

Research Article

Does Patent Help to Build Investment Portfolio of China A-Shares under China-US Trade Conflict?

Yu-Jing Chiu,¹ Kuang-Chin Chen,¹ and Hui-Chung Che^{1,2} 

¹Department of Business Administration, Chung Yuan Christian University, Taoyuan, Taiwan

²Shenzhen TekGlory Intellectual Property Data Technologies, Ltd., Shenzhen, China

Correspondence should be addressed to Hui-Chung Che; drcharlie918@yeah.net

Received 25 March 2020; Accepted 17 April 2020; Published 13 May 2020

Guest Editor: Chi-Hua Chen

Copyright © 2020 Yu-Jing Chiu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Patent, a legal representation of innovation achievement, is strongly meaningful for almost every country's economic growth and technology development. China, the world's no. 2 stock market, is the world's largest patent application country. In this study, we observed 2,197 China-listed companies of Renminbi (RMB) common stocks (A-shares) from 2016 to 2018. The relationship among 570 valid patent indicators and financial indicators of the stock price, Return-on-Assets (ROA), and Return-on-Equity (ROE) was examined. We constructed patent leading indicators, patent kernel indicators, and patent prediction equations for predicting the stock price, ROA, and ROE using the Granger causality test and the dynamic time series forecast model. The stock performance of investment portfolios based on patent prediction equations was thoroughly discussed. We found that investment portfolios constructed by the higher predictive ROA, the higher predictive ROE, and the higher predictive stock price growth rate have better performance than the A-shares' average. The underlying concept behind this study is that despite the overall economic environment fluctuation and the China-US trade conflict, the patent-based algorithm proposed was proved to be useful to discover good investment portfolios.

1. Introduction

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

When the GDP of China surpassed that of Japan, China became the second-largest economy in the world since 2010. The value of China stock market (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 the Shanghai main board start with 600, 601, and 603; most stocks are of state-owned listed companies and big companies. Stock codes of the

Shenzhen main board start with 000 and 001; most stocks are of state-owned listed companies and big companies. Stock codes of the GE board start with 300; most stocks are of small and medium companies. Stock codes of the SME board start with 002; most stocks are of small & medium companies), comprising more than 3,400 listed companies of RMB common stocks (A-shares), is also ranked as the world's no. 2.

According to the World Intellectual Property Organization, for years, China has been the largest source of patent applications. In 2018, 4.32 million patent applications emanated from China (China patent applications comprise three species of submitted and undisclosed applications: invention applications, utility model applications, and design application). By the end of 2018, the cumulative number of patent publications (China patent publications comprise four species of publications: invention publications, invention grants which have been disclosed and have passed substantial examination, utility model grants which have

been disclosed and have passed the initial examination, and design grants which have been disclosed and have passed the initial examination) in China's database reached 22 million. China's patent database is the largest in the world.

The main objective of investing in technological innovation is to have commercial benefits in the future. Deng et al. [1] and Thomas [2] proposed the use of multiregression analysis for modeling the patent indicators and the price-to-book ratio in the US stock market. Though the adjusted R^2 was very poor, they still found a way to distinguish the undervalued and overvalued stocks. For China, with its enormous number of patents and the size of the stock market, is it possible to find a similar relationship between patent indicators and stock performance? Chen et al. [3] modified the abovementioned multiregression model for discussing the data of Shanghai A-shares from 2011 to 2017. The leading and lagging relationship between the patent indicators and the stock price was roughly modeled. The patent leading equation for predicting the stock price by patent leading indicators was proposed. It seems that under the normal macroeconomic environment, the investment portfolio selected based on the higher predictive stock price return rate performed better 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 the US will impose tariffs on 60 billion US dollars' worth of goods imported from China and restrict Chinese companies' investment, mergers, and acquisitions in the United States. On April 4, 2018, the US government released a list of goods subject to tariffs, which amounted to the imposition of a 25% tariff on approximately US\$50 billion worth in goods imported from China. On April 5, 2018, US President Donald Trump instructed the United States Trade Representative to consider imposing US\$100 billion in additional tariffs on imports from China. On July 6, 2018, the first batch of US\$34 billion worth of Chinese goods entering the US faced the new 25% tariff. The China-US trade conflict not only seriously affects China's exports to the US but also impacts the Chinese stock market. From the beginning to the end of 2018, the CSI 300 Index (CSI 300 Index, code 000300, is composed of 300 large-scale, liquidity, and most representative high-quality stocks selected from the entire list of A-shares and represents the top stocks in China) declined by 25.3% and the Shanghai Composite Index (Shanghai Composite Index, code 000001, comprises all Shanghai A-shares and represents the market trend of the Shanghai main board stocks) declined by 24.6%.

