Mendel University in Brno

MENDELU Working Papers
in Business and Economics
53/2015

# Sentiment and blue-chip returns. Firm level evidence from a dynamic threshold model

Jaroslav Bukovina

**MENDELU Working Papers in Business and Economics**
Research Centre
Faculty of Business and Economics
Mendel University in Brno
Zemědělská 1, 613 00 Brno
Czech Republic
http://vyzc.pef.mendelu.cz/en
+420 545 132 605

**Citation**

Bukovina, J. (2015). Sentiment and blue-chip returns. Firm level evidence from a dynamic threshold model. *MENDELU Working Papers in Business and Economics* 53/2015. Mendel University in Brno. Cited from: http://ideas.repec.org/s/men/wpaper.html

**Abstract**

Jaroslav Bukovina: **Sentiment and blue-chip returns. Firm level evidence from a dynamic threshold model**

The higher impact of sentiment on small and young companies due to higher transaction and information costs is a generally agreed notion. In contrast, the paper is focused on the relationship between sentiment and blue-chip stocks. The author contributes with the proposal of theoretical background that blue-chip stocks can reflect sentiment in periods of occasional excessive social activity despite the existence of smart-money investors. The author employs the appropriate empirical methods, specifically dynamic threshold model, to analyze such a relationship because standard regression is not appropriate. In addition, the paper contribution is the employment of unique data of Facebook activity as a sentiment gauge. Overall, the author finds the high level of social activity connected with negative sentiment with inverse but occasional influence at stock price returns.

**Key words**

Blue-chip companies, sentiment, Facebook, dynamic threshold effects

**JEL:** G14, G19

**Contacts**

Jaroslav Bukovina, Department of Finance, Faculty of Business and Economics, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic, e-mail: jaroslav.bukovina@mendelu.cz.

**Acknowledgements**

This paper was supported by the Internal Grant Agency of PEF Mendelu, Grant No. PEF_DP_2015_003.

## 1 Introduction

In general, efficient market models define the new information about the fundamentals of a security as the only source of its volatility. Stock prices are near fundamental values due to activity of rational arbitrageurs (Friedman, 1953; Fama, 1965). Since 80´s, observations of anomalies and especially excess volatility have been detected at capital markets, which occurs for no fundamental reason (Le Roy & Porter, 1981; Shiller, 1981, 2003). The efficient market hypothesis is not able to successfully explain it.

Behavioral finance argues that social mood of society or investors' sentiment can be a source of excess volatility. Therefore, behavioral finance has augmented the standard model of finance with assumptions of investors, who are subject to sentiment and existence of limits to arbitrage (De Long et al. 1990; Shleifer & Vishny, 1997; Barberis, Shleifer & Vishny. 1998).

Empirical studies try to define the impact of sentiment more closely. Edmans, García & Norli (2007) and Kaplanski & Levy, (2010a; 2010b) study the specific "sentiment-bearing" situations like famous sport events or aviation disasters which affect the asset pricing. Baker & Wurgler (2007) define the classes of stocks more prone to be affected by sentiment. Specifically, stocks of low capitalization, young and non dividend paying companies can be more sensitive to investors' sentiment, because of higher information and transaction costs (Baker & Wurgler, 2007).

In contrast, the purpose of this paper is the analysis of the relationship between sentiment and blue-chip companies. The contribution of the paper is the employment of unique data from social media Facebook which represents the sentiment of a society related to the specific blue-chip company.

Blue-chip companies (bond-like stocks) are companies with long earning history, tangible assets and stable dividends. There should be the lower influence of sentiment in comparison with small and young firms because the limits in arbitrage are lower. However, the unique data of social activity from social media Facebook can reveal new relationships between sentiment and publicly traded brand-name companies. In the context of the paper, the sentiment is a social activity or social mood tracked by Facebook. Nofsinger (2005) implies the social mood as influential to financial variables. Renn (1990) summarizes the factors of risk perception. Among others, it is judgment of reference groups (e.g. an opinion of a society) or social amplification of risk-related information due to media coverage. Nowadays, it is the web and social media coverage as well. Kaplansky & Levy (2010a) summarize the market sentiment as a general notion which incorporates any misperception that can lead to mispricing.

