Asian Journal of Information and Communications 2020, Vol.12, No. 1, 156-170

KISDI KOREA INFORMATION SOCIETY DEVELOPMENT INSTITUTE Asian Journal of Information and Communications

Patent implemented time series algorithm for building stock portfolios in China A-shares

Yu-Jing Chiu*

Chung Yuan Christian University, Taiwan

Kuang-Chin Chen**

Chung Yuan Christian University, Taiwan

Hui-Chung Che***

Shenzhen TekGlory Intellectual Property Data Technologies, Ltd., China

Abstract

Patent, a legal representation of innovation achievement, is strongly meaningful for almost every country's economic growth and technology development. China, with the world's No.2 stock market, is the world largest patent application country. In this study, we observed 2,197 China's listed companies of common stocks (A-shares). The panel data consisting of 570 valid patent indicators and the financial indicators including the stock price, the price-to-book ratio (PB), and the price-to-earnings ratio (PE) of the A-shares through 2016Q4 to 2018Q3 were examined. We constructed patent leading indicators and patent prediction equations for predicting these financial indicators via the Granger Causality test and the dynamic time series forecast model. We found that stock portfolios constructed by the higher predictive PB growth rate, the higher predictive PE growth rate, and the higher predictive stock price growth rate have outstanding performance than the market trend. The underlying concept behind this study is that though the overall economic environment fluctuated to decline and the China-US Trade conflicts, the patent implemented time series algorithm proposed was proved to be useful to discover good stock portfolios.

Keywords: stock performance, patent leading indicator, patent prediction equation, Granger Causality test, dynamic time series forecast model, China A-share

1. Introduction

* yujing@cycu.edu.tw

** jimchenw9889@yahoo.com.tw

*** Corresponding Author: drcharlie918@yeah.net

Global economic growth seems to lose its momentum in 2019. Productivity growth hits a low record, trade wars continue, and economic uncertainty remains high. Despite sluggish market sentiment, innovation is in full swing around the world. In either developed economies or developing economies, innovation activities (which can be measured by R&D spending and new patents) are booming, and innovation spending has been increasing.

As GDP of China surpassed that of Japan, China has become the second largest economy in the world since 2010. The value of China's stock market[1], comprising more than 3,400 listed companies of common stocks (A-shares), is also ranked as world's No. 2.

According to the statistics by the World Intellectual Property Organization (WIPO), China has been the largest patent application country in the world for years. In 2018, there were 4.32 million new China patent applications.[2] By the end of 2018, there were more than 22 million patent publications[3] in China's patent database, which is also world's largest patent database.

The main objective of investing in technological innovation is to have commercial benefits in the future. There should be some relevance between patents and stock market. Due to such enormous amounts of patents and the stock market in China, the relationship is quite interesting for researchers of patent informatics.

The previous related research mostly aimed at US patents and US market. Branch (1974) found that in the early US market from 1950 to 1965, an increase in the number of company's patents usually resulted in predictive growth in sales and profits. Griliches (1981) found a significant relationship between the market value of the firm and its intangible capital, proxied by past R&D expenditures and the number of patents, based on a time-series cross-section analysis of data for large US firms. Cockburn and Griliches (1988) found evidence of an interaction between industry level measures of the effectiveness of patents and the market's valuation of a firm's past R&D and patenting performance, as well as its current R&D moves.

The citation is an important indicator. Hall, Jaffe and Trajtenberg (2005) used patents and citations from 1963 to 1995 and found that citations significantly affect market value, with an extra citation per patent boosting market value by 3%. Branch and Chichirau (2010) found patent counts and patent citations are all positively associated with growth and negatively associated with profitability. Investors who effectively evaluate the quality of R&D performed, may be able profitably to exploit the risk premium applied to the stock of R&D-intensive companies. Crossan and Apaydin (2010) conducted a systematic analysis of research findings in all peer-reviewed literature on innovation in the Social Science Citation Index Database (SSCI) from 1981 to 2008 and found that a company's disclosure of its innovation results is significantly positively related to its earnings. Pandit, Wasley and Zach (2011) examined whether both R&D expenditures and patent counts and citations and their interaction associate with the level and variability of future earnings and operating cash flows.

Hirshleifer, Hsu and Li (2013) found innovative efficiency (IE), patents or citations scaled by R&D expenditures, is a strong positive predictor of future returns after controlling for firm characteristics and risk. Caner, Bruyaka and Prescott (2016) demonstrated the value of a temporal lens in explaining why diversity

[1] There are two stock exchanges in China, one in Shanghai and the other in Shenzhen. The A-share stocks comprise four stock boards: Shanghai Main Board, Shenzhen Main Board, GE board, and SME Board. Stock codes of Shanghai main board start with 600, 601, 603; most stocks are from state-owned listed companies and big companies. Stock codes of Shenzhen main board start with 000, 001, most stocks are also from state-owned listed companies and big companies. Stock codes of GE board start with 300, most stocks are from small & medium companies. Stock codes of SME board start with 002, most stocks are also from small & medium companies.

[2] China patent applications comprise three species of submitted and undisclosed applications: invention applications, utility model applications, and design applications.

[3] China patent publications comprise four species of publications: invention publications, invention grants which disclosed and passed the substantial examination, utility model grants which disclosed and passed the initial examination, and design grants which disclosed and passed the initial examination.

in a firm's patent and alliance portfolios sends flow signals that establish expectations among market observers and have performance implications. Yu and Hong (2016) found that the number of patents has more significant explanatory power than R&D expenditure; incorporating the number of patents in explaining stock returns could add value. Mama (2018) used a large international sample to see if a firm’s innovative efficiency (IE) is positively related to future returns or not. It is found that the relationship is robustly U-shaped. Long-short investment strategies based on highest and lowest IE are inefficient.

