Valuation Analysis of Stock Prices in Pharmaceutical Industry Based on HP-ARIMA Method--Taking a sample of 8721 from 12 companies as an example

Xinru Ma

Department of Surrey International Institute, Dongbei University of Finance and Economics, Dalian, Liaoning, 116012, China

Corresponding author: maxinru20010304@163.com

Abstract. Share price valuation can help investors avoid risks. Therefore, this paper selects 12 daily opening prices of stocks to study stock price volatility and make predictions. In this paper, the regression analysis was first conducted using the HP-ARIMA method, which followed by OLS model and Tobit model robustness tests. Heterogeneity analysis was then carried out using the ARIMA model, which yielded the following results: (1) The midstream industry economy has a much less radical effect on social development than the upstream and downstream. (2) The long-term trend of the daily opening price of stocks of upstream companies is more volatile, indicating a higher investment risk. The above prediction results are also used to propose countermeasures.

Keywords: Stock Valuation, Daily Opening Price of Stocks, Pharmaceutical Industry, HP Filtering, ARIMA model.

1. Introduction

Firstly, as the economy develops, the standard of living of the population improves and more attention is paid to investment in health. The ageing of the Chinese population increases, which makes people spend much more attenion on the pharmaceutical industry. This, coupled with the introduction of relevant national policies to support them, has made pharmaceutical stocks popular with investors. Secondly, the stock market is an important channel for corporate financing and the price of stocks, as an important piece of information in the stock market, is extremely volatile due to numerous factors, and the impact of these factors on stock prices is often difficult to measure. Therefore, in this paper, we analyse the pattern based on the change of share price in the past period, so as to achieve the forecast of share price. Stock valuation is of great importance to investors.First, it helps them to issue stocks that are undervalued at their current price and to buy them in time to reap the economic benefits when the price rises. Second, stock valuation can help investors determine whether the stocks they buy are overvalued or undervalued and make the decision to sell or continue to hold them, thus helping them to avoid risk and make the most of their economic benefits.

In this paper, the daily opening prices of 12 stocks from July 1, 2019 to July 1, 2022 were selected from four companies in each of the upstream, midstream and downstream sectors with different locations in the pharmaceutical industry chain to study stock price volatility and make forecasts. After decomposing short-term volatility using the HP filter method, long-term trend data was used as the explanatory variable for ARIMA model regression analysis, followed by OLS and Tobit model robustness tests. Then, ARIMA model is uesd for heterogeneity analysis.The HP filter decomposes a smoother trend than the original net value trajectory, which is more conducive to subsequent analysis and research, while the ARIMA model is theoretically mature and is a common forecasting model for processing time series data, which can be used for future economic planning. It can be used as a reference for future economic planning.

2. Review of the literature

2.1 Quantitative investment analysis of the industry

There are many studies at home and abroad on various fields using different methods to establish models for valuation,Yang Yimei (2022) In the process of enterprise valuation and investment practice, the relative valuation method (market sales method) is first chosen to find the undervalued or overvalued investment object. The absolute valuation method (EVA valuation method) is used to assess the intrinsic investment value of the company to determine whether it has the value of investment.[1] Dong Yangfan (2022) used transaction data on used car trading platforms in the valuation of used pure electric vehicles and used a random forest valuation model to find that factors such as motor power, new car price, model, and year on the market had important effects on the price of used pure electric vehicles. [2] Fischer (2018) used an LSTM model to predict the S&P 500 index. Then, he compared the predictions with standard deep neural networks and simple logistic regression, while developing a short-term stock trading strategy . It was eventually concluded that LSTM is a more suitable method for predicting this index. [3] Mingyu Xu, et al. (2021) selected a sample of 3402 A-share listed companies in China from 2018 to 2020. Then they empirically investigated the relationship between ESG performance and corporate value of listed companies by using a fixed-effect model. Finally, they obtained the conclusion that listed companies with good ESG performance have higher enterprise value.[4]