The theoretical background about the impact of sentiment at blue-chips can be addressed as follows. The blue-chip companies are well established in the society and in the collective mind of the public. These companies shape our everyday reality, e.g. because of the products we use every day. Let's imagine there is an unexpected sentiment-bearing information[1] about the specific company, not the regular information like earnings announcements. If this information is significantly amplified by activity of society through the press, web and social media then this sentiment information can drive the stock of the company. The sentiment issue here is the herding behavior of society based on the overreaction on the information. It has a psychological foundation in the availability heuristic (Tversky & Kahneman, 1973; Schwartz, 1998). Brand-name companies are well-known in the society and information is easily connected with them. The activity of society and media makes the information more visible at least in a short run. The perception of information is more vivid and intense.

The less rational investors or noise-traders can be easily influenced by the amplified activity of society. In opposite, smart-money investors should drive the stock back to fundamentals. However, to follow the studies of De Long et al. (1990), Shleifer & Vishny, (1997), Barberis, Shleifer & Vishny. (1998), arbitrageurs can be limited. In the context of the paper, smart-money investors do not know which information will be the one which cause the overreaction or how long it will last. In addition, they have to bet against the opinion of a society. Therefore, arbitrageurs who bet against mispricing run the risk, at least in a very short run, that overreaction becomes more extreme and prices can move further away from fundamental value.

However, this paper does not define the deviation in the stock price or the range of mispricing. The author addresses the simple idea. If social activity due to sentiment information overcomes some level (estimated by a threshold below), it can influence the stock price. It is in contrast with the efficient market hypothesis which generally argues that the information about fundamentals is the only source of volatility.

The author employs the unique data of social media Facebook as a proxy for social activity and sentiment related to the specific company. The author has been inspired by studies of Bollen, Mao & Zeng, (2011) who forecasted the Dow Jones Index based on mood of social network Twitter. Sprenger & Welpe (2011) propose the online stock forums of Twitter as an alternative to traditional media sources. Choi & Varian (2009) provide the early indicators of consumer spending based on Google search queries. Da, Engelberg & Gao (2011, 2015) propose new sentiment indices based on

[1] A customer's death announcement due to product usage which is not proved by fact.

search volume data. These studies analyze the impact of sentiment at aggregate level (stock index, portfolios, households or consumer groups). This paper puts an intense focus at the micro level, specifically to the analysis of sentiment at the level of individual blue-chip companies.

In general, social media can be seen as a virtual mirror of our society. In the context of this paper, every well-known company has a social network profile, where millions of users share their opinions. The advantage of social media is the ability to track the behavior of users according to specific information. In addition, social media contain especially the users' mood and sentiment. The regular or expected information connected with fundamentals like company earnings are not an issue here. The rest of the paper is addressed as follows. The next section describes the methodology and data. Section 3 provides the results and the last section summarizes the discussion and conclusion.

## 2 Methodology and Data

This paper studies the impact of excessive social activity on 21 large-cap companies which belongs among the 100 biggest international companies according to market capitalization. An every company in sample is traded at US capital market. Analyzed period has 94 weeks (22 months), from February 2013 till December 2014. The time period is given based on availability of social media data. This paper provides the unique analysis because it employs the data of Facebook as a gauge of social activity and sentiment in the society. However, bond-like stocks in general are less influenced by sentiment as is proposed by Baker & Wurgler (2007). Therefore, the definition of proper variables which captures the company fundamentals in the model is important as well.

The author considers the volume of trade (liquidity) and stock indices as the only reasonable variables which can capture company fundamentals in a short run. Changes in the liquidity can represent the important signal about changes of fundamentals and the interest of investors. The liquidity is important too, especially in the view of current monetary policy because US equity market has been significantly influenced by unconventional monetary tools like quantitative easing. Volume of trade is measured at the level of the firm. Regarding the influence of macro environment at the individual company, the author applies the stock index S&P 500 which is constructed as the representative indicator of US economy. The impact of stock indices on individual stock prices has been shown by King (1966), Livingstone (1977) or Roll (1992).

The author applies the liquidity and stock price index as a proxy variable for fundamentals in the short run, but these variables are control variables for the investors' sentiment or current market mood as well. Positive or negative investors' sentiment is reflected in the trading volume Baker &

Wurgler (2007) and current mood of the whole market is captured by stock indices as well. This situation makes the results about the impact of sentiment measured by social media data more robust. Specifically, social media data captures occasional exogenous events which are not reflected in capital market variables.