Regarding to the quantitative measure of stock performance by patents, Deng, Lev and Narin (1999) and Thomas (2001) proposed to use quantitative patent indicators in modeling company price-to-book ratio via multi-regression analysis for US stock market. Though the adjusted R^2 of the patent modeled equation is low, they applied the patent modeled equation to build the stock portfolio and concluded the high investment return of the stock portfolio. The researchers, especially outside the US, in patent informatics field are inspired by Deng et al. (1999) and Thomas (2001). However, the challenge for researchers is the stock market data retrieval, the patent data collection and patent indicators processing.

Though China is the largest patent application country in the world, related research is little due to the delay of patent data release. He, Tong, Zhang and He (2018) constructed a Chinese patent database which linked China Intellectual Property Administration (CNIPA) patents to all China A-shares and their subsidiaries in China. Chen, Wei and Che (2018) also constructed a database linking China A-shares, their subsidiaries, their patents and their stock price. Chen et al. (2018) discussed patent indicators and the stock price for the first time based on the patent data and the stock price data of A-shares in Shanghai main board from 2011 to 2017. The concepts of patent leading indicators and the patent leading equation used to predict the stock price are proposed. It seems that under the normal macroeconomic environment, the stock portfolios selected by the higher predictive stock price return rate have better performance than the market trend.

Unfortunately, on March 22, 2018, the US government launched a trade war against China through the tariff system. US President Trump officially signed a trade memorandum, announcing that it would impose tariffs on 60 billion US dollars of goods imported from China and restrict Chinese companies' investment, merges and acquisitions in the United States. On April 4, 2018, the US government released a list of goods subject to tariffs, which would impose a 25% tariff on approximately 50 billion USD of goods imported from China. On April 5, 2018, US President Trump requested to impose additional tariffs on 100 billion USD of Chinese imports. On July 6, 2018, the first batch of 34 billion USD of Chinese goods entering the US began to be subject to a 25% tariff. The China-US Trade conflict not only seriously affects China's exports to the US, but also impacts on the Chinese stock market. From the beginning to the end of 2018, the CSI 300 Index[4] fell by 25.3% and the Shanghai Composite Index[5] fell by 24.6%.

As the background of this China-US Trade conflict, do patent leading indicators proposed by Chen et al. (2018) still work in leading the stock price? Do patent leading indicators exist for other financial indicators and for the whole A-shares? How do we build stock portfolios by patent informatics to have better performance than the market trend? In this study, we revised the model of Chen et al. (2018) and proposed a preferable quantitative model for predicting various financial indicators which comprise the stock price, the price-to-book ratio (PB) and the price-to-earnings ratio (PE). Furthermore, we proposed preferable stock selection strategies based on the predictive financial indicators and discussed the stock performance under the impact of the China-US trade conflict.

[4] CSI 300 Index, code 000300, composed of 300 large-scale, liquidity and most representative high-quality stocks selected in the whole A-shares, represents the top stocks in China.

[5] Shanghai Composite Index, code 000001, composed of all Shanghai A-shares, represents the market trend of the Shanghai Main Board stocks.

2. Methodology

2.1 Panel data, population and sample

A total of 8 quarters from 2016Q4 to 2018Q3 are selected before and after the China-US trade conflict. The population of this study is China A-shares over Shanghai and Shenzhen stock exchanges. But Chinese companies listed in the Hong Kong Stock Exchange and/or overseas are excluded. As of now, the number of whole China A-shares is more than 3,400 and is still increasing.

An effective sample of this study must meet two conditions:

(1) During the eight quarters from 2016Q4 to 2018Q3, it remained listed; and

(2) In each quarter from 2016Q4 to 2018Q3, it had at least one new patent publication for last one year, but no restriction for patent species.

For those A-shares whose subsidiaries' revenue merged with the parent company in the annual report, we assume that patents of subsidiaries have corresponding contributions to the parent company, so patents of such subsidiaries are also merged with the parent company for processing patent indicators.

The panel data comprises all effective samples with their corresponding patent indicators and financial indicators in each quarter from 2016Q4 to 2018Q3.

Table 1 shows effective samples and A-shares statistics. A total of 2,197 effective samples are extracted from the whole 3,467 A-shares. Among them, Shanghai main board has the most of effective samples with 776 samples whose proportion is 35.3% of all effective samples, but the SME board has the highest effective sampling rate, which is equal to 74.8%.

Table 1. China A-share statistics of 2018Q3

	Shanghai main board	Shenzhen main board	GE board	SME board	Total
A-shares	1,389	465	710	903	3,467
Effective samples (proportion)	776 (35.3%)	258 (11.7%)	488 (22.2%)	675 (30.7%)	2,197
Sampling rate for effective samples	55.9%	55.5%	68.7%	74.8%	63.4%

2.2 Patent indicator

Though China is the largest patent application country in the world, most China A-shares have filed foreign patents which are less than 5% of China patents. We only focused on patent indicators processed by China patents because the amount of foreign patents, such as US patents, PCT patents, European patents, etc., is too small to be included.