Against the backdrop of the China-US trade conflict, do the patent indicators proposed by Chen et al. [3] hold good in leading the stock price? Do patent leading indicators also work to correctly predict other financial indicators? Do patent-based stocks still perform better than the market trend? This study aims at answering the above questions.

2. Literature Reviews

R&D capability and market structure are important factors for driving a company's growth and maintaining

competitive advantages. Branch [4] found that an increase in the number of patents of a company usually resulted in predictive growth in sales and profits. Griliches [5] found a significant relationship between the market value of a company's shares and the number of patents the company holds. Cockburn and Griliches [6] found an interaction between the effectiveness of patents and the market's valuation of a firm's past R&D and patenting performance. Hall et al. [7] found that citations significantly affect market value. Every additional citation per patent boosts the market value of the stock by 3%. Branch and Chichirau [8] found that patent counts and patent citations are both positively related to growth and negatively related to profitability. Crossan and Apaydin [9] found that the company's innovation results are significantly and positively related to its earnings. Pandit et al. [10] examined whether R&D expenditures and patent counts and citations and their interaction are associated with the level and variability of future earnings and operating cash flows. Fabrizi et al. [11] proposed that venture capitalists are fairly good at judging the value of an idea and whether it is marginally costlier to patent low-value ideas than high-value ideas. Hirshleifer et al. [12] found that innovative efficiency—patents or citations scaled by R&D expenditures—is a strong positive predictor of future returns. Caner et al. [13] theorized that diversity in a firm's patent and alliance portfolios sends contrasting flow signals impacting its market value in a nuanced way. Yu and Hong [14] found that the number of patents has more significant explanatory power than R&D expenditures, incorporating the number of patents in explaining stock returns could add value. Mama [15] found that the relationship of a firm's innovative efficiency and future return is robustly U-shaped.

Deng et al. [1] and Thomas [2] proposed that the main objective of technology analyses is to understand how investing in technological innovation can have commercial benefits. They used quantitative patent indicators in modeling the company price-to-book ratio (PB) by multiregression analysis for the US stock market. They proposed that those companies with patent-modeled PB higher than actual PB are undervalued and those companies with patent-modeled PB less than actual PB are overvalued. The highly undervalued stocks might have high investment potential.

All the above studies are about the US stock market. The relationship between the China stock market and patents is rarely discussed. For this purpose, Chen et al. [3] modified the abovementioned multiregression model to discuss Shanghai A-shares in China from 2011 to 2017. They attempted to use the Granger causality test, time series model, and patent prediction equation for predicting the stock price by patent leading indicators. They showed that the stock selection strategy by the predictive stock return rate yields a better stock performance than the market trend.

In this study, we applied the model of Chen et al. [3] to all A-shares rather than to Shanghai A-shares only to check the model's effectiveness. We also modified the model by setting ROA and ROE as the predictive targets because these are more relevant indicators of the company's managerial and operating performance than the stock price. Furthermore,

we discussed the stock portfolio performance based on predictive stock price, predictive ROA, and predictive ROE. Finally, we proposed the investment strategies that may be preferred as workable under the impact of the China-US trade conflict.

3. Methodology

3.1. Population and Sample. China has two stock exchanges, one in Shanghai and the other in Shenzhen, where China A-shares are traded and are spread into Shanghai main board, Shenzhen main board, GE board, and SME board, and these stocks are included in this study. Chinese companies listed on the Hong Kong Stock Exchange and overseas are excluded. As of now, the total number of 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
- (2) In each of the eight quarters from 2016Q4 to 2018Q3, it had a new patent publication within one year back through the end of each quarter, the patent may be of any of the patent species mentioned in footnote 3 on page 3

For those A-shares of companies whose subsidiaries' revenue is merged with that of 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 patents of the parent company for processing patent indicators.