### 2.1 Model

According to proposed theory, blue-chip stocks can be driven by sentiment only in periods of occasional excessive activity. In the model construction, the author follows the methodology of event studies, like Brown & Warner (1985), Kamstra, Kramer & Levi (2003) and Kaplanski & Levy (2010a; 2010b). Specifically, the author considers the excessive social activity as an event due to sentiment information. This situation requires appropriate empirical methods. The standard panel[2] model $y_{it} = \alpha + X'_{it}\beta + \varepsilon_{it}$ is not appropriate. It averages across all observations, therefore it cannot capture observations which fall into discrete classes. Specifically, the proposed theory requires a method, which can identify the periods of excessive social activity. This problem can be addressed by a threshold regression technique for panels proposed by Hansen (1999) extended for dynamic relationships to deal with the endogeneity by Caner & Hansen (2004) and applied by Bick (2010) and Kremer et al. (2012). The dynamic model with threshold effects which captures the excessive social activity considered as an event can be addressed as:

$$ret_{it} = \alpha_i + \delta_m sent'_{it} + \beta_m(sent'_{it} \leq \gamma) + \beta_n(sent'_{it} > \gamma) + \varepsilon_{it}, \tag{1}$$

or in the compact representation for threshold variable $\gamma$ only:

$$sent'_{it} = \begin{pmatrix} sent'_{it} \leq \gamma \\ sent'_{it} > \gamma \end{pmatrix}, \tag{2}$$

and if $\beta = (\beta_m; \beta_n)$ and $\alpha_i = 0$ then (1) augmented with the rest of variables equals:

$$ret_{it} = ret_{it-1} + \delta_m sent'_{it} + \beta sent'_{it} + \beta_k liq'_{it} + \beta_l ind'_{it} + D_{1,\dots,21} + \varepsilon_{it}, \tag{3}$$

where subscripts $i = 1, \dots, N$ represents the firm and $t = 1, \dots, T$, index the time, $ret_{it}$ denotes the stock return, $ret_{it-1}$ is a lagged dependent variable, *sent* captures the social activity tracked by Facebook so called *talking about* where $\delta_m$ is the estimate of regime dependent regressor and $\beta$ is threshold variable, specifically the marginal effect of sentiment on stock return, $liq_{it}$ is the liquidity or volume of trade, $ind_{it}$ represents the stock index S&P 500, $D_{1,\dots,21}$ denotes 21 time dummies for

[2] The author applies panel data due to nature of analysis for observations of individual companies in time

every month[3] with aim to capture unobserved changes in sentiment through the time and $\varepsilon_{it}$ is the error term.

In the context of the paper, when social activity in the society overcome threshold $\gamma$ then it drives stock prices. Observations of social activity are divided into two "regimes" depending on whether the variable *sent* is higher or lower than threshold $\gamma$. The "regimes" are defined by different regression slopes of $\beta_m$ and $\beta_n$ (eq. 1). The computational issues and technical details of dynamic threshold effects in this paper follow the approach provided by Hansen (1999), Caner & Hansen (2004), Bick (2010) and Kremer, Bick & Nautz (2012).

### 2.2 Data

Data of stock prices, volume of trade for individual firm and stock price index S&P 500 have been collected from the Yahoo! Finance database. Data of Facebook activity, (*talking about*) have been collected by algorithm. *Talking about* can be described as the amount of comments, likes and sharing of information among users. Facebook data are unique because it is not publicly available. Facebook data are in quantitative form, therefore, there is no need to apply the statistic standardization of qualitative data to quantitative form in a scale range proposed by Kapounek & Lacina (2011).

Only companies which belong among 100 biggest international companies according to market capitalization are analyzed. The final sample consist of 25 companies. Four companies are excluded from the analysis because they are not traded in US capital markets. The final sample is given according to benchmark which represents the amount of "likes" on the Facebook profile of a firm. Only companies where the amount of likes is higher than 75% quantile are a part of the sample. The amount of "likes" relatively represents the amount of followers (people who follow the activity of individual firms). This benchmark diminishes the bias that excessive social activity will be tracked by Facebook company profile which has only a few hundreds or few thousands of followers. Small groups of followers can excessively react to information which do not have the all-society impact.

All data are expressed as average value per week. The raw data were transformed to logarithms and then to first differences.