For boosting industry innovation, the Chinese government has carried out a fee-subsidy policy for new patent applications. Many companies file a large amount of patents to get subsidies, and abandon unimportant patents when the annual fees are due. Therefore, only valid patents[6] are discussed in this study. The valid

[6] Valid patents include: issued and annual fee maintained patents of invention grants, utility model grants and design patents; and unissued invention publications which are under examination.

patent indicators PAij applied in this study are modified by those proposed by Chen et al. (2018). For PAij, i=1 to 10 years data collection interval (for avoiding confusion, 10 is represented by X hereinafter), j=1 to 41, 45 to 60, a total of 57 valid patent indicators per year, and a total of 570 patent indicators for 10 years data collection intervals. The definitions are shown in Appendix. Meanwhile, the data collection interval is calculated based on the last day of each quarter from 2016Q4 to 2018 Q3.

The valid patent indicators are processed by China patent raw data which is officially published by the China Intellectual Property Administration, including data on invention publications, invention grants, utility model grants, design grants, and legal status data thereof.

According to the Kolmogorov-Smirnov test, the original data distribution of valid patent indicators is seriously skewed, so all valid patent indicators applied in this study have been cox-box transformed to reduce the skewness.

2.3 Financial indicator

Three price based financial indicators of A-shares are observed in this study:

(1) Stock price: the closing price of the last trading day of each quarter is selected in this study.

(2) Price-to-book ratio (PB): the ratio of the price per share to the book value per share. PB has been favored by value investors for decades and is widely used by market analysts. Traditionally, any value under 1.0 is considered a good PB value, indicating a potentially undervalued stock. Value investors often consider stocks with a PB value under 3.0, indicating a potentially undervalued stock. In this study, PB on the last trading day of each quarter is selected.

(3) Price-to-earnings ratio (PE): the ratio of the price per share to earnings per share. Strictly speaking, any PE between 15 to 25 is a decent valuation. When companies have PE in single-digits, they tend to be seen as good buys by investors who watch this PE. When a stock gets PE around 30 or 35, generally investors who watch this PE think the company is overpriced. In this study, PE on the last trading day of each quarter is selected.

The data of abovementioned financial indicators are retrieved from the formal annual reports, semi-annual reports, quarterly reports and public information announced by the stock exchanges.

2.4 Patent Leading Indicator (PLI)

The Patent Leading Indicator (hereinafter, PLI) is some specific patent indicator which leads the financial indicator for a leading period. The PLI is obtained via the The Granger Causality test. The Granger Causality test is a statistical hypothesis test for determining whether one time series variable is useful in forecasting another. It is not for determining a true cause-and-effect relationship but for finding a probabilistic account of causality. It uses empirical data sets to find leading/lagging patterns of correlation. The null hypothesis for the Granger Causality test is applied. It assumes that one time series variable x(t) doesn't Granger-cause the other time series variable y(t). If the result of F test gets $p<0.05$, then the null hypothesis is rejected. That means the time series variable x(t) does Granger-cause the time series variable y(t).

In this study, take PB for example, each patent indicator in panel data is applied as one time series variable, and PB in panel data is applied as the other time series variable. If the patent indicator satisfies the Granger Causality test (F test, *$p<0.05$) under the Lag condition, it is defined as the PLI for PB. Lag=1 means that the leading period of the PLI to the financial indicator is one quarter, Lag=2 means 2 quarters, and so on. For the stock price and PE, the PLI is obtained via the the Granger Causality test following the same process.

2.5 Patent prediction equation

The patent prediction equation for each financial indicator is generated via the dynamic time series forecast model as follows:

$$y_t = y_{t-4} + \sum_{i=1}^{n} c_i x_{i,t-4} + e_t$$

wherein, -4 represents Lag=4 quarters; y_t is the financial indicator applied as the dependent variable; $x_{i,t\text{-}4}$ is the PLI applied as the dependent variable when Lag=4 while F test, $p^*<0.05$ satisfied; e_t is the error term.

By previously communicated with financial investment institutions, it is suggested to invest at least one year because the reasonable investment behavior is not short-term speculation. Therefore, Lag=4 for the patent prediction equation is applied in this study. The patent prediction equation is used to generate the predictive value for each of financial indicators after four quarters.

3. Result and discussion

3.1 Patent Leading Indicator

Table 2 shows statistics of our PLI analysis. For the stock price, PB, and PE, not all of 570 patent indicators have statistical significance for prediction. However, we found some PLIs statistically exist for each of Lags for predicting each of financial indicators. For predicting the stock price, the number of the most PLIs is 408 when Lag=2 and the number of the least PLIs is 74 when Lag=4. For predicting PB, the number of the most PLIs is 272 when Lag=1 and the number of the least PLIs is 83 when Lag=4. For predicting PE, , the number of the most PLIs is 198 when Lag=1 and the number of the least PLI is only one when Lag=2 and Lag=3. To sum up, the most PLIs usually occur when Lag=1. As the length of the Lag increases, the number of PLIs tends to decrease.

Table 2. Number of PLIs for financial indicators

Financial Indicator	Number of PLIs			
	Lag=1	Lag=2	Lag=3	Lag=4
Stock Price	369	408	127	74
PB	272	147	129	83
PE	198	1	1	12

3.2 Patent prediction equation

The patent prediction equation is constructed by integrating four prediction modeling periods, namely Period I: 2016Q4 predicting 2017Q4; Period II: 2017Q1 predicting 2018Q1; Period III: 2017Q2 predicting

2018Q2; and Period IV: 2017Q3 predicting 2018Q3. Details of patent prediction equations for the stock price, PB, and PE are respectively shown in Tables 3 to 5.