Table 1 shows effective samples and A-share statistics. A total of 2,197 effective stocks were extracted from all 3,467 A-shares. Of these, the most, that is, 776 effective samples (35.3%) were of the Shanghai main board. However, the largest percentage of effective stocks out of the stocks on any board to be picked was 74.8% from the SME board.

3.2. Patent Indicator. Though China is the leading applicant for patents in the world, only 5% of China's patent applications are filed in foreign countries by A-share companies. The number of patent applications filed abroad for US patents, PCT patents, European patents, etc. is negligible. Therefore, we focused only on patent applications filed in China as the patent indicator.

For boosting industrial innovation, the China government has been subsidizing the fee for filing new patent applications. Many companies apply for a large number of patents to get subsidies and give up unimportant patents when the annual fees for those patents are due. Therefore, only valid patents (valid patents include issued, for which the annual fee has been paid, patents with invention grants, utility model grants, and design patents, and unissued invention publications which are under examination) were included in this study. The valid patent indicators applied in this study are identified as PA_{*ij*}, where *i* = 1 to 10-year data collection interval (to avoid confusion, 10 is represented by

X hereinafter) and *j* = 1 to 41, 45 to 60, a total of 57 valid patent indicators for each data collection interval, and a total of 570 patent indicators for all data collection intervals, and the definition is shown in Table 2. The valid patent indicators are extracted from China patent raw data which published by the China National Intellectual Property Administration and included data on invention publications, invention grants, utility model grants, design grants, and legal status data thereof.

When subjected to the Kolmogorov–Smirnov test, the original data distribution of valid patent indicators was observed to be seriously skewed. Therefore, all valid patent indicators applied in this study were Cox-Box transformed to reduce the skewness.

3.3. Financial Indicator. There are three financial indicators of A-shares used in this study, including the stock price, ROA, and ROE. The financial indicators are calculated based on the information revealed by the stock exchanges, quarterly reports, half-yearly reports, and annual reports.

For the stock price, the closing price of the last trading day of each quarter is considered.

ROA is the ratio of the net profit after tax to total assets. ROA indicates how well a company utilizes its assets. ROA is the best indicator for comparing similar companies or comparing a company's ROA at one point in time with the ROA at various points in the past. In this study, ROA at each quarterly settlement is considered.

ROE is the ratio of net profit after tax to shareholders' equity. ROE is considered a measure of how effectively management is using a company's net assets to create profits. In this study, ROE at each quarterly settlement is considered.

3.4. Panel Data-Modeling Period. 2016Q4~2018Q3, a total of eight quarters, is considered before and after the commencement of the China-US trade war.

3.5. Patent Leading Indicator and Patent Kernel Indicator. The Granger causality test is applied for finding the patent leading indicator (hereinafter, PLI). The Granger causality test is a statistical test of hypothesis for determining whether a 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.

For finding the PLI, each of the PA_{*ij*} is applied as a one-time series variable, and each of the financial indicators is applied as the other time series variables. When any of PA_{*ij*} satisfies the Granger causality test (*F*-test, $p^* < 0.05$) under the lag condition, the PLI is obtained. Lag = 1 means that the leading period of the PLI to the financial indicator is one quarter; Lag = 2 means two quarters; Lag = 3 means three quarters; and Lag = 4 means four quarters. When a PLI satisfies the Granger causality test (the *F*-test, $p^* < 0.05$) for all lags 1 to 4, it is defined as the patent kernel indicator (hereinafter, PKI).

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 (percentage of total)	776 (35.3%)	258 (11.7%)	488 (22.2%)	675 (30.7%)	2,197
Sampling rate for effective samples percentage of top row	55.9%	55.5%	68.7%	74.8%	63.4%

TABLE 2: Valid patent indicator PA_{ij}.