2.2 .Valuation methods for stock prices and financial derivatives

Stock valuation refers to the intrinsic value of listed companies for estimation. There are many domestic and foreign studies related to the valuation of financial derivatives. Bin Zhangand Dianli Zhao (2016) obtained the idea of calculating the value of convertible bonds with special terms by amending the traditional B-S model on the basis of complying with the no-arbitrage theory and adding the redeemable clause of convertible bonds to the theoretical calculation of the model.[5] Ailie Heather Charteris and Barry Strydom analysed the sub-Saharan capital market through a GARCH model and obtained the conclusion that the more liquid the capital is, the greater the volatility of the interest rate of treasury bonds. [6] Pandya (2019) compared the valuation results of the equity free cash flow model with the market price of the stock. In order to judge the extent to which the stock prices of the selected six companies are undervalued or overvalued. [7] Chang Hui (2019) used a B-S model with a GARCH model to value 15 convertible bonds listed from January 1 to June 30, 2018 and obtained the conclusion that adding a resale clause and a redemption clause to the model could more accurately reflect the real value of market convertible bonds. [8] With reference to the operational and financial characteristics of Minsheng Bank, Xie Chenyu (2020), after valuing its share price with a dividend discount model and validating the valuation results using a price-to-net ratio model and a price-to-earnings model. Ultimately, he concluded that Minsheng Bank's current share price is significantly below its intrinsic value. Therefore, the stock has high investment value. [9]

2.3 Review of the literature

In the valuation of different fields, different scholars use relative valuation method, absolute valuation method, etc. In the valuation of financial derivatives,many valuation methods from other fields are still applicable to the valuation of financial derivatives, such as absolute valuation method. Also common in the valuation of financial derivatives are the GRACH model and the B-S model.

Among the existing stocks valuation models for pharmaceutical industry, the absolute valuation model is highly sensitive to subjective factors. The relative valuation model has logical deficiencies in the way comparable companies are selected and the analytical frameworks such as industry analysis and risk analysis in the fundamental factor valuation model are not yet complete. And this paper differs from the previous studies in the following two aspects: (1)research data innovation, the data is time-sensitive and industry-specific representative. (2) Innovation in research methodology, i.e.

interdisciplinary and econometric approaches are combined to estimate the long-term trends of individual stocks.

3. Pharmaceutical industry and research data

3.1 Overview of the pharmaceutical industry

Based on the structure of the industry chain, different companies in the pharmaceutical industry are classified into upstream, midstream and downstream. Their stocks are selected for analysis. The upstream of the pharmaceutical industry includes Chinese herbal medicines, raw materials.The midstream includes various biological drugs, Chinese medicine and device manufacturing. The downstream includes various offline channels and online channels. This paper choosing upstream enterprises: 300347, 600276, 603259 and 300759, midstream enterprises: 002007, 000538, 300003 and 600518, and downstream enterprises: 300015, 600763, 603882 and 603883, a total of 12 stocks of 12 pharmaceutical enterprises for research.

3.2 Study data and descriptive statistics

This paper selects a total of 12 stocks of companies in the upstream, midstream and downstream of the pharmaceutical industry chain as the research sample. The data related to different companies' stock products are obtained from CSMAR, company annual reports, etc. See Table 1 for specific values.

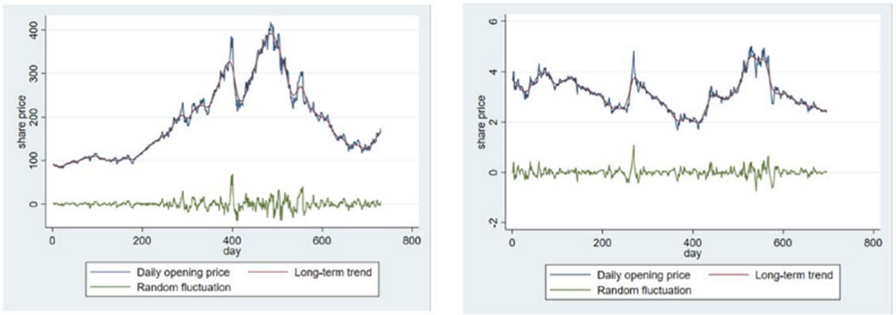
Table 1 Descriptive statistical analysis of daily opening prices per unit for a sample of stocks in the pharmaceutical industry

Share Price Code	Obs.	Mean	Std.Dev.	Min.	Max.
300347	730	110.730	39.368	49.100	206.180
600276	730	73.808	22.927	27.010	116.000
603259	730	113.545	26.340	61.810	188.000
300759	730	112.098	52.351	33.290	233.080
002007	730	37.236	10.307	15.420	70.000
000538	730	94.747	18.411	53.420	162.010
300003	730	29.162	6.551	16.310	45.850
600518	697	3.131	0.655	1.680	5.010
300015	724	48.612	14.675	26.840	92.590
600763	730	190.746	85.482	83.010	417.550
603882	730	91.002	32.601	34.990	176.000
603883	730	63.378	17.271	28.740	113.600

