Using sentiment to predict GDP growth and stock returns

Guzman, Giselle C.

Economic Alchemy LLC

29 June 2007

Using Sentiment to Predict GDP Growth and Stock Returns

Giselle Guzmán
Columbia University[1,2]

Draft, Preliminary and Incomplete

This Version: June 29, 2007

---

[1] Prepared for publication in *The Making of National Forecasts,* edited by Lawrence R. Klein.

[2] I would like to thank Lawrence Klein, Joseph Stiglitz, Marshall Blume, Marc Giannoni, Charles Jones, and Lawrence Glosten for their insights and guidance. Thanks also to Randy Moore, William Dunkelberg, Richard Curtin, Lynn Franco, and Dennis Jacobe for generously providing data. All equations estimated with EViews Version 6 for Windows PC by Quantitative Micro Software, LLC. Any errors are solely my own.

Using Sentiment to Predict GDP Growth and Stock Returns

Giselle Guzmán
Columbia University

Abstract:

This study sheds new light on the question of whether or not sentiment surveys, and the expectations derived from them, are relevant to forecasting economic growth and stock returns, and whether they contain information that is orthogonal to macroeconomic and financial data. I examine 16 sentiment surveys of distinct respondent universes and employ the technique of principal components analysis to extract the common signals from the surveys. I show that the ability of different population groups to anticipate correctly economic growth and excess stock returns is not identical, implying that not all sentiment is the same, although there exist some common components. I demonstrate that sentiment surveys have significant predictive power for both GDP growth and excess stock returns, and that the results are robust to the inclusion of information pertaining to the macroeconomic environment and momentum. Furthermore, the findings reject the conventional wisdom that the effect of sentiment is apparent exclusively in small-capitalization stocks.

Using Sentiment to Predict GDP Growth and Stock Returns

Giselle Guzmán
Columbia University

Section 1 Introduction

The American Heritage Dictionary defines "sentiment" as "a thought, view, or attitude, especially one based mainly on emotion instead of reason." By the same token, it defines something that is "not endowed with reason" to be "irrational." Hence, "sentiment" is largely regarded as "emotional" and "irrational." Classical asset pricing theory makes no provision for such an irrational component in determining asset prices, particularly in long-run equilibrium. Yet, it remains a favorite statistic for financial media and popular press and is the source of endless commentary by market pundits and economists alike. Indeed, the financial press often credits or blames "sentiment" for a rising or falling stock market. If markets do, in fact, react to reports of changes in sentiment, then this indicates that the reality of asset pricing contradicts the theory of asset pricing. This suggests an oversight on the part of the academic literature in failing to give sentiment the importance it may warrant in the theory of asset pricing.

Academics have only recently begun to examine what role, if any, sentiment may have in the theory of asset pricing. However, consensus is lacking regarding its most basic characteristics. The literature remains divided not only about whether or not sentiment matters for asset prices, but also about what sentiment actually is, and how best to measure and incorporate it in a theoretical framework. I focus here on the empirical aspects of sentiment, its measurement, and its predictive power for the real economy as well as for financial markets.

Sentiment has no explicit role in traditional asset pricing models. The omission of sentiment from classical finance is rather curious, considering the key role played by emotion in the theories of Bentham (1781), one of the most influential early utilitarian philosophers. Bentham's concept of utility "…meant that property in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness…or … to prevent the happening of mischief, pain, evil, or unhappiness to the party whose interest is considered…" As Lowenstein (2000) notes, neoclassical economists later rendered the utility construct devoid of its emotional content in a process that "…culminated in the development of ordinal utility and the theory of revealed preference which construed utility as an index of preference rather than of happiness." Classical finance has evolved around the mathematical concepts of mean-variance optimization, rational maximization of preferences, equilibrium analysis, and no-arbitrage arguments, but it has largely neglected a key ingredient of financial markets: human emotion.

The pioneering work of Katona (1951, 1957, 1975) seeks to address the confluence of emotions and economics. His psychological approach to consumption prescribes that both *capacity* and *willingness* to buy are primary determinants of the consumption function. From this treatment one can infer that sentiment, i.e., something generally regarded as irrational, should be considered a bona fide component of expectations formation. Katona's theories build upon the notion of "animal spirits" put forth by Keynes (1936). Notable contributions to the theory of emotions in economics are made by Elster (1998), Lowenstein (2000), Thaler (2000), and Romer (2000). Romer succinctly echoes Katona's theories by stating that "…economists can usefully segregate decision mechanisms into two broad categories: those based on thoughts and those based on feelings…," and suggests that the profession should "…treat thoughts and feelings more symmetrically."

In this chapter, I aim to establish a definitive role for sentiment in macroeconomic forecasting and asset pricing by answering the following question: do the attitudinal data obtained from sentiment surveys contain any predictive power for economic growth and asset prices beyond the predictive information contained in macroeconomic and financial data? To answer this question, I examine 16 popular sentiment surveys of businesses and households, and test their ability to predict GDP growth as well as aggregate excess stock returns. I construct a composite sentiment factor using all of the sentiment survey indexes and then create separate factors for the sentiments of businesses and of households, for a total of three composite measures – All (APC), Business (BPC), and Household (HPC). I use the technique of principal components analysis to extract the common elements from sentiment surveys and create the composite factors. This signal-extraction technique allows a large collection of dynamic factors to be distilled into a few key measures that illustrate the joint effects of many popular surveys.

Section 2 presents a review of the related literature. Data and methodology are discussed in Section 3. The relation between sentiment indexes and macroeconomic factors is the focus of Section 4, as is the use of sentiment surveys in conjunction with the CQM model of Klein and Sojo (1989), a high-frequency model used to forecast GDP growth.[3] Section 5 investigates the predictive power of sentiment for excess returns of aggregate stock indexes, controlling for macroeconomic factors and lagged stock returns, and section 6 concludes.