Valid patent indicator ($i = 1 \sim 10$)	Indicator definition
PA _{i01}	Counts of valid invention publications for i years
PA _{i02}	Counts of valid utility model grants valid for i years
PA _{i03}	Counts of valid design grants for i years
PA _{i04}	Counts of valid invention grants for i years
PA _{i05}	Average of the patent examination duration of valid invention publications for i years
PA _{i06}	Total International Patent Classification (hereinafter, IPC) counts of valid invention publications for i years
PA _{i07}	Total IPC counts of valid utility model grants for i years
PA _{i08}	Total IPC counts of valid invention grants for i years
PA _{i09}	Average IPC counts of valid invention publications for i years
PA _{i10}	Average IPC counts of valid utility model grants for i years
PA _{i11}	Average IPC counts of valid invention grants for i years
PA _{i12}	Total specification words of valid invention publications for i years
PA _{i13}	Total specification words of valid utility model grants for i years
PA _{i14}	Total specification words of valid invention grants for i years
PA _{i15}	Average specification words of valid invention publications for i years
PA _{i16}	Average specification words of valid utility model grants for i years
PA _{i17}	Average specification words of valid invention grants for i years
PA _{i18}	Total claim counts of valid invention publications for i years
PA _{i19}	Total claim counts of valid utility model grants for i years
PA _{i20}	Total claim counts of valid invention grants for i years
PA _{i21}	Average claim counts of valid invention publications for i years
PA _{i22}	Average claim counts of valid utility model grants for i years
PA _{i23}	Average claim counts of valid invention grants for i years
PA _{i24}	Total independent claim counts of valid invention publications for i years
PA _{i25}	Total independent claim counts of valid utility model grants for i years
PA _{i26}	Total independent claim counts of valid invention grants for i years
PA _{i27}	Average independent claim counts of valid invention publications for i years
PA _{i28}	Average independent claim counts of valid utility model grants for i years
PA _{i29}	Average independent claim counts of valid invention grants for i years
PA _{i30}	Total drawing counts of valid invention publications for i years
PA _{i31}	Total drawing counts of valid utility model grants for i years
PA _{i32}	Total drawing counts of valid invention grants for i years
PA _{i33}	Average drawing counts of valid invention publications for i years
PA _{i34}	Average drawing counts of valid utility model grants for i years
PA _{i35}	Average drawing counts of valid invention grants for i years
PA _{i36}	Total abstract words of valid invention publications for i years
PA _{i37}	Total abstract words of valid utility model grants for i years
PA _{i38}	Total abstract words of valid invention grants for i years
PA _{i39}	Average abstract words of valid invention publications for i years
PA _{i40}	Average abstract words of valid utility model grants for i years
PA _{i41}	Average abstract words of valid invention grants for i years
PA _{i45}	All valid patent counts for i years
PA _{i46}	Proportion of valid invention publications in all invention publications for i years
PA _{i47}	Proportion of valid utility model grants in all utility model grants for i years
PA _{i48}	Proportion of valid design grants in all design grants for i years
PA _{i49}	Proportion of valid patents in all invention grants for i years
PA _{i50}	Average lifespan of valid invention publications for i years
PA _{i51}	Average lifespan of valid utility model grants for i years
PA _{i52}	Average lifespan of valid design grants for i years
PA _{i53}	Average lifespan of valid invention grants for i years
PA _{i54}	Total backward patent citation counts of valid invention grants for i years
PA _{i55}	Proportion of inventions publication patents in all valid patents for i years

TABLE 2: Continued.

Valid patent indicator ($i = 1 \sim 10$)	Indicator definition
PAi56	Proportion of utility model grants in all valid patents for i years
PAi57	Proportion of design grants in all valid patents for i years
PAi58	Proportion of inventions grants in all valid patents for i years
PAi59	Total forward patent citation counts of valid patents for i years
PAi60	Total backward nonpatent citation counts for valid invention grants for i years

3.6. *Patent Prediction Equation.* The patent prediction equation for each of the financial indicators is generated via the dynamic time series forecast model as follows:

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

wherein -4 is Lag = 4; y_t represents each of the financial indicators applied as the dependent variable; and $x_{i,t-4}$ are the PKIs applied as the dependent variables for Lag = 4, while in the F -test, the requirement $p^* < 0.05$ if satisfied e_t is the error term.

Though we found the PKI for Lag = 1 to 4 quarters, the patent prediction equation is used to generate the predictive value for each of the financial indicators after four quarters. Earlier when we communicated with financial investment institutions, they advised us that an investment must last for at least one year because short-term investment for speculation is not the reasonable investment behavior. Therefore, Lag = 4 is used in the patent prediction equation in this study.

