



Research on Insider Trading Based on Sentiment Analysis of Annual Reports

Yu Qi^{1,*}

¹ Business master of commerce , University of Sydney, Camperdown/Darlington NSW 2006, 1466, Australia

*Corresponding author. Email: yuqi8843@uni.sydney.edu.au

ABSTRACT

Along with the split share structure reform, the phenomenon of insider trading was born. In addition, with the continuous improvement of our country's laws, insider trading has increased year by year, and its impact on the capital market cannot be ignored. This paper takes the number of insider transactions before the annual report of the year as the research object, studies the impact of annual report sentiment on insider transactions and analyses whether insiders have information advantages. The results of the study show that when listed companies release positive news, insider trading increases significantly, while for bad news, insider trading is less. This is since insiders in my country follow the "timed trading theory" before the earnings announcement. The success of the above research can not only make up for the vacancy of domestic A-share insider trading, but also improve the current research direction of insider trading information advantages and lay the foundation for further research on insider trading in the future.

Keywords: *Sentiment Analysis, Insider trading, Information advantage.*

1. INTRODUCTION

The activity of insider trading improves the effectiveness of securities market information, and at the same time, it also has a certain degree of restraint on insider trading. While allowing insider trading, the China Securities Regulatory Commission continues to gradually improve insider trading behavior through various laws and regulations. Especially in 2015, the domestic securities market was turbulent. Since the insider trading behavior has an important reference role for market participants, The China Securities Regulatory Commission has repeatedly issued policies to restrict insiders from selling, and actively encourage insiders to buy stocks.

On the one hand, the activity of insider trading enhances the effectiveness of the securities market and imposes constraints on insider trading. The China Securities Regulatory Commission requires listed companies to disclose major events in a timely manner to prevent insiders from trading with insider information and to reduce information asymmetry between insiders and general investors. At the same time, the disclosure of information such as the number, time, and price of insider transactions on the Shanghai Stock Exchange and Shenzhen Stock Exchange is conducive to ordinary investment and supervision of the legality and

compliance of insider trading behavior and restricts insiders from using listed companies' major non-public information. information to trade.

With the rapid development of artificial intelligence and machine learning technology and the improvement of big data analysis capabilities, all kinds of financial data, such as annual reports, quarterly earnings announcements, prospectuses, and performance press conferences, have been widely used in many fields. In finance, textual data can be used to analyze issues such as annual report readability, investor attention, media sentiment, management tone, news implied volatility, and investor disagreement. For structured and numerical financial statements, unstructured texts occupy a larger space in external release channels and contain richer information connotations. Therefore, financial text analysis has a great impact on the improvement of capital market information quality and supervision efficiency. important value.

2. LITERATURE REVIEW

As an important source of financial information disclosure, the annual report is a bridge between listed companies and investors, and it is also the first-hand information for investors to grasp the latest developments of the company, reducing insider trading, and

maintaining market order are of great significance (Yu Wei, 2021). From the perspective of transparency, companies or corporate executives tend to conduct strategic information manipulation, such as hiding negative information, exaggerating positive information (Biggerstaff et al., 2020) to seek personal benefits such as position compensation, political promotion, and option exercise. Relevant research results show that the executives of companies with low transparency can obtain significant abnormal reports through insider transactions, while executives of companies with high information transparency have no significant reports (Dong Jichang et al., 2021). From a normative perspective, the degree of standardization of information disclosure will affect the operational risk of financial companies and platforms. The higher the degree of standardization of information disclosure, the lower the operational risk. However, at present, there are no unified requirements for the standardization of financial information disclosure, and the problem of information lack is more serious. Compared with transparency and standardization, readability is the most basic requirement of the annual report text, which determines whether investors can accurately and quickly understand the company's true financial status and operating results. The "Securities Law of the People's Republic of China" puts forward "concise, clear and easy-to-understand" requirements for the information disclosed by listed companies. Revised in 2017) clearly pointed out that it is necessary to strengthen the information disclosure of the annual report text, to facilitate investors to understand, see and see through the listed company. To improve the decision-making quality of information users, annual reports of listed companies need to convey the content of accounting information accurately and clearly to information users (Yang Dan et al., 2018). At present, scholars at home and abroad have carried out a series of studies on the readability of annual reports. Miller (2010) believes that easy-to-read annual reports can effectively reduce the information processing cost of investors and help them to grasp the information related to the company's value more quickly. Yang 's (2018) study found that investors are more willing to trade stocks with highly readable annual reports, because simple and clear annual reports will make investors more trustworthy. The higher the readability of the annual report of a listed company, the higher the earnings response coefficient, that is, the readability of the annual report is conducive to the better reflection of the company's earnings information in the company's stock price. There is a strong positive correlation between the readability of annual reports and the quality of corporate information disclosure, that is, the higher the readability you guarantee, the higher the quality of corporate information disclosure, and the more able to attract investors' attention and analysts to follow up. Enter.