Section 2 Literature Review

Section 2.1 Sentiment Measures

[Insert Section 2.1 here]

Section 2.2 Survey Data

[3] Please refer to the chapter "The Making of National Economic Forecasts – Introduction" by Lawrence R. Klein for more details regarding the CQM high frequency forecasting model for the United States.

The most obvious way to measure investor sentiment is by directly polling market participants and soliciting their opinions. Surveys ask respondents to report probabilistic expectations of significant personal financial or general economic events. Financial market participants can be broadly categorized as either Households or Businesses. Households participate in markets as both investors and consumers. Consumers influence stock prices not only by purchasing goods and services from publicly traded companies thereby affecting sales and reported earnings, but also because consumer spending represents approximately two-thirds of GDP in the U.S. With the increased popularity of discount brokerages and online investing over the past couple of decades, many consumers are now also individual investors involved in direct trading, in addition to participating in the stock market via mutual funds and 401Ks. Households, consisting of consumers and investors, have gained increased exposure to, and influence in, the stock market in recent years.

Businesses are also important market participants. They are the listed companies themselves, or the suppliers, customers, or strategic partners of the listed companies, and their economic health determines general economic growth as well as the discount rates used in asset valuation. Thus, today's stock market brings together Households and Businesses like never before, highlighting the need to identify a direct and sensible way to measure the perceptions, sentiments, and expectations of these market participants and determine whether or not they contain any predictive power for economic quantities of interest.

Over the years, survey data has had its share of detractors who sought to discredit its use in economic forecasting. Opponents argue that people do not always do as they say, and many economists dismiss the use of subjective data out of hand. Nevertheless, there are some legitimate concerns about the quality of the data elicited from surveys. Campbell (2004) points out that the most serious concern is whether respondents answer survey questions accurately. Most surveys cannot be used to track expectations of particular individuals through time, since they are series of cross-sections and not complete panels. Additionally, surveys are always subject to sampling error and other measurement issues. Additional concerns pertain to the manner in which questions are posed and responses elicited. Dominitz and Manski (2003) point out that certain phrases in survey questions such as "better off" may be subject to interpretation. Survey data are imprecise because they represent an attempt to construct a *quantitative* measure of human attitudes, which are inherently *qualitative*.

But, it is clear that observed choice data alone are insufficient for empirical analysis of decisions made when the information set is imperfect, as in the sense of Grossman and Stiglitz (1980), or incomplete. Manski (2004) contends that the assumption of rational expectations is implausible in decision-making with partial information. He advocates measuring expectations with survey data using subjective probabilities rather than the standard practice of revealed preference analysis, i.e., inferring decision processes from data on observed choices. The goal is to use self-reported data on expectations to relax or validate the assumptions regarding expectations that underlie economic models.

It appears that survey data may indeed be on the verge of a renaissance, with many researchers now taking an interest in sentiment surveys, and exploring their ability to explain how investors form expectations and their usefulness in forecasting the economy and asset returns.

Section 2.3 The Predictive Power of Survey Data – The Evidence So Far

Section 2.3.1 Consumer Sentiment – Macroeconomic Literature

Most of the early work on expectations derived from surveys examines the University of Michigan's Surveys of Consumers and its ability to predict consumer spending. Klein and Lansing (1955) find that survey questions on buying intentions, feelings of financial well-being and price expectations are predictive of consumer expenditures on durable goods.[4] Mueller (1963) reports that lagged values of the University of Michigan surveys have predictive power for household expenditures on durable and non-durable goods.

Some researchers contend that surveys lose their predictive power once other financial and macroeconomic variables enter the specification. Hymans, et. al. (1970) find that the University of Michigan Index of Consumer Sentiment (ICS) can forecast automotive spending, but that lagged values of income, consumer prices, and changes in stock prices can forecast the ICS. Mishkin (1978) finds that the University of Michigan ICS significantly predicts consumer expenditures on durable goods, but this relationship does not hold once financial variables are taken into account. Leeper (1992) uses a vector autoregression (VAR) framework to examine the relationship between consumer sentiment, industrial production, and unemployment, and also finds that the relationship significantly weakens once stock prices and T-bill rates are included in the analysis.

Garner (1991) asserts that consumer confidence indexes aid in forecasting aggregate consumption only during major economic or political events. Similarly, Throop (1992), using a five-variable vector-error-correction model (VECM), finds that in times of turbulence such as the Gulf War and the 1987 stock market crash, consumer sentiment can move independent of economic fundamentals, thus providing unique insights about future consumer expenditures. However, Throop noted that during normal periods, forecast results are slightly worse when sentiment is included in the specification than when it is omitted. The notion that sentiment is a particularly valuable forecasting tool during times of turmoil concurs with Katona's (1975) suggestion that the University of Michigan Surveys of Consumers reflect psychological factors that become pronounced during extraordinary periods of social, political, or economic upheaval.

---

[4] Klein and Lansing (1955) studied a "re-interview" sample of the 1953 Survey of Consumer Finances conducted by the Survey Research Center of the University of Michigan, a precursor to the University of Michigan's Index of Consumer Sentiment.