[3] Analysed period has 22 months but one month has been removed due to collinearity.

## 3 Results

Table 1 summarizes the details of threshold estimate, Table 2 provides the regression estimates of equation 3 and Figure 1 shows the graphical results of threshold variable *talking about.*

The threshold has been estimated at level $0.1394 \doteq 14\% \text{ week change in Facebook activity}$. In the context of the paper, when social activity in Facebook profile of specific company reaches the 14% change in comparison with previous week the sentiment information is amplified in society in such level that it can influence the stock price returns.

According to proposed theory, there should be the occasional impact of excessive social activity at blue-chip companies. We can consider the impact of social activity as occasional because aproximatelly 2/3 of observations are below the threshold. However, the threshold estimate 14 % is quite low to be considered as the existence of excessive social activity. The reason of this low threshold can be connected with employment of data in the form of the first differences and better rationale is provided in the Table 2 and by the graphical result in Figure 2.

Table 1: Threshold estimate

| | |
|---|---|
| Threshold estimate: | 14 % |
| Confidence interval: | [13 - 36 %] |
| *Number of observations:* | |
| sentiment ≤ 14 % | 1 260 |
| sentiment > 14 % | 672 |

Estimates of regime dependent variable *sent* show the negative impact of regressor $\delta_m'$ across the all cross-sections and the negative impact of threshold variable $\beta'$ above the level 14 %. This threshold has to be explained as whenever there is the increase in the Facebook activity there is a negative sentiment related to specific companies. Figure 1 clearly shows that below the threshold, there is fall in social activity. When it starts to increase, it soon reaches and overcome the threshold level 14 % with negative influence, according to the regression results in Table 2.

The lagged $return_{t-1}$ is not significant. Stock index *S&P 500* has positive, statistically significant and the largest influence at stock returns. It was an expected result in line with asset pricing models like CAPM or current after-crisis period where new liquidity from FED caused the growth period in stock markets and individual asset prices as well. *Liquidity* or *trade volume* is significant with a positive sign. It can be explained in the light of unconventional monetary policy as in the case of S&P 500.

However, a better explanation in the spirit of De Long et al. (1990) is the idea of noise-trader activity. Less rational investors mismatch the sentiment as a signal and smart-money investors tries to drive the stock prices back to fundamentals due to potential negative deviation caused by the excessive negative sentiment in the society. Therefore, there is the positive impact on liquidity. Table 2 shows the result only for the dummy for the 13th week, which captures the unobserved negative impact at stock returns in time. The rest of the dummies is not statistically significant and they were removed from the model.

Table 2: Regression estimates

| $return_t$ | coefficient | stand. error |
|---|---|---|
| *Regime-dependent:* | | |
| $(\widehat{\delta_m'})$ $sentiment_t$ | -0.006*** | 0.002 |
| $(\widehat{\beta'})$ sentiment ≤ 14 % | -0.001 | 0.002 |
| $(\widehat{\beta'})$ sentiment > 14 % | -0.004*** | 0.002 |
| *Regime-independent:* | | |
| $return_{t-1}$ | 0.007 | 0.031 |
| S&P $500_t$ | 0.956*** | 0.048 |
| $liquidity_t$ | 0.009*** | 0.003 |
| $dummy_{13}$ | -0.005* | 0.003 |

