# The Impact of 5G-Relevant Factors on the Stock of Media Industry Based on Multivariate Regression

Hongbin Shi[1](✉) and Xiaoyang Yuan[2]

[1] Tourism College, Hainan University, Hainan, China
171847285@masu.edu.cn
[2] School of Logistics and E-commerce, Zhejiang Wanli University, Ningbo, China

**Abstract.** Contemporarily, the 5G technology is rapidly, and the media industry (e.g., Tiktok and micro-blog software) are growing rapidly and playing an increasingly important role in society. Intrinsically, the media and 5G technology industries might have a strong correlation with each other in stock market. On this basis, we analyze the time-series data of the corresponding underlying assets to investigate the relationship in deep. Based on the regression analysis, there is a positive correlation between the income impact of 5G network sector stocks and media sector stocks. To quantitively explore and analyze the situation of the media sector in the stock market, multivariate regression models are constructed. According to the statistical metrics, the factors will greatly affect the stock price and can be used to improve the performance of the corresponding models. Overall, these results shed light on stock price prediction of media industry in terms of the 5G relevant information.

**Keywords:** 5G Industry · Stock Prediction · Multivariate Regression

## 1 Introduction

The fifth-generation mobile communication technology is a new generation of broadband mobile communication technology with the characteristics of high speed, low delay, and large connection. With the continuous improvement and update of the 5G base station architecture, the existing 5G network architecture, from the traditional BBU, RRU, antenna reconfiguration, the original BBU non-real-time part of the LI was split out separately, and support the backpropagation synergy needs, comprehensive support needs, low latency needs, can create the conditions for the ability to further for the front, backpropagation and core network transmission network between the high-speed conversion of communication signals [2, 3, 5].

It is the network infrastructure to realize man-machine and object interconnection. 5G has great impacts on various aspects of society, such as the media industry. The

H. Shi and X. Yuan—These authors contributed equally.


D. Qiu et al. (Eds.): ICBBEM 2022, AHIS 5, pp. 791–799, 2023.
https://doi.org/10.2991/978-94-6463-030-5_78

media industry refers to the industrial cluster composed of media entities that disseminate all kinds of information and knowledge. It is a special industry that produces and disseminates all kinds of information products in the form of text, graphics, art, language, image, sound, digital, and symbols, as well as providing all kinds of value-added services. There is a close relationship between the media industry and the 5G network. The modern media industry continues to grow relying on the development of science and technology. At the same time, the 5G industry plays a great role in promoting the media sector. The progress of information technology can strengthen the information communication ability of the media industry [7].

Jiao of Guizhou University of Finance and economics investigates the motivation of stock repurchase of Internet media enterprises and its impact on performance, focusing on the emerging asset industry of Internet media [6]. Based on exploring the motivation of Sanqi mutual Entertainment's stock repurchase and its impact on performance, it provides some reference for the stock repurchase of Listed Companies in this industry in the future. Firstly, the text summarizes the relevant research results, theoretical basis, and legal system background of stock repurchase, and analyzes the research path of the impact of stock repurchase on corporate performance. Combined with the financial data of Panax not ginseng mutual entertainment and the motivation of repurchase, it is concluded that the main motivation of stock repurchase in the Internet media industry is the implementation of equity incentive calendar: share repurchase can enhance the confidence of external investors; share repurchase can bring good governance correctors for listed companies, and it can optimize the financial locking effect of listed companies. According to the characteristics of the industry, this paper puts forward four suggestions. It is necessary to improve and innovate the equity plan and encourage the company to make rational use of stock repurchase (Jiao 2021). We should make full use of national policies to choose the time point of Stock Repurchase: improve and develop relevant laws and regulations on the stock repurchase and strengthen the punishment of illegal stock repurchase.