Matsusaka and Sbordone (1995) find that fluctuations in consumer sentiment account for between 13 percent and 26 percent of the variance of GNP innovations, even after controlling for a collection of economic indicators, demonstrating that expectations play a non-trivial role in forecasting output. Bram and Ludvigson (1998) run a horserace of the University of Michigan ICS versus the Conference Board's Consumer Confidence Index (CCI) and compare their relative abilities to forecast five categories of household expenditures: total, motor vehicles, all goods excluding motor vehicles, services, and durable goods excluding motor vehicles. The authors report that sentiment can help predict consumption, even after including control variables, and also suggest that consumer attitudes may provoke economic fluctuations.[5] Howrey (2001) finds that the University of Michigan ICS is a statistically significant predictor of real GDP growth and provides an informative signal about the probability of recession. Klein and Ozmucur (2002) and (2004) demonstrate that models incorporating sentiment surveys of consumers, producers, or managers to forecast economic quantities such as personal consumption expenditures, personal income, and industrial production perform significantly better than models that do not include the surveys.

Section 2.3.2 Surveys and the Stock Market – Asset Pricing Literature

Despite their potential methodological shortcomings, surveys uniquely provide a direct measure of investor expectations. Yet, most of the literature concerning the predictive power of surveys has focused on macroeconomic forecasting and limited use has been made of examining the predictive power of sentiment surveys in forecasting stock market returns. De Bondt (1993) considers the American Association of Individual Investors (AAII) survey and finds that the sentiment of small investors displays a bias towards extrapolation of past market trends. Otoo (1999) examines the relation between stock

[5] Bram and Ludvigson (1998) estimate the relation between the difference of logs in consumption and the ICS and CCI sentiment indexes, and include a vector of control variables that contains the lagged dependent variable, lagged growth in real labor income, lagged log first difference in the real S&P500 index, and lagged first difference of the three-month T-bill rate.

returns and the University of Michigan ICS and Conference Board CCI surveys, and reports that returns share a strong contemporaneous relation with the surveys, but lagged changes in sentiment have no explanatory power for stock returns. Fisher and Statman (2000) investigate the Merrill Lynch survey of sell-side strategists, the AAII survey of individual investors, and the Investors Intelligence (II) survey of investment newsletter writers, and conclude that the sentiments of these three groups of market participants are not identical.[6] Lee, Jiang, and Indro (2002) employ a GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model to examine the relation between the II survey and stock returns and report that sentiment is a significant factor in explaining both excess returns and the conditional volatility of returns.

Guzmán (2003) finds that the Union Bank of Switzerland/Gallup Index of Investor Optimism surveys have significantly more predictive power than either the University of Michigan ICS or the Conference Board CCI surveys. Brown and Cliff (2005) study the relationship between the II survey and market-implied pricing errors from an independent valuation model, and find that the relationship is positive. Additionally, they report that sentiment is negatively related to future returns over multi-year horizons. Charoenrook (2005) examines the University of Michigan ICS and finds a negative relation with future excess returns at horizons of one month and one year. In addition, the author reports that the predictive power of consumer sentiment appears to be unrelated to economic cycles or time-varying expected returns. Lemmon and Portniaguina (2006) investigate the relationship between returns and the University of Michigan ICS and Conference Board CCI, and determine that the surveys forecast returns of small stocks and stocks with low institutional ownership. Finally, Verma and Verma (2007) use the II survey as

[6] These three surveys were omitted from this study because they are of questionable value. The Merrill Lynch survey of sell-side strategists is likely to have a pronounced optimistic bias towards over-weighting stocks in its recommended asset allocation. The AAII survey suffers from self-selection bias (members can take the survey as often as they wish on the AAII website), while the II survey depends on a subjective classification of newsletter writers, which can be influenced by the personal opinions or cognitive biases of the newsletter readers who determine the classification.

a proxy for the sentiments of institutional investors and the AAII survey as a proxy for the sentiments of individual investors, and conclude that the former are more rational than the latter.

To my knowledge, this chapter is the first to examine a large collection of sentiment surveys of distinct respondent universes, extract their common signal, and test its ability to predict GDP growth and excess stock returns.

Section 3 Data and Methodology

[Insert Section 3.1 here]

This study examines the *changes* in survey-derived expectations and their relation to GDP growth and excess stock returns. Therefore, most of the survey data are transformed using the difference of logs. The exceptions are The Philadelphia Federal Reserve Business Outlook Survey, which is transformed using the first difference, and Union Bank of Switzerland/Gallup Index of Investor Optimism - Personal Financial and Economic indexes, which are transformed using the first relative difference.[7] Note that the Union Bank of Switzerland/Gallup survey was conducted sporadically from October 1996 through January 1999, and has been conducted monthly since February 1999.[8] For the purposes of this study, data from October 1996 through January 1999 are interpolated to create a monthly series of comparable length to the other series.[9, 10]

[7] This is because these series contained negative values, hence it was not possible to use difference of logs.
[8] As of January 2008, UBS and Gallup dissolved their partnership to conduct the surveys.
[9] Regression results were virtually identical using the interpolated series beginning in October 1996 and the non-interpolated monthly series beginning in February 1999.
[10] The sample period under study begins in February 1997 due to the availability of TIPS data, which are used to calculate the implied inflation expectation that is included among the indicators employed in constructing the macroeconomic factor.

Table 1 provides summary statistics for the transformed series of sentiment survey changes, quarterly observations from February 1997 to May 2007.[11] The time series properties are also presented. Autocorrelations for each series are provided at 1, 3, and 12 lags, and none display significant autocorrelation. Augmented Dickey-Fuller tests reject at better than the 1% level the hypothesis that any of the series has a unit root. Some of the data are highly correlated, hence it is intuitively appealing to extract the common elements from this group of surveys and test the predictive power of the shared components.

[Insert Equations 1-3 here]

Table 2 shows the eigenvectors and the variance proportions captured by the first three principal components of each group: All, Business, and Household. The first principal component of All surveys, APC1, explains 42.3% of the variance, while the first three principal components, APC1, APC2, and APC3, together capture 72.6% of the total variance of the system. The first principal component of Business surveys, BPC1, explains 52.9% of the variance, while the first three principal components, BPC1, BPC2, and BPC3 reflect 90.2% of the total variance of the Business group. The first principal component of Household surveys, HPC1, explains 64% of the variance, while the first three principal components, HPC1, HPC2, HPC3 collectively represent 89.8% of the total variance of the Household group.