***/**/* is 1/5/10 % significance level

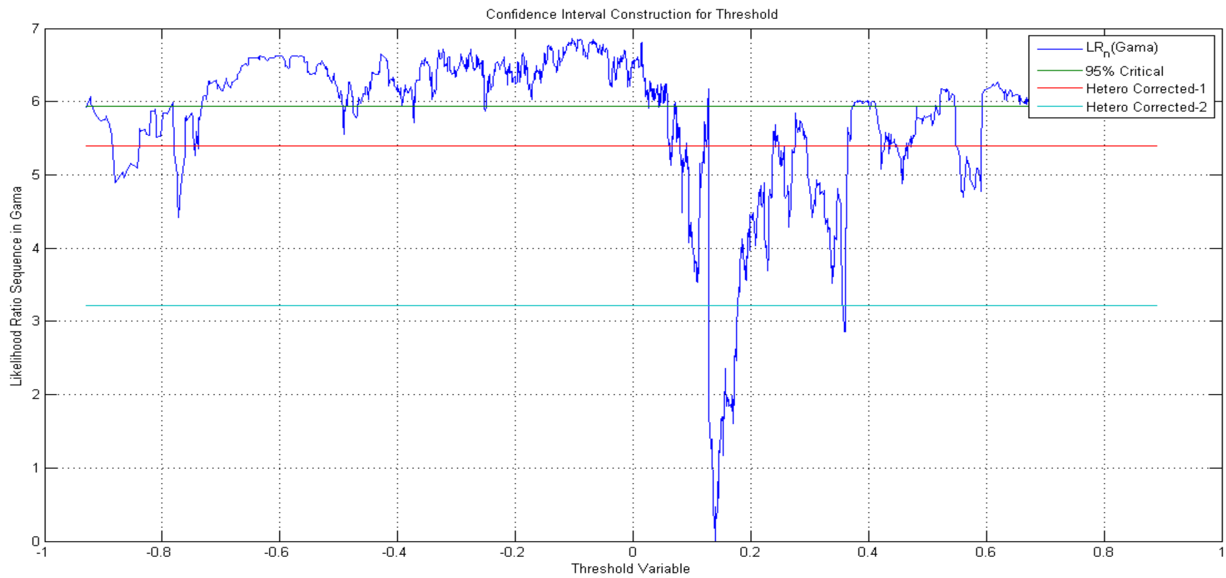


Figure 1: Threshold variable

The dynamic threshold estimation is connected with instrumental variables. Instruments are lags of dependent variable (Arellano & Bover, 1995). The author has applied four lags of the dependent variable (in total one month lag) because this paper employs weekly data. Either the application of one or four instruments has not significant impact on the above results. The author also tried to augment equation 3 with “fear” index VIX to control for negative sentiment. VIX variable is not significant and it does not change the value and significance of the rest of the coefficients.

## 4 Discussion and Conclusion

This paper studies the impact of excessive social activity on bond-like stocks. The author presents the theory that bond-like stocks can be under the influence of occasional excessive activity in the society. The rationale lies in the availability heuristic (Tversky & Kaheman, 1973) with the result in the herding behavior as an overreaction due to sentiment information.

The author as well proposes the appropriate empirical methods to estimate the excessive levels of sentiment. The paper provides conclusion about the negative impact of excessive social activity. Specifically, social activity which overcomes the threshold is connected with negative sentiment with negative reflection in stock returns. These results are in line with theories and empirical results describing the negative events as more influential Kahneman & Tversky, (1979), De Bondt & Thaler, (1985), Tetlock (2007), Tetlock, Saar-Tsechansky & Mackassy (2008), Da, Engelberg & Gao (2011).

The paper is unique by employment of unique data tracked by social media Facebook. The author proposes the social media as an appropriate source of sentiment in the society. This paper accomplished the analysis on a sample of 21 companies. The further research should be extended to bigger sample. An extension of the model with control variables like media coverage, which should capture events like new product release can be beneficial as well. Further research with social media data should be accompanied with proper evaluation of data and the distinction between positive and negative reactions of society to news. This can reveal new features in the analysis of sentiment and stock returns.

## References

ARELLANO, M. and BOVER, O., 1995: Another look at the instrumental variables estimation of error-components models. Journal of Econometrics 68 (1). 29-51.

BARBERIS, N., SHLEIFER, A. and VISHNY, R., 1998: A model of investor sentiment. Journal of Financial Economics, 49: 307-343.

BAKER, M. and WURGLER, J., 2007: Investor Sentiment in the Stock Market. National Bureau of Economic Research. Working Paper 13189: 1-38.

BICK, A., 2010: Threshold effects of inflation on economic growth in developing countries. Economic letters 108 (2): 126-129.

BOLLEN, J., MAO, H. and ZENG, X.-J., 2011: Twitter mood predicts the stock market. Journal of Computational Science, 2 (1): 1-8.

BROWN, S. and WARNER, J., 1985: Using daily stock returns: the case of event studies. Journal of Financial Economics 14. 3-31.

CHOI, H. and VARIAN, H., 2009: Predicting the Present with Google Trends, Tech. rep., Google, 2009.

CANER, M., and HANSEN, E. B., 2004: Instrumental variable estimation of a threshold model. Econometric Theory. 20: 813-843.

DA, Z., ENGELBERG, J. and GAO, P., 2011: In search of attention. Journal of Finance, 66, 1461–1499.