3.7. *Research Steps.* The analysis in this study was conducted through the following steps:

- (1) Collecting effective samples with their financial indicators and processing the corresponding valid patent indicators in the course of panel data modeling
- (2) Using Granger causality test to obtain PLIs and PKIs for each of the financial indicators
- (3) Obtaining patent prediction equations for each of the financial indicators using PKIs and the dynamic time series forecast model
- (4) Discussing the performance of investment portfolios selected by predictive values and predictive growth rates of various financial indicators and comparing with the market trend

4. Result and Discussion

4.1. *Patent Leading Indicator.* Table 3 shows the statistics derived from PLI and PKI analysis. For financial indicators of the stock price, ROA, and ROE, not all 570 patent indicators are statistically significant for the prediction. However, we find the PLIs that are statistically significant for predicting each of the financial indicators, namely, the stock price, ROA, and ROE. PKIs also exist for every financial indicator. Details of PKIs are shown in Table 4.

TABLE 3: Number of PLIs and PKIs for financial indicators.

Financial indicator	PLI & PKI	Lag			
		Lag = 1	Lag = 2	Lag = 3	Lag = 4
Stock price	Number of PLIs	369	408	127	74
	Number of PKIs	29			
ROA	Number of PLIs	305	258	162	148
	Number of PKIs	92			
ROE	Number of PLIs	165	107	95	89
	Number of PKIs	56			

For predicting the stock price, there are a number of PLIs for each of the lags. Among these, the one with the largest number of PLIs is Lag = 2 and it has 408 PLIs. On the other hand, Lag = 4 has 74 PLIs, which is the least number. The number of PKIs is 29. When considering the data collection interval, the longest interval is one year and it has 10 PKIs, and the least is the six-year interval, which has one PKI. As regards patent features, the largest number of PKIs, 10 (PA160~PAX60), is related to “total backward nonpatent citation counts for valid invention grants.” The next is “total independent claim counts of valid invention publications” with 4 PKIs (PA324, PA424, PA524, and PA724).

For predicting ROA, each of the lags has PLIs. The largest number, 305 PLIs, is for Lag = 1, and the least number, 148 PLIs, is for Lag = 4. There are 92 PKIs, the data collection interval, of which the most, 11 PKIs each, are for one-year and two-year intervals. The least number of PKIs, 6, is for the six-year interval. As regards the patent features, the largest number of 10 PKIs (PA130~PAX30) relates “total drawing counts of valid invention publications” and “average drawing counts of valid invention publications” (PA133~PAX33). 9 PKIs (PA160~PA760, PA960~PAX60) relate to “total backward nonpatent citation counts for valid invention grants.”

For predicting ROE, each of the lags has PLIs. The largest number, 165 PLIs, is for Lag = 1, and the least, 89 PLIs, is for Lag = 4. There are 56 PKIs found. Considering the data collection interval, the most, 8 PKIs, are with the two-year interval, while the least number of PKIs, 5, is with one-year, four-year, five-year, six-year, eight-year, nine-year, and ten-year intervals. As for patent features, the most, 10 PKIs (PA135~PAX35), relate to two features, “average drawing counts of valid invention grants” and (PA160~PAX60) “total backward nonpatent citation counts for valid invention grants.” Next, 9 PKIs (PA248~PAX48) relate to “proportion of valid design grants in all design grants” and 9 PKIs (PA252~PAX52) relate also to “average lifespan of valid design grants”.

TABLE 4: PKIs for financial indicators.