A macroeconomic factor $M_{i,t}$ is also constructed using principal components analysis. The goal is to determine if sentiment surveys merely reflect macroeconomic information, as some researchers have postulated, or if the surveys contain unique information. A broad collection is formed of 30 macroeconomic variables that are generally regarded as indicative of the economic cycle, such as new

[11] Quarterly observations are calculated as quarterly averages.

orders, housing starts, inflation, unemployment, industrial production, etc. The set of 30 macroeconomic indicators, hypothesized to have *a priori* importance for economic growth and asset prices, is loosely based on Matsusaka and Sbordone (1995), Klein and Ozmucur (2002), and Stock and Watson (2002). I construct the composite macroeconomic factor $M_{i,t}$ using principal components at time t with each indicator $I_{i,,t-n}$ lagged appropriately to reflect its own n-period data reporting lag. For example, CONSUMER_CREDIT enters the information set with a two-period reporting lag, while INDUSTRIAL_PRODUCTION is reported with a one-period lag. Thus, the principal components are calculated with each variable lagged to reflect how it enters the agent's information set at time t. Let the *i-th* principal component of the macroeconomic indicators be denoted:

$$M_{j,t} = \sum_{i=1}^{16} \gamma_{ij} I_{i,t-n} \qquad \text{for } i = 1, 2, \ldots 30 \quad (4)$$

Table 1 presents descriptive statistics for the macroeconomic indicators, as well as their associated data reporting lags. Most of the indicators were transformed using the difference of logs, with the exception of ratios and yield data, which were transformed using first differences. Whenever possible, the pre-update series is used rather than the revised historical series, as the goal is to replicate as closely as possible the real-time information set that is available to market participants. The Augmented Dickey-Fuller tests reject at the 1% level the hypothesis that any series has a unit root, with the exception of the three-month U.S. Treasury Bill (TBILL_3M) and total non-farm employment (NON-FARM_EMPLOYMENT) for which the unit root hypothesis is rejected at the 5% level, and federal government net receipts as a percent of GDP (NET_GOV_RECEIPTS), for which the hypothesis cannot be rejected. The data are transformed and standardized prior to computing the principal components. Once again, quarterly average values are utilized.

Only the first principal component is employed in the analysis, and it captures 17.3% of the total variance of the system of macroeconomic indicators. The eigenvector of the first principal component of macroeconomic indicators is given by:

M1 = 0.057 NEW_ORDERS(t-2) + 0.046 UNFILLED_ORDERS(t-2) - 0.099 HOUSING_STARTS(t-1)
+ 0.093 CONSTRUCTION(t-2) - 0.079 BUILDING_PERMITS(t-1) - 0.132 HOURLY_EARNINGS(t-1)
+ 0.050 AVG_WEEKLY_HOURS(t-1) + 0.155 CPI(t-1) + 0.147 PPI(t-1) - 0.047 RETAIL_SALES(t-1)
- 0.025 TRADE-WEIGHTED_EXCHANGE_RATE(t-1) - 0.230 MONEY_SUPPLY(t-1)
- 0.143 CONSUMER_CREDIT(t-2) + 0.008 INVENTORY_SALES_RATIO(t-2) - 0.052 NET_GOV_RECEIPTS(t-1)
- 0.156 UNEMPLOYMENT_RATE(t-1) + 0.175 TIPS-IMPLIED_INFLATION(t-1) + 0.284 FED_FUNDS_RATE(t-1)
+ 0.280 PRIME_RATE(t-1) + 0.269 CORPORATE_BOND_RATE(t-1) + 0.328 TBILL_3M(t-1)
+ 0.060 EXPORT_IMPORT_RATIO(t-2) + 0.371 TBOND_YIELD_1YEAR(t-1) + 0.282 TBOND_YIELD_10YEAR(t-1)
+ 0.253 TBOND_YIELD_20YEAR(t-1) + 0.281 30-YR_FIXED_MORTGAGE(t-1)
+ 0.025 DISPOSABLE_PERSONAL_INCOME(t-1) + 0.161 INDUSTRIAL_PRODUCTION(t-1)
+ 0.200 NON-FARM_EMPLOYMENT(t-1) - 0.009 MANUFACTURING_AND_TRADE_SALES(t-1)

The eigenvector of the first principal component reveals that the factor is essentially a proxy for the slope of the yield curve. The coefficient loadings of the eigenvector elements are also the correlation coefficients between the principal component and the underlying variables.

Due to serial correlation in the residuals, Newey-West (1987) heteroscedasticity and autocorrelation consistent (HAC) standard errors are employed throughout the analysis.

Section 4 Sentiment and GDP Growth

I begin by testing the ability of the sentiment factors to explain future GDP growth. The baseline regression measures the relation between GDP growth at time t, lagged GDP growth, and the lagged composite macroeconomic factor. A fixed lag of one period and the PDL are each tested. The baseline model is:

[Insert Equation 5 here]

The baseline results for the fixed lag are presented in Panel A of Table 3 and the results for the PDL are presented in Panel B. The one-period lag specification results in an adjusted R-squared of 0.079, while the PDL specification gives an adjusted R-squared of 0.277. The quarterly average macroeconomic factor, i.e., the first principal component of the macroeconomic indicators, is denoted MPC1 in Table 3, and it is significant at the 10% level for the one-period lag specification.