For predicting the stock price, the patent prediction equation is shown in Table 3, in which SP represents the stock price; subscript -4 represents Lag=4. The adjusted R^2=0.6568 shows that though the patent prediction equation is not good, it might have some prediction capability. Though the number of PLIs in Table 2 is 74, there are only 7 PLIs in Table 3 because in the patent prediction equation, all PLIs are linearly combined and each PLI must satisfy significance p*<0.05 in the combination.

Table 3. Patent prediction equation for the stock pice

Dependent variable	SP			
Independent variable	Coefficient	Std. error	t-statistic	p
C	-5.8913	0.4060	-14.5118	0.0001***
SP_{-4}	0.9106	0.0071	128.5839	0.0001***
$PA104_{-4}$	0.9661	0.3327	2.9038	0.0037**
$PA108_{-4}$	-0.5771	0.2213	-2.6075	0.0091**
$PA160_{-4}$	0.5419	0.1785	3.0361	0.0024**
$PA424_{-4}$	0.4124	0.1666	2.4749	0.0133*
$PA454_{-4}$	-0.7103	0.1715	-4.1411	0.0001***
$PA960_{-4}$	-2.8174	1.0804	-2.6078	0.0091**
$PAX60_{-4}$	3.7089	1.0664	3.4780	0.0005***
Patent prediction equation	$SP = -5.8913 + 0.9107*SP_{-4} + 0.9663*PA104_{-4} - 0.5771*PA108_{-4} + 0.5419*PA160_{-4} + 0.4124*PA424_{-4} - 0.7103*PA454_{-4} - 2.8174*PA960_{-4} + 3.7089*PAX60_{-4}$			
Adjusted R^2	0.6568		p (F-statistic)	0.0001***

p*<0.05, p**<0.01, p***<0.001

For predicting PB, the patent prediction equation is shown in Table 4. The adjusted R^2=0.4052 shows that the prediction capability of the patent prediction equation is not good. Though the number of PLIs in Table 2 is 83, there are only 21 PLIs in Table 4.

Table 4. Patent prediction equation for PB

Dependent variable	PB			
Independent variable	Coefficient	Std. error	t-statistic	p
C	0.4180	0.0255	16.3750	0.0001***
PB_{-4}	0.6016	0.0084	71.6389	0.0001***
$PA110_{-4}$	-0.0477	0.0167	-2.8654	0.0042**
$PA111_{-4}$	-0.0294	0.0092	-3.1900	0.0014**
$PA156_{-4}$	0.1300	0.0528	2.4623	0.0138*
$PA160_{-4}$	0.0479	0.0068	7.0721	0.0001***
$PA202_{-4}$	-0.1261	0.0372	-3.3884	0.0007***
$PA207_{-4}$	0.1250	0.0273	4.5717	0.0001***
$PA256_{-4}$	-0.2436	0.1142	-2.1332	0.0329*
$PA260_{-4}$	-0.0195	0.0060	-3.2425	0.0012**
$PA302_{-4}$	0.1050	0.0486	2.1599	0.0308*

$PA307_{-4}$	-0.0837	0.0306	-2.7325	0.0063**
$PA331_{-4}$	-0.0620	0.0248	-2.4985	0.0125*
$PA337_{-4}$	0.0142	0.0058	2.4728	0.0134*
$PA456_{-4}$	0.8013	0.1902	4.2138	0.0001***
$PA525_{-4}$	0.1444	0.0280	5.1579	0.0001***
$PA556_{-4}$	-1.3070	0.2540	-5.1450	0.0001***
$PA602_{-4}$	-0.1894	0.0354	-5.3520	0.0001***
$PA619_{-4}$	-0.1053	0.0531	-1.9827	0.0474*
$PA631_{-4}$	0.2045	0.0523	3.9134	0.0001***
$PA719_{-4}$	0.1090	0.0521	2.0919	0.0365*
$PA931_{-4}$	-0.1466	0.0461	-3.1809	0.0015**
$PAX56_{-4}$	0.4600	0.1654	2.7802	0.0054**
Patent prediction equation	$PB = 0.4180 + 0.6016*PB_{-4} - 0.0477*PA110_{-4} - 0.0294*PA111_{-4} + 0.1300*PA156_{-4} + 0.0479*PA160_{-4} - 0.1261*PA202_{-4} + 0.1250*PA207_{-4} - 0.2436*PA256_{-4} - 0.0195*PA260_{-4} + 0.1050*PA302_{-4} - 0.0837*PA307_{-4} - 0.0620*PA331_{-4} + 0.0142*PA337_{-4} + 0.8013*PA456_{-4} + 0.1444*PA525_{-4} - 1.3070*PA556_{-4} - 0.1894*PA602_{-4} - 0.1053*PA619_{-4} + 0.2045*PA631_{-4} + 0.1090*PA719_{-4} - 0.1466*PA931_{-4} + 0.4600*PAX56_{-4}$			
Adjusted R^2	0.4052		p (F-statistic)	0.0001***

p*<0.05, p**<0.01, p***<0.001

For predicting PE, the patent prediction equation with two PLIs in it is shown in Table 5. The adjusted R^2=0.0370 shows the prediction capability is poor.