Xu mentioned in Hengyang daily about 5G that at present, the domestic 5G network is developing rapidly, the speed and level of infrastructure construction are constantly improving, and China continues to make phased achievements in high-tech industries [4]. In the meantime, the development of 5G information technology has also brought great promotion to the development of related industries, and the future of 5G technology is immeasurable. Eliopoulos focuses on the methods of the multivariate regression analysis and its suitable environment. He mentioned that multivariate regression analysis has been widely applied in biostatistics, including the medical field. Besides, it is also pointed out that any data analysis has one common goal of figuring out accurate data from the raw and real data. The multivariate regression analysis should be taken as linear when the data is continuous and normal distribution. Imai argues that multivariate regression analysis can also be applied in other survey methodologies. The article suggests new nonlinear least squares and maximum likelihood estimators for the multivariate regression analysis. Based on the analysis, it is also found that the maximum likelihood estimators are more efficient than the alternative estimators (Imai 2011). According to Chung et al., multivariate regression analysis has been developed in geography. It is assumed that the future landslide will be possibly predicted by the past landslides,

based on the GIS technology. All the scholars show that the methodology of multivariate regression analysis is quite useful in the various field for its adaptability [1]. Uyanık et al. give an example of how to correctly use the multiple linear regression. They collect the data from a certain department of a university, defining several factors to estimate the relationship between the variables and the dependent variable. ANOVA statistics are used in predicting the KPSS score among the students [8]. Zhou of Harbin Institute of Technology analyzes the impact of the equity pledge of controlling shareholders of listed cultural media companies on the value of the company. It is believed that the assets of cultural media industry companies are mostly soft, e.g., copyrights and property rights of film and television work, with few heavy assets and a lack of fixed assets, which can hardly be financed by conventional traditional financing methods. Due to the excellent characteristics of easy delivery and high liquidity of equity, equity pledges can facilitate financing for listed companies. The number of listed companies in the culture and media industry that choose to raise funds by pledging their equity is increasing because of the relatively good quality of their shares. Unlike other industries, the equity of listed companies in the culture and media industry is easily liquid when used as pledges for financing [10]. Yang et al. mention that the development of information technology has greatly promoted the progress of human society. 5G era provides more convenience for people's life and work, while its computer network information security issues are gradually attracted people's attention. On this basis, the paper firstly elaborates on the concept and advantages of 5G network, briefly analyzes the 5G network control technology, and then triggers the consideration of computer network information security problems in the 5G era, and seeks countermeasures that can effectively solve the 5G network information security problems, to help the smooth development of 5G network technology in China [9].

In this paper, we chose the Price Earnings ratio (PE) and Market Value (MV) indicators of the 5G sector and the CMCC stock. Moreover, we selected several stocks from the 5G sector, concentrating on close price as the influential factor. Meanwhile, one typical stock in the media industry is chosen as a regression target in the multiple linear regression models. Based on analyzing the data, we conclude some important views on it and finally discussed deeply the short of our research. Besides, the future work outline is also proposed for this research.

The rest part of the paper is organized as follows. Section 2 will introduce the origination of the data and demonstrate the analysis model. Subsequently, Sect. 3 will present the results of different models and evaluations accordingly. Eventually, a summary is given in Sect. 4.

## 2 Data and Method

The data is gathered from multiple sources at various time points from January 3, 2020, to December 24, 2021. We collected the data from the Wind with the total samples of 103, and the sampling frequency is weekly. Specifically, the 5G sector's Market value, CMCC Market value, 5G sector's Price Earning ratio, CMCC Price Earning ratio, the close price of 002786, 6032285, 300287, 300620, 002815,300571, 000607 is collected. Figure 1 gives the visualization of the data.