Next, the model is augmented to test whether any of the three composite sentiment factors (All, Business, or Household) has predictive power for future GDP growth over the baseline equation. The inclusion of a composite macroeconomic factor in a model that tests the ability of sentiment to forecast an economic variable such as GDP growth efficiently addresses the concerns of researchers such as Mishkin (1978), Leeper (1992), Carroll, et.al. (1994), and Bram and Ludvigson (1998), who hypothesized that sentiment may be made redundant by macroeconomic and financial information. If sentiment is merely a reflection of macroeconomic and financial information, then the sentiment-augmented regression should not have any incremental predictive power over the baseline equation. The sentiment-augmented model is estimated as:

[Insert Equation 6 here]

In the interest of parsimony, only the first principal component is utilized. Panel C of Table 3 reveals that APC1, the composite factor for All sentiment, is significant at the 5% level for one lag. The sign of the coefficient is negative. However, Table 2 indicates that the eigenvector for APC1 is negative since all of the elements have negative coefficients. Thus, APC1 is positively predictive of future GDP growth at one lag, even after controlling for the persistence of GDP and a lagged composite macroeconomic factor. When sentiment is high, future GDP growth is high. The addition of APC1 to the model increases the adjusted R-squared by 5.8%. This result is driven mainly by the sentiment of Households rather than Businesses, since the specification with HPC1 is incrementally predictive, but the specification with

BPC1 is not. The Household sentiment factor, HPC1, is statistically significant at the one percent level, and increases the adjusted R-squared by 6.2% for the one-period fixed lag. Note again that the eigenvector of HPC1 is negative, hence the factor is positively predictive of future GDP growth. The reverse is true for the PDL specification, displayed in Panel D: BPC1 is statistically significant at the one percent level and adds 2% to the adjusted R-squared, while HPC1 and APC1 are not statistically significant.

Next, I examine the predictive power of the composite sentiment factors when used in conjunction with the Current Quarter Model (CQM) of Klein and Sojo (1989), for forecasting GDP growth in the United States. The baseline relationship is estimated as:

[Insert Equation 7 here]

Panel A of Table 4 presents the results of the baseline model of GDP growth regressed on the average of the CQM high frequency forecasts made throughout the quarter. The relevant null hypothesis for the baseline regression is $\beta_1 = 1$. If the CQM is a good forecasting model, it should almost perfectly explain GDP growth. Indeed, the coefficient on the CQM model forecasts is statistically significant at better than the 1% level, with $\beta_1 = 1.045$. The adjusted R-squared is 0.107.

The equation is then augmented with the composite sentiment factors to determine whether sentiment can improve the performance of the CQM model. In this specification, only the first principal component of the sentiment surveys is utilized. The macroeconomic factor is not included since this information would already be reflected in the CQM forecast. The sentiment-augmented model is estimated as:

[Insert Equation 8 here]

A fixed lag of zero (contemporaneous relation), one period, and the PDL are each tested. Panel B of Table 4 reveals that the composite factor for All sentiment, APC1, is significant in all specifications, adding as much as 10.7% to the adjusted R-squared. Once again, the coefficients for APC1 are negative, but the negative eigenvector indicates a positive relationship. The result appears to be driven in the short-run by the sentiments of Households, as HPC1 is significant contemporaneously, at one lag, and with the PDL specification, increasing the adjusted R-squared by as much as 11.4%. The sentiment of Businesses appears to have more effect at longer lags since only the PDL of BPC1 is significant, adding 5.7% to the adjusted R-squared. Both now-casting and forecasting of GDP growth are aided by the addition of sentiment data to the CQM model, since both the contemporaneous and lagged sentiment factors are statistically significant. The results suggest that macroeconomic forecasters should not hesitate to incorporate sentiment measures in their efforts to predict future GDP growth.

Section 5 Sentiment and Stock Returns

Do the sentiment factors have any predictive power for aggregate excess stock returns? In order to investigate this question, a baseline model is presented that controls for the composite macroeconomic factor and momentum, i.e., lagged stock returns. If sentiment is nothing more than a reflection of recent stock returns and macroeconomic information, then the sentiment-augmented model should not have any incremental predictive power over the baseline model. The baseline equation is estimated as:

[Insert Equation 9 here]

The results of the baseline model are presented in Panel A of Table 5. Cumulative three-month (one quarter), six-month (two-quarters), and nine-month (three quarters) excess returns (i.e., the gross return minus the risk-free rate, Rf_t) are examined for the S&P500 index (SPQ), the Russell 1000 Growth index (R1GQ), the Russell 1000 Value index (R1VQ), the Russell 2000 Growth index (R2GQ), and the Russell 2000 Value index (R2VQ), for a total of 15 test portfolios (five stock indexes and three time horizons, Q1, Q2, and Q3.) The Russell 1000 indexes contain large-capitalization stocks, whereas the Russell 2000 indexes contain small-capitalization stocks.

Next, the baseline model is augmented with the composite sentiment factors to determine whether or not sentiment has any incremental predictive power. In this specification, the *first three* principal components of the sentiment surveys are utilized, relying on the arguments of Stone (1947). The augmented model controls for lagged excess returns of the relevant portfolio and the lagged composite macroeconomic factor. The inclusion of a composite macroeconomic factor in a model that tests the ability of sentiment to forecast stock returns efficiently addresses the concerns of researchers such as Brown and Cliff (2005), Lemmon and Portniaguina (2006), and Verma and Verma (2007), who hypothesized that sentiment may be made redundant by macroeconomic and financial information. If the lagged sentiment factors have any incremental ability to predict stock returns, then the *increment* to the adjusted R-squared should be positive. The sentiment-augmented model is estimated as:

[Insert Equation 10 here]

The results are presented in Panels B, C, and D of Table 5 for the sentiment of the All, Business, and Household groups, respectively. A fixed lag of one period and the polynomial distributed lag are both tested, and both specifications have predictive power, but the results for the polynomial distributed lag are more robust. The robustness of the PDL specification indicates that the sentiment factor possesses non-linearity. Due to space limitations, only the results of the PDL are reported.