4. Empirical analysis of daily opening price prediction for stocks in the pharmaceutical industry

4.1 Long-term trends in the daily opening prices of stocks in the pharmaceutical industry

The HP filter was first proposed by Hodrick and Prescott in their analysis of the post-war economic boom in the U.S. The most classic application of this method is in the field of macroeconomics, where the HP filter is based on the theoretical approach of spectral analysis of time series, decomposing the trend and cyclical terms to facilitate the corresponding study. The HP filter is a high-pass filter in which the time series is considered as a superposition of components of different frequencies. The High-Pass filter of the time series is to separate out the cyclical fluctuations and trend elements of different nature from all these components of different frequencies. Firstly, this paper makes HP filter analysis on the original data of daily opening price of stock. Then obtaining short-term fluctuation data and long-term trend data is the second step. Finally, separate the long-term trend data for follow-up research. The software chosen for data processing and analysis in this paper is Stata16, and the results obtained are shown in Figure 1.



600763 600518

Figure 1 Long-term trend in daily opening prices of selected sample stocks

From Figure 1, we can see that the HP filter decomposes a smoother trend compared to the original net trajectory, i.e. The red curve represents the long term trend. The green curve represents the short term fluctuation below the setting. Analysis of the HP filtered trend shows that the daily opening prices of 600763 are more volatile in relation to the long-term trend. 600518 daily opening prices are more stable in relation to the long-term trend.

4.2 Forecast analysis of daily opening prices of stocks in the pharmaceutical industry

The ARIMA model is a well-known time series forecasting method proposed by Box and Jenkins. In this paper, the ARIMA model is used for forecasting analysis because it combines the advantages of AR and MA models and meets the condition that the data is time series data. Once this model is identified it is possible to predict future values from the past and present values of the time series. The ARIMA (p, d, q) model can be

$$(1-\sum_{i=1}^{p}\phi_i L^i)(1-L)^d x_t = (1+\sum_{i=1}^{q}\theta_i L^i)\varepsilon_t \quad (1)$$

where L is the lag operator, d∈ Z, d > 0 [10]. The selected stocks' daily opening price long-term trend data was analysed using Stata16 and the results obtained are shown in Table 2.

Table 2 Forecast analysis of unit-day opening prices of sample stocks in the pharmaceutical industry

Stock Code	Ytitle (Daily opening price_Long-term trend) Xtitle (Day) Coef.	cons Coef.	P>|Z |	Wald
300347	1.08E-01	7.12E+01	0.000	146.33
600276	-7.51E-02	1.01E+02	0.000	535.35
603259	6.18E-02	9.10E+01	0.000	167.96
300759	1.81E-01	4.58E+01	0.000	213.81
002007	-2.27E-02	4.55E+01	0.000	212.01
000538	3.06E-03	9.36E+01	0.332	0.62
300003	-1.87E-02	3.60E+01	0.000	320.91
600518	-6.26E-04	3.35E+00	0.000	10.51
300015	9.79E-03	4.51E+01	0.000	9.69
600763	1.84E-01	1.24E+02	0.000	87.84
603882	8.08E-02	6.15E+01	0.000	162.99
603883	-5.92E-02	8.50E+01	0.000	455.57

As can be seen from Table 2, the explanatory variable is the long-term trend in the daily opening price of the stock and the explanatory variable is the number of days the stock is traded. The core coefficients, constant terms as well as the p-values and White's test values are obtained from the ARIMA model analysis. The core coefficient represents the average increase (decrease) in the net value of fund units by a

unit for each additional day of trading (a represents the core coefficient for each stock after regression analysis). The White's test indicates the level of significance of the equation, while the White's test for 000538, 600518, 300015 and 600763 is relatively small compared to the other stocks and is therefore less significant. The p-value for all constant terms is 0. The p-value for the core coefficient of 000538 is 0.332, which is greater than 0.1 and is not significant compared to the core coefficients of the other stocks.

4.3 Robustness tests for daily opening price prediction analysis of pharmaceutical industry stocks

After the HP filter analysis and ARIMA analysis of the selected stock data, the OLS model and Tobit model are chosen to ensure the validity of the prediction. The results are shown in Table 3, where the explanatory variable is the long-term trend of daily opening price and the explanatory variable is the number of trading days of the stock. As can be seen from Table 3, the R^2 and T-values for 000538, 600518, 300015 and 600763 are small when the OLS model analysis is conducted, indicating that the regression curves are poorly fitted to the true data points and were less significant. The remaining stocks are largely significant at their level of significance. This is generally consistent with the results obtained from the ARIMA model analysis.