DA, Z., ENGELBERG, J. and GAO, P., 2015: The Sum of All Fears Investor Sentiment and Asset Prices. The Review of Financial Studies, 28 (1), 1–32.

DE BONDT, W. F. M. and THALER, R. H., 1985: Does the Stock Market Overreact? Journal of Finance 40(3), 557-581.

DELONG, J., B., SHLEIFER, A., SUMMERS, L., H. and WALDMANN, R., J., 1990: Noise Trader Risk in Financial Markets. Journal of Political Economy, 98: 703-738.

EDMANS, A., GARCÍA, D., and NORLI. Ø., 2007: Sports Sentiment and Stock Returns. Journal of Finance, 62: 1967-1998.

FAMA, E. F., 1965: The Behavior of Stock-Market Prices. The Journal of Business 38 (1): 34-105.

FRIEDMAN, M., 1953: The case for flexible exchange rates. In: Essays in Positive Economics. University of Chicago Press, Chicago.

HANSEN, E. B., 1999: Threshold effect in non-dynamic panels: Estimation, testing, and inference. Journal of Econometrics. Vol. 93 (1999): 345-368.

KAHNEMAN, D. and TVERSKY, A., 1979: Prospect theory: An Analysis of decisions under risk. Econometrica, 47. 313-327.

KAPLANSKY, G. and LEVY, H., 2010a: Sentiment and stock prices: The case of aviation disasters. Journal of Financial Economics. 95(2): 174-201.

KAPLANSKY, G. and LEVY, H., 2010b: Exploitable Predictable Irrationality: The FIFA World Cup Effect on he U.S. Stock Market. Journal of Financial and Quantitative Analysis. 45(2): 535-553.

KAMSTRA, M. J., KRAMER, L. A., and LEVI, M. D., 2003: Winter blues: a SAD stock market cycle. American Economic review 93, 324-343.

KAPOUNEK S. and LACINA L., 2011: Inflation perceptions and anticipations in the old Eurozone members states. Prague economic papers. 2 (2011): 120-139.

KING G, B., 1966: Market and industry factors in stock price behavior. Journal of business, University of Chicago Press. January 1966. No. 39.

KREMER, S., BICK, A., and NAUTZ, D., 2012: Inflation and growth: New evidence from a dynamic threshold analysis. Empirical Economics (2013) 44, 861-878.

LEROY, S. and PORTER, R., 1981: The present-value relation: tests based on implied variance bounds. Econometrica, 49, 97−113.

LIVINGSTONE, M., 1977: Industry Movements of Common Stocks. Journal of Finance, 32 (3), 861–874.

NOFSINGER, J., 2005: Social Mood and Financial Economics. Journal of Behavioral Finance, 6 (3), 144–160

RENN, O., 1990: Risk perception and risk management: A review. Risk Abstracts, 7 (1), 1– 9.

ROLL, R., 1992: Industrial Structure and the Comparative Behavior of International Stock Market Indices. The Journal of Finance. 47(1): 3-41.

SCHWARTZ, H., 1998: Rationality Gone Awry? Decision Making Inconsistent with Economic and Financial Theory. Westport, CT: Greenwood Publishing Group, Inc.

SHILLER, J. R., 1981: Do stock prices move too much to be justified by subsequent changes in dividends? American Economic Review, 71, 421−436.

SHILLER, J. R., 2003: From Efficient Markets Theory to Behavioral Finance. Journal of Economic Perspectives. 17(1): 83-104.

SHLEIFER, A. and VISHNY, W, R., 1997: The limits of arbitrage. The Journal of Finance. 52(1): 35-55.

SPRENGER, T., O. and WELPE, I., M., 2011: News or noise? The Stock Market Reaction to Different Types of Company-Specific News Events. SSRN Working Paper, January 4, 2011.

TETLOCK, P. C., 2007: Giving content to investor sentiment: The role of media in the stock market. Journal of Finance, 62(3) 1139-1168.

TETLOCK, P. C., SAAR-TSECHANSKY, M. and MACSKASSY, S., 2008: More than words: Quantifiying language to measure firms' fundamentals. Journal of Finance, 63: 1437-1467.

TVERSKY, A. and KAHNEMAN, D., 1973: Availability: A heuristic for judging frequency and probability. Cognitive Psychology. 5(2): 207-232.