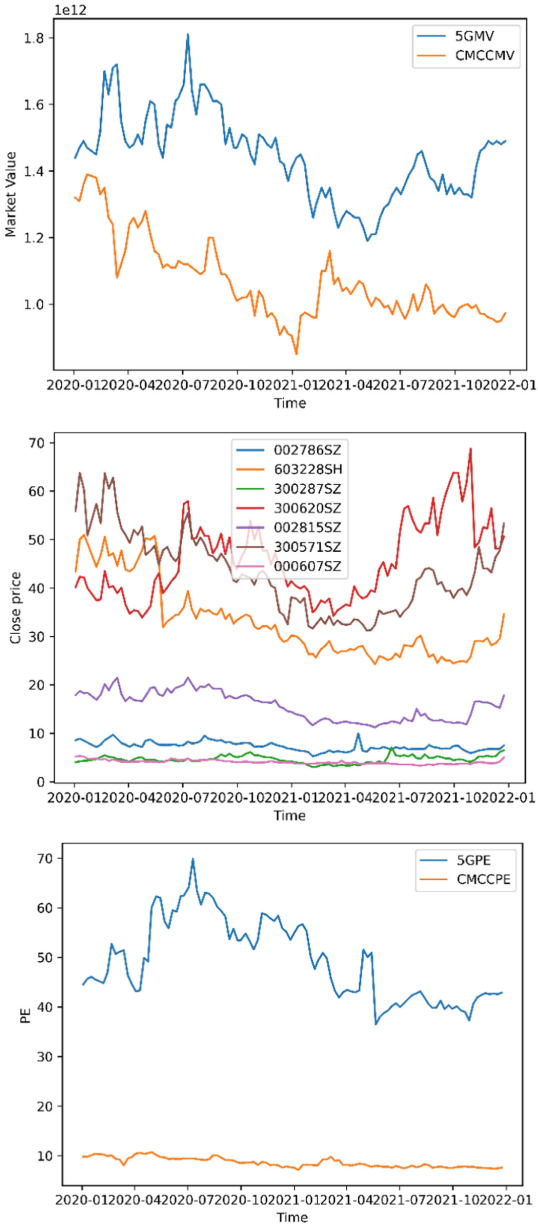


**Fig. 1.** The evolution of variables

The multivariate linear regression is one of the most common procedures for determining how the event will be changed by the moving of the potential factors. The multivariate linear regression approach was chosen to exhibit the relationship between the influential factors and target. The advantage of this particular method is that it allows

us to make predictions condition of the dependent variable. The basic descriptions can be written as:

$$y = \beta_0 + \Sigma_i \beta_i x_i + \xi \tag{1}$$

where y is the dependent variable, β is the coefficient, $x_i$ is independent variables and ξ represents the error. To comparison, we design 3 models. In model 1, the x data includes the Market Value (MV) and the Price Earning ratio (PE) of the 5G sector with another stock CMCC. In model 2, we used the close price of 6 stocks (002786, 603228, 300287, 300620, 002815) as factors. The 000607 is defined as y and the last model contains all the variables above.

There are several assumptions about multiple linear regression.

- Normal distribution: the variables should obey the normal distribution.
- Linearity: the relationship between the variables and target is linear
- The mean of ξ is 0 and there are no autocorrelations between ξ.

To evaluate the performance (accuracy) and statistically significance of the regression models, several metrics are applied including

- R-Squared scores: also known as the coefficient of determination. The values in single regression are between 0 and 1. The higher the R-Squared, the better the model fits the data. The mathematical description is $R^2 = 1 - SS_{res}/SS_{total}$.
- Mean squared error (MSE): used for evaluating model fit. It measures the average of the squared difference between the observed value and the actual value. The lower the value is, the more accurate the data predict. The minimum is 0. The mathematical description is $MSE = \sum_{i=1}^{m} \left(y_i - \widehat{y_i}\right)^2 / m$.
- Correlation: measures the degree of linear dependence between y and all of x. The mathematical description is $\rho = Corr(X, Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$.

## 3 Results and Discussion

Base on the regression analysis results, the regression results can be described as below. The Model 1 is

$$\begin{aligned} P_y = 6.279MV_{5G} \times 10^{-13} + 0.018PE_{5G} + 3.575MV_{CMCC} \\ \times 10^{-12} - 0.241PE_{CMCC} + 0.497 \end{aligned} \tag{2}$$

The Model 2 is

$$\begin{aligned} P_y = 0.056P_{x1} - 0.010P_{x2} - 0.082P_{x3} - 0.012P_{x4} \\ +0.059P_{x5} + 0.026P_{x6} + 2.877 \end{aligned} \tag{3}$$

