European Journal of Business and Management
ISSN 2222-1905 (Paper) ISSN 2222-2839 (Online) DOI: 10.7176/EJBM
Vol.11, No.24, 2019

Research on Management Over-optimistic Measurement：Based on Text Sentiment Analysis

Guo-tong Wu
Faculty of Management and Economics, Dalian University of Technology
NO.2 Linggong Road, Ganjingzi District, Dalian 116024, China

Abstract
Over-optimism managers have an important impact on corporate financing decisions.Based on the text sentiment mining technology, this paper calculates the optimism of management by constructing the emotional lexicon and using the frequency of optimistic and pessimistic vocabulary in the PFP language statistical annual report. In this paper, by constructing a theoretical model reflecting the degree of optimism of management, the residual of the model is used as the proxy variable of management over-optimism, and the rationality judgment is carried out, which provides a good research tool and method for the subsequent investigation.
Keywords: Over-optimistic ; Text Sentiment Analysis; Theoretical Model
DOI: 10.7176/EJBM/11-24-14
Publication date: August 31st 2019

1. Introduction

Traditional economic theory is based on the manager's rational assumptions and believes that managers are rational in the decision-making process. However, there are many "visions" in the real economic life that cannot be explained by traditional economic theories. More and more scholars began to examine management's behavioral decisions from an irrational perspective, such as overconfidence, risk appetite, cognitive bias, optimism, conservatism, etc. These bold attempts have effectively promoted the development of academe.Over-optimism is an important form of management's irrational characteristics, and it is universally present in real economic life. Over-optimism comes from psychology, which refers to overestimating the possibility of a better future prospect and underestimating the possibility of a poor future. Over-optimism of the management mainly means that the management will think optimistically of the future business development prospect of the enterprise, overestimate the possibility of obtaining a good development prospect of the enterprise, and underestimate the possibility of problems existing in the business development of the enterprise. Over-optimism as a management personality trait will be reflected in business decision-making. Over-optimism managers will have a more optimistic view of the company's future prospects, which will also have an important impact on corporate financing decisions.

There has been a lot of literature around the management's over-optimistic attempts, mainly using management shareholding changes (Malmendier and Tate, 2005), business climate index (Oliver, 2005), earnings forecast bias (Ben-David et al., 2006).) and the questionnaire survey (Scheier and Carver, 1985; Ben-David et al., 2006) to measure the over-optimism of management, these methods can not more comprehensively and reasonably measure the degree of management over-optimism. Based on the text sentiment mining technology, this paper calculates the optimism of management by constructing the emotional lexicon and using the frequency of optimistic and pessimistic vocabulary in the PFP language statistical annual report. In this paper, by constructing a theoretical model reflecting the degree of optimism of management, the residual of the model is used as the proxy variable of management over-optimism, and the rationality judgment is carried out, which provides a good research tool and method for the subsequent investigation.

2. Review of existing measurement methods

2.1 Management shareholding changes

Malmendier and Tate (2005) believe that the increase in management shareholdings represents management is more optimistic, the more management increases the company's stocks, the greater the degree of over-optimism. However, most of the management shareholdings have a limited sale period. The management shareholding does not indicate management's optimistic and pessimistic expectations for the company's future performance.

2.2 Business climate index

Oliver (2005) think that the higher the business climate index, the higher the management's optimism based on the survey data research method. The business climate index reflects the overall judgment of the entrepreneurs on the macroeconomic environment and can indicate the optimism of entrepreneurs on the future economic development trend. There are two problems in the business climate index: First, the disclosure of the business climate index is cyclical. The business climate index is quarterly, which makes data users unable to reasonably obtain the annual business climate index. Second, the business climate index reflects the overall mood of

entrepreneurs. The business climate index cannot reflect the attitude of individual enterprises to the economic trend, and it cannot reflect the optimism of the management of different enterprises in different industries.

2.3 Earning forecast deviation

Ben-David et al. (2006) believe that the actual profit level of the company is lower than the predicted profit level, which indicates that the management of the company is more optimistic. The greater the actual profit level is higher than the predicted profit level, the higher the management's over-optimism. However, the factors affecting the deviation of corporate earnings forecasts in real economic life are complex, including not only the optimism of management, but also the management ability, economic environment and other factors.

2.4 Earning forecast deviationQuestionnaire survey

Scheier and Carver (1985), Ben-David et al. (2006) obtained information on management optimism through questionnaires, and on this basis, judged the degree of optimism of management. Questionnaires can effectively obtain first-hand data and have been widely used in the field of behavioral science research. The limitation of this kind of questionnaire is that it cannot meet the large sample size of empirical research needs.