Financial indicator	PKI
Stock price	PA104, PA107, PA108, PA114, PA120, PA126, PA132, PA138, PA154, PA160, PA259, PA260, PA324, PA360, PA404, PA412, PA424, PA454, PA460, PA524, PA560, PA660, PA701, PA724, PA760, PA860, PA960, PAX59, PAX60
ROA	PA106, PA107, PA109, PA110, PA130, PA131, PA132, PA133, PA134, PA135, PA160, PA230, PA231, PA232, PA233, PA234, PA235, PA246, PA248, PA252, PA259, PA260, PA330, PA332, PA333, PA334, PA335, PA346, PA348, PA352, PA359, PA360, PA403, PA425, PA430, PA432, PA433, PA435, PA448, PA452, PA460, PA503, PA530, PA533, PA534, PA535, PA548, PA552, PA560, PA630, PA633, PA635, PA648, PA652, PA660, PA711, PA730, PA733, PA734, PA735, PA748, PA752, PA758, PA760, PA811, PA830, PA833, PA834, PA835, PA848, PA852, PA858, PA903, PA911, PA930, PA933, PA934, PA935, PA948, PA952, PA958, PA960, PAX03, PAX11, PAX30, PAX33, PAX34, PAX35, PAX48, PAX52, PAX58, PAX60
ROE	PA107, PA109, PA110, PA135, PA160, PA231, PA233, PA234, PA235, PA248, PA252, PA259, PA260, PA333, PA335, PA346, PA348, PA352, PA359, PA360, PA433, PA435, PA448, PA452, PA460, PA533, PA535, PA548, PA552, PA560, PA633, PA635, PA648, PA652, PA660, PA733, PA735, PA748, PA752, PA758, PA760, PA833, PA835, PA848, PA852, PA860, PA933, PA935, PA948, PA952, PA960, PAX33, PAX35, PAX48, PAX52, PAX60

Regarding the lag, the most PLIs usually exist for Lag = 1, except for the stock price. As the lag increases, the number of PLIs tends to decrease.

Regarding PKIs, they are present for all financial indicators, of which the largest number, 92 PKIs, relates to ROA and the least, 29 PKIs, relates to stock price. The number of PKIs in each data collection interval is shown in Figure 1. The one-year interval has 26 PKIs, the largest number, and the two-year interval has 21 PKIs, the next largest number. As the data collection interval increases, the number of PKIs tends to decrease.

For different financial indicators, PKIs change and are not all the same. Figure 2 shows the top PKI statistics. Among all PKIs, most of them are relating to “total number of nonpatent backward citations of patents for valid invention grants (PAi60)” of 29 PKIs; the second most are relating to “average drawing counts of valid invention grants (PAi35)” of 20 PKIs; the third most are relating to “average drawing counts of valid invention publications” of 19 PKIs.

4.2. Patent Prediction Equation. Though lots of PKIs are found, the positive and/or negative contribution to financial indicators are not yet revealed. We, therefore, construct the patent prediction equations for quantitatively modeling the relationship.

The patent prediction equation is constructed by four predicting modeling periods, namely, period I: 2016Q4 predicting 2017Q4; period II: 2017Q1 predicting 2018Q1; period III: 2017Q2 predicting 2018Q2; period IV: 2017Q3 predicting 2018Q3.

Details of patent prediction equations are shown in Tables 5–7, though some of them are provided with poor goodness of fit.

For predicting the stock price, the patent prediction equation is shown in Table 5, in which PRICE represents the stock price and (−4) represents Lag = 4. The adjusted $R^2 = 0.6568$ shows that though the patent prediction equation is not perfect, it might have some prediction capability. The number of PKIs in Table 3 is 29, here only 7 PKIs left in Table 5 because they are linearly combined in the patent prediction equation, and each of the PKIs must satisfy significance $p^* < 0.05$.

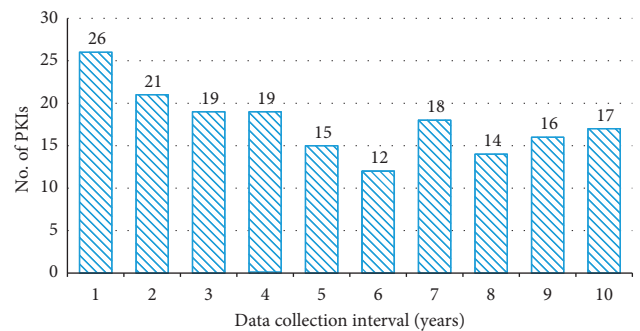


FIGURE 1: The number of PKIs in each data collection interval.