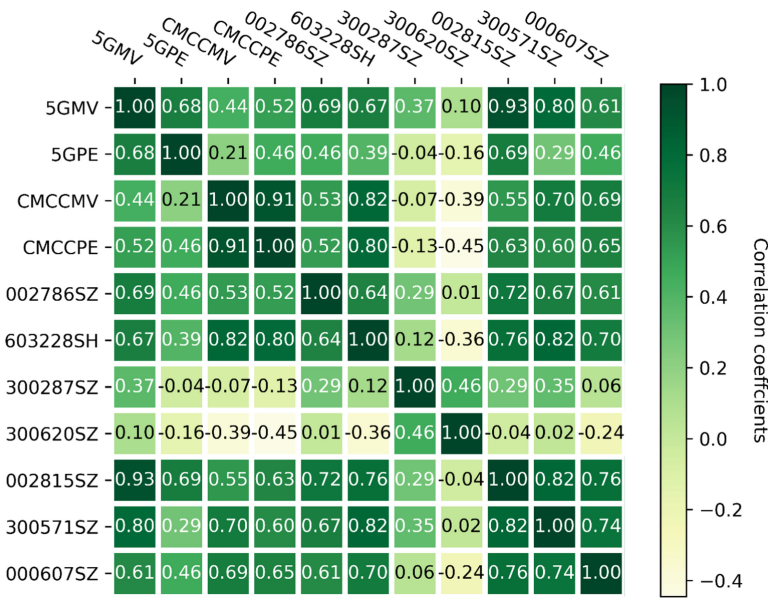


**Fig. 2.** The correlation coefficients of the variables

The Model 3 is

$$\begin{aligned} P_y = &-2.193MV_{5G} \times 10^{-12} + 0.014PE_{5G} \\ &+2.188MV_{CMCC} \times 10^{-12} - 0.169PE_{CMCC} \\ &+0.023P_{x1} - 0.018P_{x2} \\ &-0.022P_{x3} - 0.007P_{x4} + 0.129P_{x5} \\ &+0.025P_{x6} + 3.392 \end{aligned} \tag{4}$$

Here the P represents for price, the footnote y of the price represents for 000607SZ and the footnote x1–x6 represents for, 002786, 603228, 300287, 300620, 002815, 300571 respectively.

The first step in formulating the equations is to analyze the correlations between the independent variables (seen from Fig. 2). Several pairs of independent variables show a higher correlation. Finally, we then look at the exact coefficient of each variable in different models, the variables in the model1 consists of the 5G sector's Price Earning ratio (PE) and the Market value(MV), as well as the stock of CMCC. In the model2, the factors are chosen in the 5G concept companies, which are the most typical among the industry. Lastly, model3 is the combination of the former models we see it as a whole. According to the accomplished models, we made tabulation of the metrics. As summarized in Table 1, the correlation in three models is 0.797, 0.833, 0.882, respectively. The MSE is 0.059, 0.049 and 0.778 respectively. The R-Squared is 0.636, 0.694 and 0.036.

In terms of correlation, Pearson Correlation is the most commonly used correlation coefficient, the higher absolute value is, the stronger correlation it has. All the models here are greater than 0.6. In this case, all three models have a strong correlation with the dependent variable. Model 3 correlation is the maximum among them. According

**Table 1.** Metrics of different models

| | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Correlation | 0.797 | 0.833 | 0.882 |
| MSE | 0.059 | 0.049 | 0.778 |
| R2 | 0.636 | 0.694 | 0.036 |

to the MSE, which is to measure the gap between the parameter estimation and the true observation value. Although the correlation in Model 3 is the biggest, the MSE is also the highest, i.e., Model 3 has produced significant error as well. Comparing the rest of the models, Model 2 shows a lower value of MSE. Concerning the fitting degree, R-Squared is the most appropriate indicator to measure. It is not difficult finding from the table that Model 3 has poor performance at R-Squared, however, Model 2 has the highest fitting degree. To summary, despite the high correlation, Model 3 does not well-fitting to the dependent variable. In the rest of the models, Model 2 has lower MSE, both the correlation and the R-Squared are the higher one, i.e., model 2 is the relatively best model to predict the return of the dependent variable.