Table 5. Patent prediction equation for PE

Dependent variable	PE			
Independent variable	Coefficient	Std. error	t-statistic	p
C	2.8011	0.0525	53.3194	0.0001***
PE_{-4}	0.1625	0.0092	17.6970	0.0001***
$PA526_{-4}$	-0.0455	0.0122	-3.7368	0.0002***
$PA558_{-4}$	0.7386	0.1847	3.9989	0.0001***
Patent prediction equation	$PE = 2.8011 + 0.1625*PE_{-4} - 0.0455*PA526_{-4} + 0.7386*PA558_{-4}$			
Adjusted R^2	0.0370		p (F-statistic)	0.0001***

p*<0.05, p**<0.01, p***<0.001

By setting Lag=4, the patent prediction equations are successfully constructed for predicting the stock price, PB, and PE. Regarding to the goodness of fit, such as the adjusted R^2 is compared, the stock price patent prediction equation has the best one equal to 0.6568; the PE patent prediction equation has the worst one equal to only 0.0370. Though PB patent prediction equation has the most PLIs, the number of 21 PLIs does not show high relevance to the adjusted R^2.

Since the adjusted R^2 of patent prediction equations is between 0.0370 and 0.6568, it might be inappropriate to use any patent prediction equation for predicting any specific financial indicator. However, the predictive value resulted from the patent prediction equations might be applied for constructing specific stock portfolios which consist of some specific stocks.

3. 3 Stock portfolio performance

The main objective of patent analyses is to understand how investing in technological innovation can have commercial benefits. As the patent prediction equations for financial indicators are established, we try to propose stock selection criteria based on the patent prediction equations. Since Lag=4 is applied in the patent prediction equations, that is, four quarters, so the annual stock return rates are set as the performance evaluator of the stock portfolios.

Because four predicting modeling periods are applied for constructing the patent prediction equations, the stock portfolio performance is also observed over these four predicting modeling periods. In order to objectively compare the performance, we also compare the stock performance of the whole A-shares average and the whole effective samples average over the same periods.

We set up two different stock selection criteria. For the criteria I, the stocks are selected by the higher predictive values of each financial indicator. The patent prediction equations are applied to generate the predictive values of the financial indicators in the first quarter of each period. Top 100, Top 200, and Top 300 stocks selected by the higher predictive values are set as the stock portfolios, and then the averages of actual stock annual return rates are examined. For the criteria II, the stocks are selected by the higher predictive growth rates of each financial indicator. The patent prediction equations are applied to generate the predictive values of financial indicators in the first quarter of each period. When compared with real values in the first quarter of each period, the predictive growth rates correspondingly result. Top 100, Top 200, and Top 300 stocks selected by the higher predictive growth rates are set as the stock portfolios, and then the averages of actual stock annual return rates are examined.

The performance comparison is shown in Table 6. Because of the decline in the overall economic environment, the annual stock return rates of the whole A-shares average from period I to period IV are all negative. Especially due to the impact of the China-US trade conflict, the decline in the periods III and IV is more serious. However, the performance of the whole effective samples average in each period is higher than the whole A-shares average, with an average 1.14% higher. It means that the patent-based stock selection basically has a good effect.

Table 6. Performance comparison of stock portfolios

Stock portfolio	Actual stock return rate				
	Period I	Period II	Period III	Period IV	Average
A-share	-19.42%	-22.77%	-27.85%	-36.86%	-26.73%
Effective samples	-18.12%	-21.51%	-26.94%	-35.79%	-25.59%
SP100	-43.16%	-32.62%	-18.29%	-31.19%	-31.32%
SP200	-40.08%	-33.18%	-24.06%	-33.86%	-32.80%
SP300	-35.92%	-30.16%	-25.06%	-34.72%	-31.47%
PB100	-30.28%	-26.94%	-24.34%	-35.31%	-29.22%
PB200	-32.07%	-27.08%	-25.85%	-36.38%	-30.35%
PB300	-31.59%	-27.31%	-26.27%	-36.71%	-30.47%
PE100	-21.50%	-25.15%	-33.17%	-38.56%	-29.60%
PE200	-23.68%	-27.74%	-32.13%	-38.61%	-30.54%
PE300	-25.41%	-28.01%	-29.83%	-39.73%	-30.75%

SP100R	5.02%	-7.23%	-15.93%	-19.97%	-9.53%
SP200R	4.16%	-1.98%	-14.30%	-22.09%	-8.55%
SP300R	-0.26%	-6.08%	-15.04%	-24.35%	-11.43%
PB100R	10.95%	-4.62%	-12.98%	-23.20%	-7.46%
PB200R	5.07%	-7.54%	-17.38%	-25.30%	-11.29%
PB300R	2.67%	-9.99%	-19.36%	-27.02%	-13.43%
PE100R	14.95%	0.95%	-11.56%	-20.10%	-3.94%
PE200R	11.82%	-1.61%	-14.84%	-23.86%	-7.12%
PE300R	6.94%	-6.03%	-16.76%	-25.15%	-10.25%

SP100, SP200, SP300 stand for stock portfolios of top 100, top 200, top 300 stocks selected by the higher predictive stock price, respectively; PB100, PB200, PB300 stand for top stocks selected by the higher predictive PB; PE100, PE200, PE300 stand for top stocks selected by the higher predictive PE; SP100R, SP200R, SP300R stand for top stocks selected by the higher predictive stock price growth rate; PB100R, PB200R, PB300R stand for top stocks selected by the higher predictive PB growth rate; PE100R, PE200R, PE300R stand for top stocks selected by the higher predictive PE growth rate.

Regarding to the stock portfolios selected by the higher predictive stock price, among four periods, SP100 has the best performance in two periods (periods III and IV), SP300 has the best performance in two periods (periods I and II), and SP100 is preferable for the average of the four periods. However, the higher predictive stock price seems not to be a good investment strategy. Because SP100 is 4.59% less than the whole A-shares average, and is 5.73% less than the effective samples average.