Panel B of the Table 5 reveals that the sentiment factors of All respondent groups have some modest predictive power, mostly, but not exclusively, for small-capitalization stocks. Large-capitalization value stocks have some limited predictability. The sentiment factors of All survey respondents show significant predictive power for small-capitalization stocks. The increments to the adjusted R-squared for the Russell 2000 Growth and Value indexes range from 3.6% for the cumulative two-quarter return on the Russell 2000 Growth index to 23.3% for the cumulative three-quarter return on the Russell 2000 Value index.

Panel C of Table 5 presents the results for the sentiment factors of the Business group. The Business sentiment factors show predictive power for all five of the major stock market averages. The sentiment-augmented equation for the cumulative three-quarter return on the S&P500 index is improved by 4.8% relative to the baseline equation. The increments to the adjusted R-squared for the Russell indexes range from an improvement of 1.4% for the cumulative three-quarter return of large-capitalization growth stocks to 14.5% for the cumulative three-quarter return of large-capitalization value stocks. Note that the significant coefficients in Panel C are positive. Table 2 reveals that the eigenvector of the third principal component of Business sentiment is mostly negative, and loads heavily on FED, the Philadelphia Federal Reserve Business Outlook Survey, which has a negative coefficient. This suggests that Business sentiment inversely anticipates the excess return on large-capitalization value and small-capitalization growth stocks. One possible interpretation is that business managers have a keen sense of the pulse of the economy. If managers detect improved business conditions, they may become more optimistic about future economic growth and respond to survey questions accordingly. As their optimism rises, their level of risk aversion declines, and thus they demand lower returns on their investments, creating a negative relation between changes in Business sentiment and future aggregate excess stock returns.

The results for the sentiment of Households are presented in Panel D of Table 5. The Household sentiment factors display significant predictive power for all portfolios except the large-capitalization growth stocks. The improvements to the adjusted R-squared range from 0.6% for the one-quarter return on the S&P500 to as much as 36.5% for the cumulative three-quarter return on small-capitalization value stocks. Note that the significant coefficients in Panel D are negative, and occur mostly for HPC2. Table 2 reveals that HPC2, the second principal component of Household sentiment, has mostly positive elements on the eigenvector, with the positive elements loading most heavily on the University of Michigan Index of Consumer Sentiment - Current Conditions Preliminary (MCP) and the Conference Board Consumer Confidence Index - Present Situation (CBP). The combination of the positive loadings on the principal component with the negative coefficient in the regression suggests a negative relation between changes in these surveys and subsequent excess returns. However, note that HPC2 also loads *negatively* (and heavily) on all three of the Union Bank of Switzerland/Gallup Indexes of Investor Optimism – Headline, Personal Financial, and Economic (UBS, UBP, and UBE). The combination of the negative loadings on the principal component with the negative coefficient in the regression suggests a positive relation between changes in the Union Bank of Switzerland/Gallup surveys and subsequent excess returns.

One interpretation could be that this dichotomy is consistent with systematic overreaction by naïve investors such as that postulated by De Bondt and Thaler (1989), with an associated subsequent return reversal. This explanation is plausible given that the respondent groups of the University of Michigan and Conference Board surveys are households of ordinary consumers, who may not be particularly adept at interpreting economic data or anticipating stock market trends. Conversely, the respondents to the Union Bank of Switzerland/Gallup surveys are *investor* households, with a minimum of $10,000 in investable assets. The minimum asset requirement of the Union Bank of Switzerland/Gallup surveys may act as a filtering mechanism, creating a strategic universe of respondents who are sophisticated in financial matters, pay attention to economic trends, and correctly anticipate the direction of the stock market.

Section 6 Conclusion

The economic magnitude of the predictability demonstrated in this chapter is significant. Consider the GDP growth regressions in Table 3. Panel C shows that the one-period lag specification for the All sentiment factor, APC1, has a coefficient of -0.232. From Table 1, note that the standard deviation of APC1 is 2.634. Multiplication of the coefficient and the standard deviation indicates that a one-standard deviation rise (decline) in APC1 predicts a decline (rise) of -0.611% in the following quarter's GDP growth. Similarly, from Panel D, a one-standard deviation rise (decline) in the Business sentiment factor, BPC1, predicts a decline (rise) of -0.205% in GDP growth over the following quarter.

Next, consider the excess stock returns regressions presented in Table 5. Panel B shows the results for the All sentiment factor, APC2. In the regression for R1VQ1, the one-quarter excess return on the Russell 1000 Value index, APC2 has a coefficient of -0.444. The standard deviation of APC2, given in Table 1, is 1.746. Multiplication of the regression coefficient and the standard deviation indicates that a one-standard deviation rise (decline) in APC2 predicts a decline (rise) of -0.775% in the following quarter's excess return on a broad portfolio of large-capitalization value stocks from its unconditional mean. In the regression for R2GQ3, the cumulative three-quarter excess return on the Russell 2000 Growth index, APC3 has a coefficient of -3.620. Table 1 shows that the standard deviation of APC3 is 1.386. This implies that a one-standard deviation rise (decline) in APC3 predicts a decline (rise) of -5.017% in the following nine-month's excess return on a broad portfolio of small-capitalization growth stocks from its unconditional mean. Conversely, in the same regression for R2GQ3, Panel C shows that the Business sentiment factor, BPC3, has a regression coefficient of 4.575. Multiplication of the coefficient for BPC3 with its standard deviation of 1.017 obtained from Table 1 indicates that one-standard deviation rise (decline) in BPC3 predicts a rise (decline) 4.653% in the following three-quarter's excess return on a broad portfolio of small-capitalization growth stocks from its unconditional mean.