Managing annual report tone is another way company

insiders manage textual information. In recent years, more and more studies have begun to pay attention to the role of emotional information in texts in information transmission. Companies can use a more positive annual report tone to create an atmosphere of company development and prosperity. The operation is less difficult, and the cost is low, which will make investors feel the company's operating performance is more confident, and it is easier to demonstrate negative news. Just like the readability of the management annual report, the behavior of manipulating the tone of the annual report will not only seriously mislead the decision-making judgment of information users, but also seriously affect the information asymmetry in the furniture capital market. Manipulation of the tone of annual reports also has a range of economic consequences. The more the tone was manipulated, the persistence of the company's earnings at the end of the month. found that the more positive the tone of the annual report, the higher the risk of stock price crash. The tone of the annual report can reduce the synchronization of stock prices, and this effect is more pronounced the more analysts track it (Liang ,2021).

It can be seen from the above research that the text of the annual report is easy to modify and difficult to supervise, so that the readability of the annual report may become a more concealed information disclosure method for the company. The readability of annual reports is likely to become a more covert information disclosure manipulation method for companies. The readability and tone of the annual report, as two relatively hidden features of the annual report text, are more difficult to be detected by regulators and investors. Therefore, the in-depth study of the two is of great significance and value.

3. APPLICATION OF MODELS

Multiple regression analysis is a statistical technique that analyzes the relationship between two or more variables and uses the information to estimate the value of the dependent variables. In multiple regression, the objective is to develop a model that describes a dependent variable y to more than one independent variable.

In linear regression, there is only one independent and dependent variable involved. But, in the case of multiple regression, there will be a set of independent variables that helps us to explain better or predict the dependent variable y .

The multiple regression equation is given by

$$y = a + b_1 x_1 + b_2 x_2 + \dots + b_k x_k$$

where $x_1, x_2 \dots x_k$ are the k independent variables and y is the dependent variable.

4. EXPERIMENT

4.1. Feature Extraction

This paper takes the annual reports of listed companies in Shanghai and Shenzhen A-shares from 2010 to 2019 as a research sample, and studies the relationship between the readability of annual reports, the tone of annual reports and the behavior of company insiders buying and selling the company's stocks in the secondary stock market within a period after the annual report is disclosed. It is proposed that financial listed companies and relevant indicators are indeed companies. Considering that newly listed companies have a certain lock-up period for the stocks held by their executives, there is no insider trading during this period. Therefore, it is proposed that the listing date should be the date of publication of the annual report. Samples with intervals less than 1095 (3 years). The annual reports of listed companies and company transaction data and financial data used are all from the CSMAR database. This paper uses the text analysis method to extract, process and analyze the text information of the annual report, to obtain the measure of the readability of the annual report.

4.2 Construction of annual report readability indicators

The complexity of the annual report text is reflected in multiple levels. This paper draws on the research of Dong et al. (2020) and selects the following indicators to measure the readability of the annual report: Average sentence length, which is divided by dots After the sentence, the number of tie characters in the annual report sentence. After research, the optimal length of a sentence is 7-12 sub-lengths. If the length exceeds this length, it will be more difficult for readers to understand . Therefore, this paper uses the length of the tie sentence as an indicator to measure the readability of the annual report.

4.3 The Construction of Tone Index of Annual Report

Drawing on the research of Xie Deren and Lin Le (2015), the model (1) is used to calculate the net tone of the annual report.

$$Tone = \frac{Positive - Negative}{Positive + Negative} \tag{1}$$

Among them, Positive is the number of positive intonation words, and Negative is the number of negative intonation words. Tone is the net tone of the management. The larger the value, the more positive the management's words.

4.4 Display of running results

The results of using the multiple regression model are shown in Table 1. The implicit information of the annual report such as Tone and sentences number has a certain influence on the number of insider transactions, while the public information such as changes in stock transactions and average transaction price has no significant impact on the number of insider transactions.