However, when OLS analysis is performed, there are problems of contingency and low degree of curve fitting. The robustness test of the Tobit model is therefore also carried out and the results are shown in Table 3. We can conclude from the robustness test that the core coefficients obtained from the robustness test are consistent with the OLS core coefficients, indicating good robustness. There is no significant difference in the t-test to values obtained from the OLS and Tobit model tests, indicating that their significance levels are basically the same.

Table 3 Robustness tests for the analysis of unit-day opening price forecasts for the pharmaceutical industry sample stocks

	OLS model				Tobit model		
Core Variable	Coef.	Standard Error	T-value	R^2	Coef.	Standard Error	T-value
300347	1.08E-01	5.52E-03	19.60	0.3454	1.08E-01	5.51E-03	19.62
	7.12E+01	2.33E+00	30.55		7.12E+01	2.33E+00	30.59
600276	-7.51E-02	2.87E-03	-26.20	0.4853	-7.51E-02	2.86E-03	-26.23
	1.01E+02	1.21E+00	83.72		1.01E+02	1.21E+00	83.84
603259	6.18E-02	3.85E-03	16.07	0.2618	6.18E-02	3.84E-03	16.09
	9.10E+01	1.62E+00	56.03		9.10E+01	1.62E+00	56.11
300759	1.81E-01	6.16E-03	29.44	0.5436	1.81E-01	6.15E-03	29.49
	4.58E+01	2.60E+00	17.65		4.58E+01	2.59E+00	17.67
002007	-2.27E-02	1.56E-03	-14.59	0.2263	-2.27E-02	1.55E-03	-14.61
	4.55E+01	6.57E-01	69.34		4.55E+01	6.56E-01	69.43
000538	3.06E-03	3.15E-03	0.97	0.0013	3.06E-03	3.15E-03	0.97
	9.36E+01	1.33E+00	70.46		9.36E+01	1.33E+00	70.55
300003	-1.87E-02	8.91E-04	-21.04	0.3781	-1.87E-02	8.90E-04	-21.07
	3.60E+01	3.76E-01	95.81		3.60E+01	3.75E-01	95.95
600518	-6.26E-04	1.15E-04	-5.45	0.0409	-6.26E-04	1.15E-04	-5.45
	3.35E+00	4.63E-02	72.28		3.35E+00	4.63E-02	72.38
300015	9.79E-03	2.53E-03	3.86	0.0203	9.79E-03	2.53E-03	3.87
	4.51E+01	1.06E+00	42.49		4.51E+01	1.06E+00	42.55
600763	1.84E-01	1.32E-02	13.96	0.2111	1.84E-01	1.32E-02	13.98
	1.24E+02	5.56E+00	22.22		1.24E+02	5.55E+00	22.25

603882	8.08E-02	4.79E-03	16.85	0.2806	8.08E-02	4.79E-03	16.87
	6.15E+01	2.02E+00	30.39		6.15E+01	2.02E+00	30.44
603883	-5.92E-02	2.02E-03	-29.22	0.5398	-5.92E-02	2.02E-03	-29.26
	8.50E+01	8.54E-01	99.54		8.50E+01	8.53E-01	99.67

Based on the results of the analysis, empirical forecasts of stock opening prices are made for the 1st and 6th trading days in the future, i.e. the 731st and 736th trading days, corresponding to the trading dates of 4th July 2022 and 11th July 2022 respectively, using the results obtained from OLS regression analysis for forecasting. The Tobit model for robustness testing. By testing, the difference between the predicted and actual values is not significant, while the model predicts the share prices of midstream companies with relatively higher accuracy.

4.4 Heterogeneity analysis of opening price prediction analysis of pharmaceutical industry stocks

Based on the size of the underlying stock company's assets, it is divided into the largest, larger and smaller. The smallest four components are used to study the stock unit price volatility relationship. The daily opening prices of stocks are filtered by HP to remove their short-term fluctuations and the data is standardised before regression analysis is conducted using the ARIMA model, the results of which are shown in Table 4.