However, there are several drawbacks to this paper. The total samples are 103, that means the rage is not wide enough to cover all the efficient data. Maybe one should focus on more samples, for example, taking the daily sampling frequency. On the other hand, accounted for the development of 5G, it is the updated trend in the world, hence there is also a limit of resources. Available data is gathered within 3 years, i.e., may cause the lack of data either. In a common situation, the multivariate linear regression should be done a further step, hypothesis testing. Multivariate linear regressions are suitable for the F-test, which allows the degree of freedom to evaluate whether the linear relationship between all independent variables and dependent variables is close. According to Fig. 2, it is discovered that some of the factors have a high correlation between the rest of the factors, indicating the existence of multicollinearity, which is a phenomenon widely discovered in multiple regression models. It can be described as the high relationship among variables, thereby the final result can be misjudged due to the occurrence of multicollinearity. However, it can be examined by specific methods, e.g., the Klein discriminant method. If the correlation coefficient between two factors is greater than the R-square, it can be said that the formula exists the multicollinearity. The process is fairly complicated, maybe it should be simplified because there is so much data to process. In the research process, by establishing an appropriate analysis model, we ultimately get the relationship between 5G sector stocks and media sector stocks. Simultaneously, one sees the correlation of relevant data through model analysis, which plays a vital role in generating the conclusion. In the model, there is a correlation between 5G technology sector stocks and media sector stocks under certain conditions. Additonally, The total independent number is 10, it is better to expand the horizon into 10 or more. The role of the number of factors is very important. Through a reasonable number of factors, we can fully analyze the research purpose. Simlutaneously, the selected ten factors adopt the random principle to fully ensure the fairness and accuracy of the test data and results. On this basis, we can draw the final research conclusion.

## 4 Conclusion

In summary, we investigate the impact of the media industry on stock returns based on the development of 5G network stocks in terms of the multivariate regression. Through the comparative analysis of 5G stocks and media stocks, as well as data research and sorting, we explore the relationship between them. In terms of 5G plate data, we chose several stocks from the 5G sector as analysis objects. Then, we selected one typical stock in media industries. Through the correlation model, the two data are compared, to realize the research of the full text. In the research, it is found that there is a close relationship between the 5G plate and the media plate. Based on the analysis, the impact of relevant social factors on the media sector. At the same time, the role of network information technology in the media field is investigated. Under certain circumstances, the limitations of network information dissemination ability will also affect the development of the media sector. These results shed light on further promotion for the development of the media sector, offering a guideline for the future development direction of information media technology.

## References

1. Chung, Chang-Jo F., Andrea G. Fabbri, and Cees J. Van Westen. "Multivariate regression analysis for landslide hazard zonation." Geographical information systems in assessing natural hazards. Springer, Dordrecht, 1995. 107–133.
2. Ezhilarasan, E., and M. Dinakaran. "A Review on Mobile Technologies: 3G, 4G and 5G." Second International Conference on Recent Trends & Challenges in Computational Models IEEE, 2017:369–373.
3. Guo Huijun, Tang Lei, Li Yong, "Introduction to 5G transmission network requirements and networking technology", Information Technology, https://doi.org/10.19695/j.cnki.cn12-1369, December 2021.
4. Imai, Kosuke. "Multivariate regression analysis for the item count technique." Journal of the American Statistical Association 106.494 (2011): 407–416.
5. Jian, F., et al. "Review and Prospect on the Research of 5G Spectrum." Telecommunications Science (2015).
6. Jiao Fanghui, "The motivation of stock repurchase of Internet media enterprises and its impact on Performance", Economic and Management Sciences, G206-F; F832.51, June, 2021.
7. People. com, "China has initially built the world's largest 5G mobile network", September, 2021.
8. Uyanık, Gülden Kaya, and Neşe Güler. "A study on multiple linear regression analysis." Procedia-Social and Behavioral Sciences 106 (2013): 234–240.
9. Yang Qingfeng, Liu Siyu, Zhao Liping, "5G bearer network technology and optimized networking", Economic and Management Sciences, TP393.08, December, 2021.
10. Zhou Yang, "Analysis on the impact of equity pledge of controlling shareholders of cultural media listed companies on corporate value" Economic and Management Sciences, F275; G206-F, June, 2021.