yahoo! Finance	
NASDAQ Composite stock index	Yahoo! Finance	

Appendix C: Dutch translation of abstract

Aangezien de bestaande literatuur tegenstrijdig is met klassieke financiële theorieën met de suggestie dat 'investor sentiment' (emotie van investeerders) invloed heeft op aandeelprijzen, en er bewijs is dat suggereert dat dat deze relatie veranderd zou kunnen zijn door de kredietcrisis (sinds 2007), gebruikt dit paper verscheidene manieren om investor sentiment te meten voor zowel individuele als institutionele investeerders in de V.S. en bevrijdt het vervolgens van macro-economische trends. De Granger-causaliteitstest wordt toegepast om te bekijken of investor sentiment de oorzaak is van veranderingen in aandeelprijzen, en niet andersom. De veranderingen in aandeelprijzen worden vervolgens geregresseerd op veranderingen in de investor sentiment data vrij van macroeconomische trends. De resulterende regressiecoefficiënten van voor en tijdens de crisis worden daarna vergeleken om de impact van de kredietcrisis op de relatie tussen investor sentiment en aandeelprijzen te bekijken. Van de vijf investor sentiment indices die gebruikt worden, zijn er twee die significante causaliteit in de gewenste richting vertonen (de Individual One-Year Confidence Index en de Individual Valuation Confidence Index, beide van Yale). Voor de Individual One-Year Confidence Index is er sterk statistisch bewijs gevonden dat de voorspellingskracht van investor sentiment voor aandeelprijzen is toegenomen in de kredietcrisis.

Appendix D: Dutch self-reflection

In mijn studie Technische Bedrijfskunde wordt in veel vakken aandacht besteed aan finance en economie. Deze onderwerpen hebben al sinds mijn middelbare schooltijd mijn interesse, maar mijn verdieping erin is tot nu toe vrij beperkt gebleven. Ik ervaar bij opdrachten en vakken vaak gebrek aan motivatie omdat de onderwerpen simpelweg mijn interesse niet hebben. Aangezien dit een van de weinige gebieden is waar mijn interesse wel ligt, wist ik meteen dat het een goed idee was om mijn bacheloropdracht voor de Luxembourg School of Finance te doen toen ik de kans kreeg. Op een donderdag kreeg ik te horen dat ik hier welkom was, en de daaropvolgende maandag zat ik om 9 uur s'ochtends al op kantoor in Luxemburg.

Wat me hier vrij snel opviel is de erg professionele, serieuze werksfeer die hier aanwezig is. De LSF is erg jong (in 2002 opgericht, en sinds een paar jaar pas sterk aanwezig op de internationale markt), maar heeft

een reusachtig ambitieniveau. Men arriveert hier voor 9 uur, en verlaat het kantoor vaak pas na 6-7 uur. Elke donderdag heeft men hier een seminar-uurtje, waarbij iemand zijn of haar onderzoek indien gewenst kan presenteren en feedback van alle professoren en PhD-studenten kan krijgen hoe het verbeterd kan worden.

Men wilde hier graag dat ik onderzoek ging doen naar de financiële crisis. Om me te helpen met het bedenken van een geschikt onderzoeksonderwerp binnen deze crisis, kreeg ik gelijk een aantal relevante boeken om de nodige kennis op te doen. Het duurde dan ook niet lang of ik had mijn onderwerp gevonden: investor sentiment. Ik heb hierover met mijn interne begeleider (Thorsten Lehnert) en een aantal van de PhD-studenten hier gesproken, en kreeg alleen maar positieve feedback. Toen ben ik de literatuur ingedoken, en heb alle methodes die er besproken worden om investor sentiment te kwantificeren en analyseren uitgebreid bekeken en een selectie gemaakt voor de methodes die ik zou gaan gebruiken.

Aan de LSF heeft men toegang tot een aantal datastreams, waaronder de Thomson Reuters datastream. Hier viel erg veel relevante data uit te halen, die gecombineerd met een paar andere bronnen voldoende basis gaven om de regressies uit te voeren. De daadwerkelijke uitvoering van de regressies vereiste enig puzzelwerk. Ik kreeg het statistiekpakket Eviews aangereikt, het duurde een paar dagen om door te krijgen hoe ik precies met het programma moest omgaan. De regressies heb ik een aantal keer opnieuw moeten doen, vanwege kleine veranderingen in mijn onderzoeksmethodologie en het corrigeren voor statistische valkuilen. Om een Wald-toets met Eviews uit te kunnen voeren om de veranderingen in regressiecoefficiënten te onderzoeken, was het noodzakelijk om twee vergelijkingen (voor en tijdens de crisis) in één regressie op te kunnen nemen. Dit heb ik met behulp van een professor bij het LSF opgelost, die mij een techniek leerde om dit statistisch verantwoord te kunnen doen. In een laat stadium kwam ik erachter dat het gebruiken van zogenaamde “generated regressors” tot heteroskedasticiteit kan leiden, iets dat mogelijk ruimte laat voor fouten in de resultaten. Van de professoren en PhD-studenten aldaar heb ik meegekregen dat gebruik maken van de White standaardfouten dit probleem oplost. Wat ik als laatste tegenkwam was het feit dat ik Granger-causaliteit moest gebruiken alvorens de tweede regressie te doen om daadwerkelijke conclusies te kunnen trekken uit de gevonden resultaten.

Ik kijk terug op deze opdracht met het gevoel dat ik hier echt iets van geleerd heb. Voorheen was ik niet in staat om finance papers te lezen en daadwerkelijk te begrijpen, maar nu ik zelf een dergelijk onderzoek heb uitgevoerd is het allemaal een stuk beter te volgen. Het kunnen bepalen van correlaties en regressiecoefficiënten is sowieso iets waar ik nog veel aan ga hebben in mijn verdere studie- en werkaangelegenheden. Ik heb de begeleiding die ik bij deze opdracht gekregen heb als zeer prettig ervaren. Aan de LSF kreeg ik wanneer gevraagd heldere uitleg, en had men er geen probleem mee om even tijd voor me te maken. De feedback vanuit de UT (meneer Kroon) was straight-to-the-point, iets wat ik erg prettig vind. Al met al heb ik zo op een erg efficiënte manier deze opdracht kunnen invullen, wat mij grote voldoening geeft.