3. Application of text sentiment analysis in corporate finance

3.1 Management tone content

A large number of literature studies have shown that corporate text information and its management tone contain valuable incremental information and the market will respond to it. Brochet et al. (2012) based on teleconference research found that the more complex the management language, the smaller the market share price and transaction volume reacted to it. Davis et al. (2012) based on a net positive tone study of performance news releases, found that the more positive the management tone, the higher the cumulative abnormal return for the three days. Price et al. (2012) based on the quarterly conference call study, management tone is an important predictor of the company's abnormal returns and trading volume. Sandulescu (2015) found that the more positive the tone of the MD&A, the more followers buy the stock. Jegadeesh and Wu (2013) found that the more positive vocal about the disclosure of risk information, the weaker the IPO bargaining. Feldman et al. (2010) found that the more positive the management tone, the higher the company's contemporaneous and future earnings. Li (2012) found that management tone was positively correlated with the company's future earnings and liquidity. Mayew et al. (2015) pointed out that the more positive the tone of MD&A text information, the lower the probability of ruin in the future.

3.2 Management tone control

Davis and Tran (2012) found that when companies have poor performance or a negative surplus, management has a strong incentive to manipulate the tone in an attempt to blind investors. Larcker and Zakolyukina (2012) found that CEOs of financial fraud companies use more positive language in their quarterly earnings conference calls. Huang et al. (2014) pointed out that the abnormal positive tone in the performance press conference text is often accompanied by the company's negative future earnings and cash flow. The management will misunderstand investors' understanding of the company's fundamentals through strategic tone management.

3.3 Analyst report

Demers and Vega (2011), Samuel et al (2013), and Huang and Mamo (2014) all agree that management's net positive tone and language certainty are significantly positively correlated with analysts' forecast changes.

4. Management over-optimistic measurement research

4.1 Collecting text data

This article takes China's A-share listed companies from 2007 to 2017 as the research object, excluding ST companies, data missing companies, financial industry companies and financial data abnormal companies, and finally obtained 22186 annual observation samples. Since the current mainstream databases do not provide data service content related to “management discussion and analysis”, this leads to a complicated data collection process.

The specific text data collection process is as follows: First, collect financial reports. This article uses the web crawler program to automatically obtain the annual financial report of the PDF version of the listed company; The second step is to convert the PDF file into a TXT file based on the Linux system platform using Xpdf software; The third step is to select the "management discussion and analysis" content.

4.2 Text preprocessing

The word segmentation mainly refers to subdividing the text into multiple independent terms according to the agreed rules. The development of English participles is relatively mature, which is closely related to the sentence structure and vocabulary composition of English. Compared with English word segmentation, the development

of Chinese word segmentation is relatively slow. This is mainly because the combination of different Chinese characters can express different meanings. This paper uses the Chinese word segmentation technology provided by LTP to carry out scientific segmentation of the collected text data, which lays a solid foundation for the subsequent text sentiment analysis.

4.3 Building an emotional lexicon

The most important part of text sentiment analysis is to construct emotional lexicon. The science of emotional lexicon design directly affects the measure of excessive optimism of management. Although there are already many sentiment lexicons available for text sentiment analysis, the dictionary professions of these "general fields" are not strong, which makes them not suitable for research in the fields of corporate finance and accounting. Loughran and Mcdonald (2011) developed the Financial Sentiment Dictionary (FSD) from the financial texts using the terms in the Harvard Dictionary (4th Edition). The sentiment classification standard has been widely used in corporate finance and accounting (Garcia, 2013; Hoberg and Phillips, 2016; Buehlmaier and Whited, 2018; Bushee et al., 2018). From the results of the study, the standard is more scientific and reasonable. Therefore, this paper refers to the word list provided by Loughran and McDonald (2011), and combines the Chinese word habits and the annual report information context to construct the "Sentiment Dictionary of Chinese Corporate Financial Reports (SDCCFR)". The "Chinese Enterprise Financial Annual Report Emotional Thesaurus" constructed in this paper contains 3992 optimistic emotional terms and 4059 pessimistic emotional terms.

4.4 Text sentiment analysis

This paper mainly uses the Bag of Words Model to analyze text emotions. The main basis is as follows: First, as Tetlock et al. (2008) put it, the bag of words model fully conforms to the four elements of text sentiment analysis. After reviewing the literature, this paper finds that most of the research in the fields of finance and accounting adopts the bag of words model (Loughran and Mcdonald, 2011; Schmidt, 2015; Bodnaruk et al., 2015; You et al., 2017; Chemmanur and Moreira, 2019). This can also verify the scientificality of the bag of words model in this paper. Secondly, this paper mainly analyzes the emotional attributes of “management discussion and analysis”, while the text content of “management discussion and analysis” is relatively small, does not require large sample analysis. Studies have confirmed that in the small sample text sentiment analysis, it is more reasonable to use the bag of words model to judge the emotional characteristics of the text (O’Connor et al., 2010).