Turning to the Household sentiment factors in Panel D, one can observe that in the regression for the cumulative three-quarter excess return on the Russell 1000 Value index, HPC2 has a coefficient of -2.378. The standard deviation of HPC2, given in Table 1, is 1.218. Multiplication of the regression coefficient with the standard deviation indicates that a one-standard deviation rise (decline) in HPC2 predicts a decline (rise) of -2.896% in the following nine-month's excess return on a broad portfolio of large-capitalization value stocks from its unconditional mean. In the regression for the cumulative three-quarter excess return on the Russell 2000 Value index, HPC2 has a coefficient of -4.590. Multiplying by the standard deviation of HPC2 implies that a one-standard deviation rise (decline) in HPC2 predicts a decline (rise) of -5.591% in the following three-quarter's excess return on a broad portfolio of small-capitalization value stocks from its unconditional mean.

The results presented herein shed new light on the question of whether or not sentiment surveys are relevant to forecasting economic growth and stock returns, and whether they contain information that is orthogonal to macroeconomic and financial data. One important benefit of survey data is that they are readily available on a high-frequency basis. Thus, researchers have at their disposal an important and relatively under-exploited tool for forecasting economic quantities and asset prices, as well as measuring expectations of different population groups. I have shown that sentiment surveys have significant predictive power for both GDP growth and excess stock returns, and that the result is robust to the inclusion of information pertaining to the macroeconomic environment and momentum. Additionally, while the sentiment surveys share some common predictive signals, the sentiments of different respondent universes can be distinguished, and have non-identical predictive power. Furthermore, the findings reject the conventional wisdom that sentiment affects only small-capitalization stocks. The results suggest that it would behoove researchers to incorporate sentiment in their forecasting models of economic growth and stock returns.

References

Baker, M. and J. Wurgler (2006). “Investor Sentiment and the Cross-Section of Stock Returns.” *The Journal of Finance*, Vol. LXI, No. 4, August 2006, pp. 1645-1679.

Bentham, J. (1781). “An Introduction to the Principles of Morals and Legislation.” Oxford, At the Claredon Press, London, New York, and Toronto, Henry Frowde. MCMVII.

Bram, J. and S. Ludvigson (1998). “Does Consumer Confidence Forecast Household Expenditure? A Sentiment Index Horse Race.” *FRBNY Economic Policy Review*, June 1998, pp. 59-77.

Brown, G.W. and M.T. Cliff. (2004). “Investor Sentiment and the Near-Term Stock Market.” *Journal of Empirical Finance*, Vol. 11, pp. 1-27.

Brown, G.W. and M.T. Cliff (2005). “Investor Sentiment and Asset Valuation.” *Journal of Business*, Vol. 78, No. 2, pp. 405-439.

Campbell, J. (2004). “Comment: Perspectives on Behavioral Finance: Does ‘Irrationality’ Disappear with Wealth? Evidence from Expectations and Actions” in *NBER Macroeconomics Annual 2003*, eds. Gertler, M. and Rogoff, K., The MIT Press, Massachusetts Institute of Technology, Cambridge, Massachusetts, pp.194-199.

Carroll, C.D., J.C. Fuhrer, and D.W. Wilcox (1994). “Does Consumer Sentiment Forecast Household Spending? If So, Why?” *The American Economic Review*, Vol. 84, No. 5, pp. 1397-1408.

Charoenrook, A. (2005). “Does Sentiment Matter?” Unpublished manuscript. The Owen Graduate School of Management, Vanderbilt University. http://www2.owen.vanderbilt.edu/fmrc/papers%20data /2005%20papers/does%20sentiment%20matter_aida.pdf

Chen, N., R. Kan, and M. Miller (1993). “Are the Discounts on Closed-End Funds a Sentiment Index?” *Journal of Finance*, Vol. 48, pp. 795-800.

De Bondt, W.F.M. (1993). “Betting on Trends: Intuitive Forecasts of Financial Risk and Return.” *International Journal of Forecasting*, Vol. 9, pp. 355-371.

De Bondt, W.F.M. and R. Thaler (1989). “Anomalies: A Mean-Reverting Walk Down Wall Street.” *Journal of Economic Perspectives*. Vol. 3, No. 1, Winter 1989, pp.189-202.

Dominitz, J. and C. Manski (2003). “How Should We Measure Consumer Confidence (Sentiment)? Evidence from the Michigan Survey of Consumers.” Working Paper 9926, National Bureau of Economic Research.

Elster, J. (1998). “Emotions and Economic Theory.” *Journal of Economic Literature*. Vol. XXXVI (March 1998), pp. 47-74.

Elton, E.J., M.J. Gruber, and J.A. Busse (1998). “Do Investors Care About Sentiment?” *Journal of Business,* Vol. 71, pp. 477-500.

Fisher, K. and M. Statman (2000). "Investor Sentiment and Stock Returns." *Financial Analysts Journal*, March/April 2000, Association for Investment Management and Research, pp. 16-23.

Garner, C. A. (1991). "Forecasting Consumer Spending: Should Economists Pay Attention to Consumer Confidence Surveys?" *Federal Reserve Bank of Kansas City Economic Review,* (May/June), pp.57-71.

Grossman, S. and J. Stiglitz (1980). "On the Impossibility of Informationally Efficient Markets." The American Economic Review, Vol. 70, No. 3, pp. 393-408.

Guzmán, G.C. (2003). "GDP Growth Signals, Investor Psychology, and Hedging Pressure: A Multivariate Approach to Forecasting Returns on the S&P500 Index." Unpublished manuscript. The Wharton School, University of Pennsylvania. April, 2003.