Regarding to the stock portfolios selected by the higher predictive PB, PB100 has the best performance in all periods. Compared with the whole A-shares and the effective samples, PB100 is 2.49% less than the whole A-shares average, and is 3.63% less than the effective samples average. However, PB100 is higher than SP100.

Regarding to the stock portfolios selected by the higher predictive PE, PE300 has the best performance in one period (period III), PE100 has the best performance in three periods (periods I, II, and IV) and is preferable for the average of the four periods. Compared with SP100 and PB100, PE100 is higher than SP100 but less than PB100.

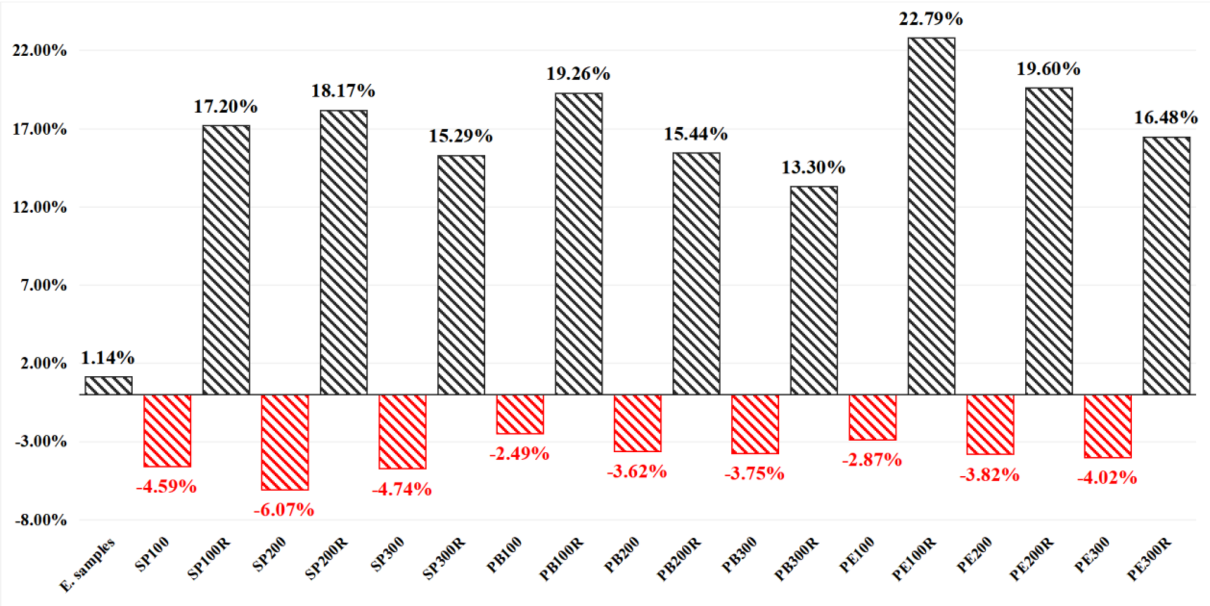
Regarding to the stock portfolio selected by the higher predictive stock price growth rate, SP100R has the best performance in two periods (periods I and IV), SP200R has the best performance in two periods (periods II and III). Compared with the whole A-shares average, each of SP100R, SP200R, SP300R is better, wherein SP200R is preferable for the average of four quarters and 18.17% higher than the whole A-shares average.

Regarding to the stock portfolios selected by the higher predictive PB growth rate, PB100R has best performance in all periods. Compared with the whole A-shares average, each of PB100R, PB200R, PB300R is better, wherein PB100R is preferable for the average of four quarters and 19.26% higher than the whole A-shares average. In addition, PB100R is higher than SP200R.

Regarding to the stock portfolios selected by the higher predictive PE growth rate, PE100R has best performance in all periods. Compared with the whole A-shares average, each of PE100R, PE200R and PE300R is better, wherein PE100R is preferable for the average of four quarters and 22.79% higher than the whole A-shares average and also higher than PB100R.

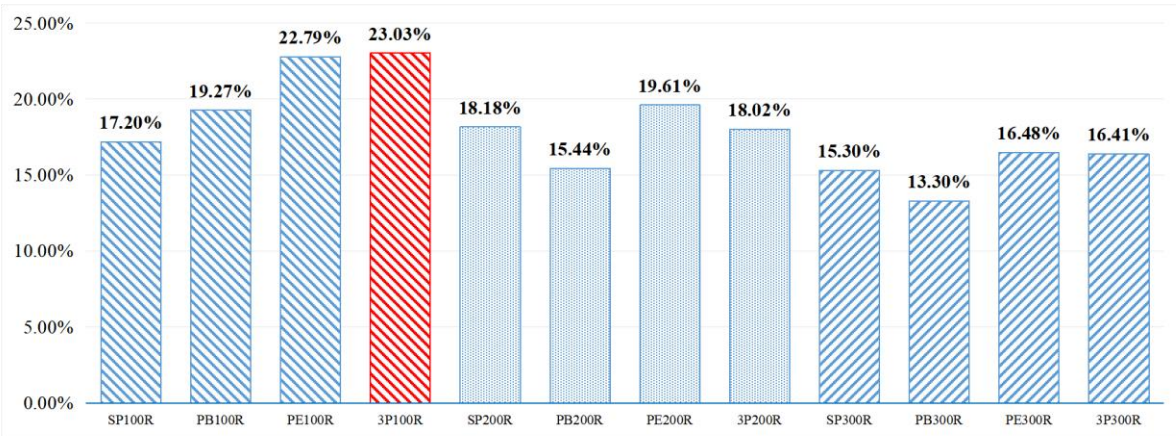
Figure 1 shows the stock performance average comparison of the effective samples, and all preferable stock portfolios when compared with the whole A-shares average, wherein, E.samples stands for the effective samples average and positive/negative values mean higher/less than the whole A-shares average.