In this paper, the text sentiment analysis program is written in PHP language. The TXT document of the “management discussion and analysis” processed in the previous article is traversed twice, and the emotional word frequency is captured and counted, and the result is stored in the MySQL database. Finally, the text emotional outcome sent from the MySQL database is exported to Excel.

4.5 Management over-optimistic measurement

The most important part of text sentiment analysis is to construct emotional lexicon. The science of emotional lexicon design directly affects the measure of excessive optimism of management. Although there are already many sentiment lexicons available for text sentiment analysis, the dictionary professions of these "general fields" are not strong, which makes them not suitable for research in the fields of corporate finance and accounting. Loughran and Mcdonald (2011) developed the Financial Sentiment Dictionary (FSD) from the financial texts using the terms in the Harvard Dictionary (4th Edition). The sentiment classification standard has been widely used in corporate finance and accounting (Garcia, 2013; Hoberg and Phillips, 2016; Buehlmaier and Whited, 2018; Bushee et al., 2018). From the results of the study, the standard is more scientific and reasonable. Therefore, this paper refers to the word list provided by Loughran and McDonald (2011), and combines the Chinese word habits and the annual report information context to construct the "Sentiment Dictionary of Chinese Corporate Financial Reports (SDCCFR)". The "Chinese Enterprise Financial Annual Report Emotional Thesaurus" constructed in this paper contains 3992 optimistic emotional terms and 4059 pessimistic emotional terms.

Drawing on the approach of Henry(2006), Price et al.(2012), Henry and Leone(2015), Lin Le and Xie Deren(2017), this article uses the following formula to measure the degree of management optimism.

$$Optimistic_{it} = \frac{POS_{it} - NEG_{it}}{POS_{it} + NEG_{it}} \quad (1)$$

Among them, Optimistic is the degree of management optimism. The larger the value, the higher the optimism of the management of the company. POS is the number of times optimistic vocabulary appears in MD&A, and NEG is the number of pessimistic vocabulary occurrences in MD&A.

Huang et al. (2013) pointed out that management's tone in the annual report is a reflection of the actual

operation and future prospects of the company. Management optimism should have the following linear relationship with the actual business operations and future prospects:

$$Optimistic_{it} = \alpha_0 + \alpha_1 EPS_{it} + \alpha_2 \Delta EPS_{it} + \alpha_3 Size_{it} + \alpha_4 Lev_{it} + \alpha_5 Growth_{it} + \alpha_6 ROE_{it} + \alpha_7 Age_{it} + \alpha_8 Loss_{it} + \varepsilon \quad (2)$$

Among them, EPS is the earnings per share; △EPS is the variable value of earnings per share; Size is the size of the enterprise, specifically the natural logarithm of the total assets; Lev is the solvency, specifically the asset-liability ratio; Growth is the growth capacity, specifically the growth rate of operating income; ROE is profitability, specifically the return on net assets, Age is the period of listing, Loss is the loss of dummy variable. The residual of the model (2) regression indicates the degree to which the management optimistic tone does not match the actual business operation and future prospects can be used as a proxy variable for the management's over-optimistic Over-optimistic. The greater the residual value, the higher the degree of over-optimism of the management. In this paper, when the management is overly optimistic, it uses the industry-by-year regression to obtain the residual of the regression model. This article is only based on the regression results of the manufacturing listed companies in 2017. It can be seen from Table 1 that the overall fit of the model is good, and the adjusted R2 is 0.112.

Table 1. Regression Results of Managerial Over-optimistic Measures

Variables	*Optimistic*
EPS_{it}	0.006***
	(5.055)
ΔEPS_{it}	0.005***
	(3.488)
$Size_{it}$	0.002***
	(6.014)
Lev_{it}	-0.070***
	(-29.937)
$Growth_{it}$	0.006***
	(7.657)
ROE_{it}	0.018***
	(4.293)
Age_{it}	-0.027***
	(-42.334)
$Loss_{it}$	-0.019***
	(-10.404)
Constant	-0.104***
	(-12.883)
Year/Ind	*Control*
Adj-R^2	0.112
N	1914

Notes: *, **,and *** denote significance at the 10%, 5%, and 1% levels, respectively, using a two-tailed t-test. The t-statistics are reported in parentheses.