Howrey, E.P. (2001). "The Predictive Power of the Index of Consumer Sentiment." *Brookings Papers on Economic Activity*, Vol. 2001, No. 1, pp. 175-207.

Hymans, S.H., G. Ackley and F.T. Juster (1970). "Consumer Durable Spending: Explanation and Prediction." *Brookings Papers on Economic Activity*, Vol. 1970, No. 2, pp. 173-206.

Katona, G. (1951). *Psychological Analysis of Economic Behavior*, McGraw-Hill, New York.

Katona, G. (1957). "Federal Reserve Board Committee Reports on Consumer Expectations and Savings Statistics." *Review of Economics and Statistics*, Vol. 39, 40-46.

Katona, G. (1975). *Psychological Economics*. Elsevier, New York.

Keynes, J.M. (1936). *The General Theory of Employment, Interest, and Money,* London: Macmillan for the Royal Economic Society.

Klein, L.R. and S. Ozmucur (2002). "The Predictive Power of Survey Results in Macroeconomic Analysis." in Wladyslaw Welfe, ed. *Macromodels 2001*, Chair of Econometric Models and Forecasts, University of Lodz, Lodz, 2002, pp. 181-197.

Klein, L.R. and S. Ozmucur (2004). "Some Possibilities for Indicator Analysis in Economic Forecasting." in Pami Dua, ed. *Business Cycles and Economic Growth: An Analysis Using Leading Indicators*, Oxford University Press, pp. 243-257.

Klein, L.R. and E. Sojo (1989). "Combinations of High and Low Frequency Data in Macroeconometric Models." in *Economics in Theory and Practice: An Eclectic Approach*, eds. Klein, L.R. and Marquez, J., Kluwer, Dordrecht, 1989, pp. 3-16.

Klein, L.R. and J.B. Lansing (1955). "Decisions to Purchase Consumer Durable Goods." *Journal of Marketing*, Vol. XX (October 1955), pp. 109-132.

Lee, C.M.C., A. Shleifer, and R.H. Thaler (1991). "Investor Sentiment and the Closed-End Fund Puzzle." *The Journal of Finance*, Vol. 46, No. 1, pp. 75-109.

Lee, W., C. Jiang, and D.C. Indro (2002). "Stock Market Volatility, Excess Returns, and the Role of Investor Sentiment." *Journal of Banking & Finance*, Vol. 26, pp. 2277-2299.

Leeper, E. M. (1992). "Consumer Attitudes: King for a Day". *Federal Reserve Bank of Atlanta Economic Review*, 77(4), pp. 1-15.

Lemmon, M. and E. Portniaguina (2006). "Consumer Confidence and Asset Prices: Some Empirical Evidence." *The Review of Financial Studies*, Vol. 19, No. 4, pp. 1499-1529.

Lowenstein, G. (2000). "Emotions in Economic Theory and Economic Behavior." *The American Economic Review*, May 2000, 90, 2, p. 426-432.

Malkiel, B.G. (1977). "The Valuation of Closed-End Investment Company Shares." *Journal of Finance*, Vol. 32, pp. 847-859.

Manski, C. (2004). "Measuring Expectations." *Econometrica*, Vol. 72, 1329-1376.

Matsusaka, J.G. and A.M. Sbordone (1995). "Consumer Confidence and Economic Fluctuations." *Economic Inquiry*, Vol. XXXIII, April 1995, pp. 296-318.

Mishkin, F.S. (1978). "Consumer Sentiment and Spending on Durable Goods." *Brookings Papers on Economic Activity*, 1, pp 217-232.

Mueller, E. (1963). "Ten Years of Consumer Attitude Surveys: Their Forecasting Record." *Journal of the American Statistical Association*, Vol. 58 (304), pp. 899-917.

Newey, W. and K. West (1987). "A Simple, Positive Definite, Heteroscedastic and Autocorrelation Consistent Covariance Matrix." *Econometrica*, Vol. 55, 703-708.

Otoo, M.W. (1999). "Consumer Sentiment and the Stock Market." Board of Governors of the Federal Reserve System.

Romer, P.M. (2000). "Thinking and Feeling." *The American Economic Review*, May 2000, Vol. 90, No. 2, pp. 439-443.

Stock, J.H. and M.W. Watson (2002). "Macroeconomic Forecasting Using Diffusion Indexes." *Journal of Business & Economic Statistics*, April 2002, Vol. 20, No. 2, pp. 147-162.

Stone, R. (1947). "On the Interdependence of Blocks of Transactions." *Supplement to the Journal of the Royal Statistical Society*, Vol. 9, No. 1, pp. 1-45.

Swaminathan, B. (1996). "Time-Varying Expected Small Firm Returns and Closed-End Fund Discounts." *Review of Financial Studies*, Vol. 9, pp. 845-887.

Thaler, R.H. (2000). "From Homo Economicus to Homo Sapiens." *Journal of Economic Perspectives*, Vol. 14, No. 1, Winter 2000, pp. 133-141.

Throop, A.W. (1992). "Consumer Sentiment: Its Causes and Effects." *Federal Reserve Bank of San Francisco Economic Review* 1, pp. 35-59.

Verma, R. and P. Verma (2007). "Are Survey Forecasts of Individual and Institutional Investor Sentiments Rational?" *International Review of Financial Analysis* (2007), doi:10.1016/j.irfa.2007.04.001.

Wiesenberger, A. (1946). *Investment Companies Services*. New York, NY: Warren, Gorham, and Lamont, (1943-1991).

Zweig, M.E. (1973). "An Investor Expectations Stock Price Predictive Model Using Closed-End Fund Premiums." *Journal of Finance*, Vol. 28, pp. 67-78.

[Insert Tables 1-5 here]