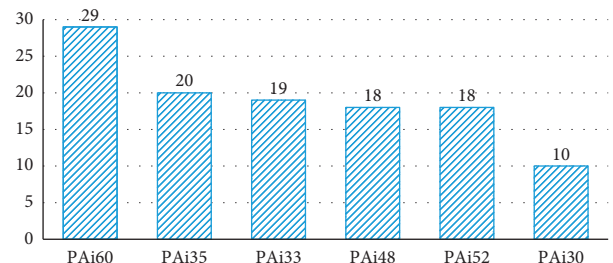


FIGURE 2: Top PKIs for financial indicators.

For predicting ROA, the patent prediction equation with the adjusted $R^2 = 0.3339$ is shown in Table 6. The patent prediction equation might not have enough prediction capability. The number of PKIs in Table 3 is 92, here only 26 PKIs left in Table 6.

For predicting ROE, the patent prediction equation with the adjusted $R^2 = 0.2606$ is shown in Table 7. The prediction capability might be not good. The number of PKIs in Table 3 is 56, here only 9 PKIs left in Table 7.

From the above analysis, it is possible to construct patent prediction equations for Lag = 4 for predicting the stock price, ROA, and ROE. In terms of the goodness of fit, the adjusted R^2 , that of the stock price patent prediction equation at 0.6568 is the best; the worst is of the ROE patent prediction equation at 0.2606. The most PKIs included, numbering 26, are in the ROA patent prediction equation;

TABLE 5: Patent prediction equation for the stock price.

Dependent var.	PRICE				
Independent var.	Coefficient	Std. error	t-statistic	p	
C	-5.8913	0.4060	-14.5118	0.0000***	
PRICE(-4)	0.9106	0.0071	128.5839	0.0000***	
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	0.0000***	
PA960(-4)	-2.8174	1.0804	-2.6078	0.0091**	
PAX60(-4)	3.7089	1.0664	3.4780	0.0005***	
Patent prediction equation	$\begin{aligned} \text{PRICE} = & -5.8913 + 0.9107 * \text{PRICE}(-4) + 0.9663 * \text{PA104}(-4) - \\ & 0.5771 * \text{PA108}(-4) + 0.5419 * \text{PA160}(-4) + 0.4124 \\ & * \text{PA424}(-4) - 0.7103 * \text{PA454}(-4) - 2.8174 * \text{PA960}(-4) + 3.7089 * \text{PAX60}(-4) \end{aligned}$				
Adjusted R ²	0.6568	p (F-statistic)		0.0000***	

p* < 0.05, p** < 0.05, and p*** < 0.05.

TABLE 6: Patent prediction equation for ROA.