5. Rationality judgment

In order to ensure the reliability of the management's over-optimistic measurement results, from the perspective of corporate investment, this paper mainly uses the following two methods to prove:

Method one, the actual investment of enterprises and the development trend of the real economy. The actual business activities of enterprises should conform to the development of macroeconomic trends, and rational management should follow the pace of macroeconomic development to make investment decisions. When the actual investment of the enterprise is far greater than the optimal investment behavior that meets the macroeconomic cycle and can maximize the return, we can assume that the management of the company is overly optimistic. The macro economy can specifically include the real economy and the capital market economy. Considering that the real economy is more closely related to corporate investment, this article mainly takes the real economy as an example. Investment behavior in the real economy mainly refers to investment in operating assets, such as new factory buildings, equipment purchase, and establishment of subsidiaries. The development of the real economy is often accompanied by cyclical fluctuations, and the period of decline and rise of the real economy alternates. Since overly optimistic management tends to overestimate its own capabilities and future development prospects, when the real economy declines, excessively optimistic management's corporate investment should continue to rise. The relationship between the specific enterprise and the real economy cycle fluctuation curve is shown in Figure 1. The points A to C in the figure indicate that the

 DOI: 10.7176/EJBM

overly optimistic management will increase the investment level of the enterprise during the period of the decline of the enterprise real economy.

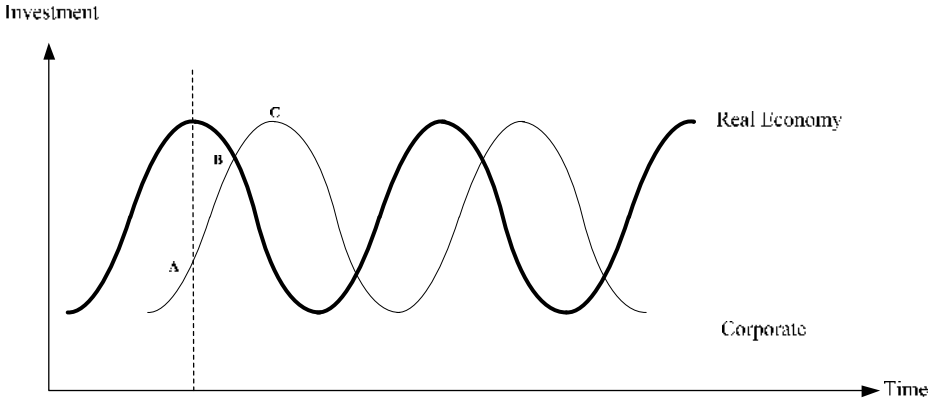

Fig1. Periodic Fluctuation Curve of Corporate and Real Economy

Method 2, the company over-investment. Overly optimistic management is more optimistic about the company's future development prospects, which makes the company's investment level tends to be higher. The investment level of the company in which the over-optimistic management is located is often higher than that of other companies in the same industry in the same year. More performance is an over-investment. Conversely, the performances of companies where excessively pessimistic management is are often under-invested. Based on this, this paper verifies the reliability of the previous metric model by examining the relationship between over-investment and management over-optimism from the perspective of over-investment in enterprises.

This paper adopts China's fixed asset investment growth rate (GFA) and GDP growth rate (GGDP) as the proxy variables of the real economy development trend. The difference between the growth rate of enterprise investment and the growth rate of fixed asset investment and GDP growth rate indicates the degree to which the enterprise investment level deviates from the optimal investment level, and is recorded as GFA and GGDP respectively. At the same time, this paper draws on Richardson's (2006) idea to measure the level of over-investment of enterprises （Over-Investment） by means of industry-by-year. Based on this, the paper uses the Pearson correlation test and the Spearman correlation test to conduct consistency test. Table 2 shows the results of the consistency test. From the table, there is a significant positive correlation between the over-optimistic management （Over-optimistic） and the other three variables, and the correlation coefficient is greater than 0.7. It can be seen that the excessive optimism of the management measured in this paper is feasible and reasonable.

Table. 2 Results of Conformity Test

Variables	*Over-Optimistic*	*GFA*	*GGDP*	*Over-Investment*
Over-optimistic	1.000	0.785***	0.791***	0.760***
GFA	0.753***	1.000	0.982***	0.774***
GGDP	0.767***	0.971***	1.000	0.784***
Over-Investment	0.709***	0.644***	0.664***	1.000

Notes: ①***、**and* denote significance at the 1%, 5%, and 10% levels; ②below the diagonal is the Pearson correlation test result, and above the diagonal is the Spearman correlation test result.

References

Ben-David I, Graham J R, Harvey C R. Managerial overconfidence and corporate policies[R]. Working Paper, 2007.