Figure 1. Performance comparison of stock portfolios



The stock selection criteria II, in which stocks selected by the higher predictive growth rate, is obviously better than the stock selection criteria I, in which stocks selected by the higher predictive values. All stock portfolios of the stock selection criteria II have better performance than the whole A-shares average and the effective samples average, wherein, PE100R is the best. In addition, as Chen et al. (2018) proposed to select stocks by the higher predictive stock price growth rate, we found the stocks selected by either the higher predictive PB growth rate or the higher predictive PE growth rate have higher performance than the stocks selected according to Chen et al. (2018).

Figure 2. Performance of stocks selected by equal weighting combinations



Meanwhile, for the number of stocks comprised in the stock portfolio, Top 100 is usually the best. As as the number of stocks increases, the performance of the stock portfolio tends to decrease.

For the stock price, PB, and PE, the preferable stock portfolios are SP200R, PB100R and PE100R, respectively. Now we try to combine these three stock portfolios and see the results. The stocks are selected by the higher values of the linear combination with equal weightings of the higher predictive stock price growth rate, the higher predictive PB growth rate, and higher predictive PE growth rate. The performance comparison is shown in Figure 2.

In Figure 2, the whole A-shares average is set as the zero line, 3P100R, 3P200R, and 3P300R stand for stock portfolios of top 100, top 200 and top 300 stocks selected by the higher value of the linear combination, respectively. For top 100, the performance of 3P100R is higher than the other three portfolios. For top 200, the performance of 3P200R is higher than PB200R but less than SP200R and PE200R. For top 300, the performance of 3P300R is higher than SP300R and PE300R but slightly less than PE300R. Obviously, by appropriately adjusting the combination weightings, 3P200R and 3P300R could be easily found to be the best.

4. Conclusion and Recommendation

Based on the patent data from 2016Q4 to 2018Q3 and the three species of financial indicators of China A-shares, this study successfully constructed algorithms for finding PLIs, the patent prediction equations, and the preferable stock selection criteria. The following conclusions were obtained:

(1) Compared with the PLIs for the stock price prediction proposed by Chen et al. (2018), we found that PLIs also exist for PB and PE prediction.

(2) The number of PLIs was usually the largest for one quarter leading, i.e. Lag=1. As the the length of Lag increased, the number of PLIs tended to decrease.

(3) This study constructed patent prediction equations, which consisted of plural PLIs, for quantitatively predicting the stock price, PB, and PE. Among them, the stock price patent prediction equation had the highest goodness of fit with the adjusted R^2=0.6568. But the goodness of fit did not show high relevance to the number of PLIs consisted in the patent prediction equations because the stock price patent prediction equation did not have the most PLIs.

(4) The stock portfolios selected by the higher predictive PB growth rate, the higher predictive PE growth rate, and the higher predictive stock price growth rate have outstanding performance than the market trend. Wherein, PE100R was the best though PE patent prediction equation had the worst goodness of fit. In addition, the stock portfolios selected by either the higher predictive PB growth rate or the higher predictive PE growth rate had better performance than the stock portfolios selected by the higher stock price growth rate which is proposed by Chen et al. (2018).

(5) The stock portfolios selected by the higher value of the linear combination of the higher predictive PB growth rate, the higher predictive PE growth rate, and the higher predictive stock price growth rate with appropriate weightings would show better performance than the stock selected by single predictive growth rate.

(6) Although the overall economic environment fluctuated to decline and the China-US Trade conflicted, the patent based investment algorithm proposed in this study was proved to be useful to discover good

stock portfolios. In addition, the patent based time series algorithm for investment might be also useful in confronting the stock market break by the COVID-19 pandemic. However, it would need more time, more data to observe and examine.

Based on this study, there might be an issue for further research, that is, how to appropriately apply AI technology for modeling patent prediction equations and stock portfolios in order to get more better performance.

References

Branch, B. (1974). Research and development activity and profitability: a distributed lag analysis. *Journal of Political Economy, 82*(5), 999-1011.

Branch, B. and Chichirau, C. (2010). Mispricing vs risk premia in R&D-intensive firms. *International Review of Financial Analysis, 19*(5), 358-367. DOI: 10.1016/j.irfa.2010.06.001

Caner, T., Bruyaka, O. and Prescott, J. E. (2016). Flow signals: evidence from patent and alliance portfolios in the US biopharmaceutical industry. *Journal of Management Studies, 55*(2), 232-264. DOI: 10.1111/joms.12217

Chen, T.-M., Wei, C.-C. and Che, H.-C. (2018). Contribution of patent indicators to China stock performance. *7th International Congress on Advanced Applied Informatics* (ICAAI 2018), 793-798.

Cockburn, I. and Griliches, Z. (1988). Industry effects and appropriability measures in the stock M. *The American Economic Review, 78*(2), 419.

Crossan, M. M. and Apaydin, M. (2010). A multi-dimensional framework of organizational innovation: a systematic review of the literature. *Journal of Management Studies, 47*(6), 1154-1191. DOI: 10.1111/ j.14 67-6486.2009.00880.x

Deng, D., Lev, B. and Narin, F. (1999). Science and technology as predictors of stock performance. *Financial Analysts Journal, 55*(3), 20–32.