Table 4 Heterogeneity analysis of the forecast analysis of the daily opening price per unit of stocks in the pharmaceutical industry sample

Asset size	Stock Code	Ytitle (Standardaized values of daily opening price) Xtitle (Day) Coef.	Cons Coef.	P>|Z |	Wald
Maximum	603259	2.35E-03	-8.58E-01	0.000	162.79
	000538	1.66E-04	-6.07E-02	0.436	0.61
	600276	-3.28E-03	1.20E+00	0.000	517.70
Larger	300347	2.75E-03	-1.00E+00	0.000	148.00
	300015	6.67E-04	-2.42E-01	0.002	9.29
	300003	-2.86E-03	1.05E+00	0.000	310.34
Smaller	300759	3.46E-03	-1.27E+00	0.000	223.73
	600518	-9.56E-04	3.34E-01	0.001	11.07
	603883	-3.42E-03	1.25E+00	0.000	436.66
Minimum	002007	-2.20E-03	8.06E-01	0.000	192.12
	603882	2.48E-03	-9.06E-01	0.000	160.34
	600763	2.15E-03	-7.86E-01	0.000	83.51

Looking at each value in turn after the regression analysis in Table 4, the core coefficients for 600276, 300003, 600518, 603883 and 002007 are negative, indicating that the number of days a stock is traded is negatively correlated with the daily opening price of the stock. The White's test shows that among the four different asset sizes, one firm in each class has a low level of stock price significance, namely 000538, 300015, 600518 and 600763. 000538's core coefficient has a P-value of 0.436, which is greater than 0.1 and is not significant compared to the core coefficients of the other stocks, while the core coefficients of the other stocks The P-values for the core coefficients of all other stocks are less than 0.01, with a high level of significance.

5. Conclusions of the Study

5.1 Conclusions of the Study

In this paper, the opening prices of 12 stocks from July 1, 2019 to July 1, 2022 are selected from four companies in different positions in the pharmaceutical industry chain, each in the upper, middle and lower reaches, to study and forecast stock price volatility. The following conclusions are drawn from the study after conducting OLS and Tobit model robustness tests.

(1) The descriptive statistical analysis of the daily opening price of stocks in the pharmaceutical industry sample shows that the average daily opening price of stocks related to upstream and downstream is relatively similar, while the average daily opening price of stocks of midstream companies is lower. The midstream industry economies are much less radical to social development than the upstream and downstream.

(2) The midstream companies' daily opening prices fit the model to a relatively higher degree. Based on the hp filter analysis, it can be seen that the long-term trend of midstream and downstream companies (except 600763) tends to be stable. The long-term trend of daily opening prices of stocks of upstream companies fluctuates more significantly. This indicates a high investment risk.

5.2 Policy Insights

In response to the above two research findings, this paper corresponds to them one by one and gives two policy insights.

(1) The uneven development of the upper, middle and lower reaches of the pharmaceutical industry chain is detrimental to the development of the pharmaceutical industry. The state can make some relevant policies to support, thus promote the coordinated development of the pharmaceutical industry.

(2) The risk-return ratio of upstream enterprises needs to be strengthened. Because the system of China's policy changes frequently, thus triggering abnormal fluctuations in stock prices. When appropriate, the government can uses legal and economic instruments to regulate the operation of the stock market

Of course there are shortcomings in this study, as the model only considers the explanatory variable of working days, and external factors are considered in a more homogeneous manner. This paper is more of an exploration of ideas in analysing the pharmaceutical industry economy through stock valuation, with a view to providing a reference for valuation studies of financial investment products.

References

[1] Yang Yimei. Study on the valuation of semiconductor companies in science and technology innovation board[J]. Business News,2022 (12):121-124.

[2] Dong Yangfan. Construction and application of used pure electric vehicle valuation model based on random forest algorithm [D]. Chongqing University of Technology, 2022. DOI :10.27753/d.cnki.gcqgx.2022.000132.

[3] Fischer T, Krauss C. Deep learning with long short-term memory networks for financial market predictions[J]. European Journal of Operational Research, 2018, 270(2): 654-669

[4] Mingyu Xu, et al. An empirical study on the impact of ESG performance of listed companies on enterprise value - An example of A-share listed companies. China Asset Appraisal .07(2021):27-37.

[5] Bin Zhang, Dianli Zhao. The Pricing of Convertible Bonds with a Call Provision[J]. Journal of Applied Mathematics and Physics,2016,4 (6).

[6] Ailie Heather CharterisBarry Strydom. Capital market openness and volatility: an investigation of five Sub-Saharan treasury bill rates[J]. International Journal of Emerging Markets. 201611 (3):438-459

[7] Bhargav Pandya. Application of Free Cash Flow to Equity Model in Valuing Mahindra Group Companies: An Empirical Study[J]. SJCC Management Research Review,2019,9(1).

[8] Chang Hui. Research on the pricing of convertible bonds in China based on B-S option pricing model [D] Shandong University 2019

[9] Chenyu Xie. Stock value analysis based on DDM model and relative valuation method—take Minsheng Bank (600016) as an example[J] Modern Business 2020(12):91-92

[10] Ying Chen, Research on financial data based on time series analysis [D] Dalian: Dalian University of technology, 2019