Bodnaruk A, Loughran T, McDonald B. Using 10-K text to gauge financial constraints[J]. Journal of Financial and Quantitative Analysis, 2015, 50(4): 623-646.

Bonsall S B, Bozanic Z, Fischer P E. The informativeness of disclosure tone[R]. Working Paper, 2013.

Brochet F, Naranjo P, Yu G. Causes and consequences of linguistic complexity in non-US firm conference calls[R]. Working Paper, 2012.

Buehlmaier M M M, Whited T M. Are financial constraints priced? Evidence from textual analysis[J]. The Review of Financial Studies, 2018, 31(7): 2693-2728.

Bushee B J, Gow I D, Taylor D J. Linguistic complexity in firm disclosures: obfuscation or information?[J]. Journal of Accounting Research, 2018, 56(1): 85-121.

Chemmanur T J, Yan A. Advertising, attention, and stock returns[J]. Quarterly Journal of Finance, 2019: 190-195.

Davis A K, Piger J M, Sedor L M. Beyond the numbers: Measuring the information content of earnings press

release language[J]. Contemporary Accounting Research, 2012, 29(3): 845-868.
Davis A K, Tran N. Earnings quality, proprietary disclosure costs and managers' use of disclosure tone to signal future performance[R]. Working Paper, 2012.
Demers E, Vega C. Linguistic tone in earnings announcements: News or noise[R]. Working Paper, 2011.
Feldman R, Govindaraj S, Livnat J, et al. Management's tone change, post earnings announcement drift and accruals[J]. Review of Accounting Studies, 2010, 15(4): 915-953.
Garcia D. Sentiment during recessions[J]. The Journal of Finance, 2013, 68(3): 1267-1300.
Henry E. Market reaction to verbal components of earnings press releases: Event study using a predictive algorithm[J]. Journal of Emerging Technologies in Accounting, 2006, 3(1): 1-19.
Hoberg G, Phillips G. Text-based network industries and endogenous product differentiation[J]. Journal of Political Economy, 2016, 124(5): 1423-1465.
Huang X, Teoh S H, Zhang Y. Tone management[J]. The Accounting Review, 2013, 89(3): 1083-1113.
Huang A G, Mamo K..Do Analysts read the news?[R]. Working Paper, 2014.
Jegadeesh N, Wu D. Word power: A new approach for content analysis[J]. Journal of Financial Economics, 2013, 110(3): 712-729.
Larcker D F, Zakolyukina A A. Detecting deceptive discussions in conference calls[J]. Journal of Accounting Research, 2012, 50(2): 495-540.
Li F. Annual report readability, current earnings, and earnings persistence[J]. Journal of Accounting and economics, 2008, 45(2-3): 221-247.
Malmendier U , Tate G. CEO Overconfidence and Corporate Investment[J]. Journal of Finance, 2005, 60.
Mayew W J, Sethuraman M,Venkatachalam M. MD&A disclosure and the firm's ability to contin- ue as a going concern［J］. The Accounting Review, 2015,90(4) : 22－36
O'Connor B, Balasubramanyan R, Routledge B R, et al. From tweets to polls: Linking text sentiment to public opinion time series. 2010[C]Proceedings of 4th International AAAI Conference on Weblogs and Social Media (ICWSM-2010). 2018.
Oliver B R. The impact of management confidence on capital structure[R]. Working Paper, 2005.
Price M K , Doran J S , Peterson D R , et al. Earnings conference calls and stock returns: The incremental informativeness of textual tone[J]. Journal of Banking & Finance, 2012, 36(4):992-1011.
Richardson S. Over-investment of free cash flow[J]. Review of accounting studies, 2006, 11(2-3): 159-189.
Sandulescu，P. M. Insiders ' incentives of using a specific disclosure tone when trading[J]. Studies in Communication Sciences,2015,15(1) : 12－36.
Scheier M F, Carver C S. Optimism, coping, and health: assessment and implications of generalized outcome expectancies[J]. Health psychology, 1985, 4(3): 219-247.
Schmidt B. Costs and benefits of friendly boards during mergers and acquisitions[J]. Journal of Financial Economics, 2015, 117(2): 424-447.
Tetlock P C, Saar-Tsechansky M, Macskassy S. More than words: Quantifying language to measure firms' fundamentals[J]. The Journal of Finance, 2008, 63(3): 1437-1467.
You J, Zhang B, Zhang L. Who captures the power of the pen?[J]. The Review of Financial Studies, 2017, 31(1): 43-96.
Loughran T, McDonald B. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks[J]. The Journal of Finance, 2011, 66(1): 35-65.