Table 1 Regression Results

	coef	std err	t	P> t	[0.025	0.975]
Intercept	576.8146	1.145	503.980	0.000	574.571	579.058
asi	-0.0029	0.010	-0.285	0.775	-0.023	0.017
Tone	-8.6088	3.631	-2.371	0.018	-15.726	-1.492
TotalWordsNum	0.0472	0.001	94.371	0.000	0.046	0.048
SentencesNum	-1.2518	0.020	-63.691	0.000	-1.290	-1.213
ChangingShares	-1.429e-06	1.12e-07	-12.702	0.000	-1.65e-06	-1.21e-06
AVGTradingPrice	6.766e-05	7.6e-06	8.902	0.000	5.28e-05	8.26e-05

From Figure 1, when the positive words in the annual report are greater than the negative words, that is, tone>0 has more insider transactions. For those listed companies that are undervalued in the market, insiders will pay more attention to the informational reason of stock returns after the company's earnings announcement when trading before the announcement, and this phenomenon is more obvious when the listed company releases good news, but for negative news. No obvious effect.

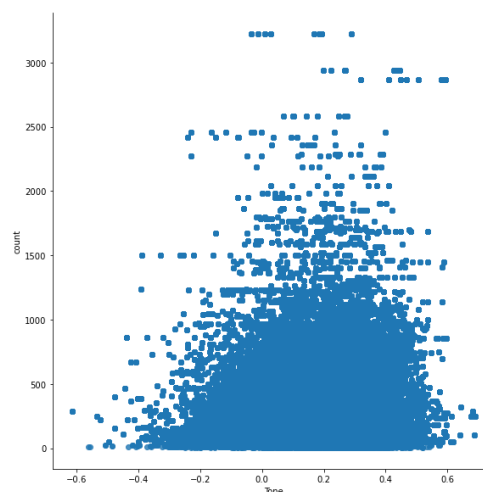


Figure 1 The relationship between Tone and count

4.5 The advantages and disadvantages of the model

According to the experimental results, the model can judge that the implicit information in the annual report does have a certain impact on the number of insider transactions, and the experiment has reference value. There is indeed information hiding in information. Although there is a certain deviation in the results, in actual stock trading, the sentiment analysis of the annual

report can indeed be used to infer the business status of the company, thereby improving the actual income of investors. The regulator can also use tone data to analyze the company's annual report. accuracy.

However, the model still has imperfections. First, for better interpretation, the model discards part of the experimental accuracy. Secondly, it does not consider the impact of the liquidity on the number of insider transactions, so it is necessary to explore the calculation method of the relevant insider transaction willingness.

5. Conclusion

This paper takes the number of insider transactions as the research object, and studies the emotional vocabulary of annual reports. Insiders use the information advantage to trade. The information advantage is mainly the insider's perception of the future performance of the company's stock and the insider's perception of the valuation of listed companies. Through research, when the annual report of listed companies is in a positive tone, insider transactions are more frequent, and when it is in a negative tone, the transactions are relatively scattered. The above conclusions can not only show that the domestic A-share insider trading will trade according to the information, but also can further distinguish the insider trading behavior by defining the two kinds of information.

REFERENCES

- [1] Ali, F., Wang, M., Jebran, K., & Ali, S. T. (2021). Board diversity and firm efficiency: evidence from China. *Corporate Governance: The International Journal of Business in Society*, 21(4), 587–607. <https://doi.org/10.1108/cg-10-2019-0312>
- [2] Dong Jichang, LEI Yanxi, LI Xiuting, DONG Zhi. Research on the Characteristics and Influencing Factors of Insider Trading in China --- Based on the Investigation of Different Periods[J]. *Journal of Systems Science and Mathematical Sciences*, 2021, 41(7): 1817-1833.
- [3] Jebran, K., Iqbal, A., Bhat, K. U., Khan, M. A., & Hayat, M. (2019). Determinants of corporate cash holdings in tranquil and turbulent period: evidence from an emerging economy. *Financial Innovation*, 5(1). <https://doi.org/10.1186/s40854-018-0116-y>
- [4] Lei, L., Zhang, Y., Wei, Y., & Zhang, Y. (2020). Forecasting the volatility of Chinese stock market: An international volatility index. *International Journal of Finance & Economics*, 26(1), 1336–1350. <https://doi.org/10.1002/ijfe.1852>
- [5] Yang, D., & Xian, J. (2018). The Correlations Among World Development Indicators. 2018 15th International Computer Conference on Wavelet

Active Media Technology and Information Processing (ICCWAMTIP).
<https://doi.org/10.1109/iccwamtip.2018.8632595>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