Griliches, Z. (1981). Market value, R&D, and patents. *Economics Letters, 7*(2), 183-187.

Hall B. H., Jaffe, A. And Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of Economics, 36*(1), 16-38.

He, Z. L., Tong, T. W., Zhang, Y. and He, W. (2018). Constructing a Chinese patent database of listed firms in China: descriptions, lessons, and insights. *Journal of Economics & Management Strategy, 27*(3), 579-606. DOI: 10.1111/jems.12186

Hirshleifer, D., Hsu, P. and Li, D. (2013). Innovative efficiency and stock returns. *Journal of Financial Economics, 107*(3), 632-654. DOI: 10.1016/j.jfineco.2012.09.011

Mama, H. B. (2018). Innovative efficiency and stock returns: should we care about nonlinearity?, *Finance Research Letters, 24*, 81-89. DOI: 10.1016/j.frl.2017.07.001

Pandit, S., Wasley, C. E. and Zach, T. (2011). The effect of research and development (R&D) inputs and outputs on the relation between the uncertainty of future operating performance and R&D expenditures. *Journal of Accounting, Auditing &Finance, 26*(1), 121-144. DOI: 10.1177/0148558X11400583

Thomas, P. (2001). A relationship between technology indicators and stock market performance. *Scientometrics, 55*(1), 319-333.

Yu, G. J. and Hong, K. H. (2016). Patents and R&D expenditure in explaining the stock price movements. *Finance Research Letters, 19*, 197-203. DOI: 10.1016/j.frl.2016.07.012

Appendix. Valid Patent Indicator

Valid Patent Indicator PAij (i = 1 ~ X) and Definitions	
PAi01	Counts of valid invention publications for last i year(s)
PAi02	Counts of valid utility model grants valid for last i year(s)
PAi03	Counts of valid design grants for last i year(s)
PAi04	Counts of valid invention grants for last i year(s)
PAi05	Average of the patent examination duration of valid invention grants for last i year(s)
PAi06	Total International Patent Classification (hereinafter, IPC) counts of valid invention publications for last i year(s)
PAi07	Total IPC counts of valid utility model grants for last i year(s)
PAi08	Total IPC counts of valid invention grants for last i year(s)
PAi09	Average IPC counts of valid invention publications for last i year(s)
PAi10	Average IPC counts of valid utility model grants for last i year(s)
PAi11	Average IPC counts of valid invention grants for last i year(s)
PAi12	Total specification words of valid invention publications for last i year(s)
PAi13	Total specification words of valid utility model grants for last i year(s)
PAi14	Total specification words of valid invention grants for last i year(s)
PAi15	Average specification words of valid invention publications for last i year(s)
PAi16	Average specification words of valid utility model grants for last i year(s)
PAi17	Average specification words of valid invention grants for last i year(s)
PAi18	Total claim counts of valid invention publications for last i year(s)
PAi19	Total claim counts of valid utility model grants for last i year(s)
PAi20	Total claim counts of valid invention grants for last i year(s)
PAi21	Average claim counts of valid invention publications for last i year(s)
PAi22	Average claim counts of valid utility model grants for last i year(s)
PAi23	Average claim counts of valid invention grants for last i year(s)
PAi24	Total independent claim counts of valid invention publications for last i year(s)
PAi25	Total independent claim counts of valid utility model grants for last i year(s)
PAi26	Total independent claim counts of valid invention grants for last i year(s)
PAi27	Average independent claim counts of valid invention publications for last i year(s)
PAi28	Average independent claim counts of valid utility model grants for last i year(s)
PAi29	Average independent claim counts of valid invention grants for last i year(s)
PAi30	Total drawing counts of valid invention publications for last i year(s)
PAi31	Total drawing counts of valid utility model grants for last i year(s)
PAi32	Total drawing counts of valid invention grants for last i year(s)
PAi33	Average drawing counts of valid invention publications for last i year(s)

PAi34	Average drawing counts of valid utility model grants for last i year(s)
PAi35	Average drawing counts of valid invention grants for last i year(s)
PAi36	Total abstract words of valid invention publications for last i year(s)
PAi37	Total abstract words of valid utility model grants for last i year(s)
PAi38	Total abstract words of valid invention grants for last i year(s)
PAi39	Average abstract words of valid invention publications for last i year(s)
PAi40	Average abstract words of valid utility model grants for last i year(s)
PAi41	Average abstract words of valid invention grants for last i year(s)
PAi45	All valid patent counts for last i year(s)
PAi46	Proportion of valid invention publications in all invention publications for last i year(s)
PAi47	Proportion of valid utility model grants in all utility model grants for last i year(s)
PAi48	Proportion of valid design grants in all design grants for last i year(s)
PAi49	Proportion of valid patents in all invention grants for last i year(s)
PAi50	Average lifespan of valid invention publications for last i year(s)
PAi51	Average lifespan of valid utility model grants for last i year(s)
PAi52	Average lifespan of valid design grants for last i year(s)
PAi53	Average lifespan of valid invention grants for last i year(s)
PAi54	Total backward patent citation counts of valid invention grants for last i year(s)
PAi55	Proportion of inventions publication patents in all valid patents for last i year(s)
PAi56	Proportion of utility model grants in all valid patents for last i year(s)
PAi57	Proportion of design grants in all valid patents for last i year(s)
PAi58	Proportion of inventions grants in all valid patents for last i year(s)
PAi59	Total forward patent citation counts of valid patents for last i year(s)
PAi60	Total backward non-patent citation counts for valid invention grants for last i year(s)