Dependent var.	ROA				
Independent var.	Coefficient	Std. error	t-statistic	p	
C	0.8407	0.0430	19.5351	0.0000***	
ROA(-4)	0.5373	0.0087	62.0963	0.0000***	
PA106(-4)	-0.0279	0.0090	-3.0935	0.0020**	
PA107(-4)	-0.0514	0.0126	-4.0975	0.0000***	
PA130(-4)	0.0255	0.0111	2.3012	0.0214*	
PA131(-4)	0.0610	0.0152	4.0166	0.0001***	
PA134(-4)	-0.0526	0.0203	-2.5951	0.0095**	
PA160(-4)	0.0275	0.0093	2.9488	0.0032**	
PA330(-4)	-0.0613	0.0223	-2.7423	0.0061**	
PA359(-4)	0.0448	0.0104	4.2888	0.0000***	
PA425(-4)	-0.0252	0.0122	-2.0702	0.0385*	
PA530(-4)	0.1066	0.0376	2.8354	0.0046**	
PA535(-4)	-0.1164	0.0531	-2.1898	0.0286*	
PA548(-4)	-0.1872	0.0525	-3.5667	0.0004***	
PA660(-4)	0.0183	0.0078	2.3495	0.0188*	
PA734(-4)	0.3742	0.1443	2.5933	0.0095**	
PA735(-4)	0.3275	0.1290	2.5392	0.0111*	
PA811(-4)	0.0575	0.0226	2.5455	0.0109*	
PA852(-4)	-0.0186	0.0084	-2.2196	0.0265*	
PA858(-4)	-2.4780	0.8502	-2.9146	0.0036**	
PA903(-4)	0.0215	0.0069	3.1319	0.0017**	
PA930(-4)	-0.1071	0.0325	-3.2966	0.0010**	
PA933(-4)	1.3999	0.3236	4.3265	0.0000***	
PA934(-4)	-0.3665	0.1448	-2.5303	0.0114*	
PAX33(-4)	-1.3631	0.3249	-4.1956	0.0000***	
PAX35(-4)	-0.2616	0.1159	-2.2565	0.0241*	
PAX48(-4)	0.1768	0.0574	3.0811	0.0021**	
PAX58(-4)	2.6567	0.8520	3.1180	0.0018**	
Patent prediction equation	$\begin{aligned} \text{ROA} = & 0.8407 + 0.5373 * \text{ROA}(-4) - 0.0279 * \text{PA106}(-4) - \\ & 0.0514 * \text{PA107}(-4) + 0.0255 * \text{PA130}(-4) + 0.0610 * \text{PA131}(-4) - \\ & 0.0526 * \text{PA134}(-4) + 0.0275 * \text{PA160}(-4) - 0.0613 * \text{PA330}(-4) + 0.0448 * \text{PA359}(-4) - \\ & 0.0252 * \text{PA425}(-4) + 0.1066 * \text{PA530}(-4) - 0.1164 * \text{PA535}(-4) - \\ & 0.1872 * \text{PA548}(-4) + 0.0183 * \text{PA660}(-4) + 0.3742 * \text{PA734}(-4) + 0.3275 * \text{PA735}(-4) \\ & + 0.0575 * \text{PA811}(-4) - 0.0186 * \text{PA852}(-4) - 2.4780 * \text{PA858}(-4) + 0.0215 * \text{PA903}(-4) - \\ & 0.1071 * \text{PA930}(-4) + 1.3999 * \text{PA933}(-4) - 0.3665 * \text{PA934}(-4) - 1.3631 * \text{PAX33}(-4) - \\ & 0.2616 * \text{PAX35}(-4) + 0.1768 * \text{PAX48}(-4) + 2.6567 * \text{PAX58}(-4) \end{aligned}$				
Adjusted R ²	0.3339	p (F-statistic)		0.0000***	

p* < 0.05, p** < 0.01, and p*** < 0.001.

TABLE 7: Patent prediction equation for ROE.

Dependent var.	ROE			
Independent var.	Coefficient	Std. error	<i>t</i> -statistic	<i>p</i>
C	1.0292	0.0358	28.7224	0.0000***
ROE(-4)	0.4873	0.0090	54.0912	0.0000***
PA110(-4)	-0.0736	0.0244	-3.0222	0.0025**
PA160(-4)	0.0540	0.0147	3.6760	0.0002***
PA233(-4)	-0.0311	0.0156	-1.9930	0.0463*
PA260(-4)	-0.0443	0.0185	-2.3888	0.0169**
PA359(-4)	0.0210	0.0089	2.3668	0.0180*
PA552(-4)	-0.0232	0.0060	-3.8484	0.0001***
PA560(-4)	0.0341	0.0127	2.6923	0.0071**
PA735(-4)	0.5444	0.1445	3.7663	0.0002***
PAX35(-4)	-0.6014	0.1455	-4.1332	0.0000***
Patent prediction equation	$\text{ROE} = 1.0292 + 0.4873 * \text{ROE}(-4) - 0.0736 * \text{PA110}(-4) + 0.0540 * \text{PA160}(-4) - 0.0311 * \text{PA233}(-4) - 0.0443 * \text{PA260}(-4) + 0.0210 * \text{PA359}(-4) - 0.0232 * \text{PA552}(-4) + 0.03406 * \text{PA560}(-4) + 0.5444 * \text{PA735}(-4) - 0.6014 * \text{PAX35}(-4)$			
Adjusted R^2	0.2606	<i>p</i> (<i>F</i> -statistic)		0.0000***

$p^* < 0.05$, $p^{**} < 0.01$, and $p^{***} < 0.001$.

the stock price patent prediction equation has the least, only seven. However, the number of PKIs does not seem to be relevant to the adjusted R^2 .

Since the adjusted R^2 of the aforementioned patent prediction equations is between 0.2606 and 0.6568, it might be inappropriate to use any patent prediction equation for predicting any specific financial indicator. However, the predictive values could be applied for constructing specific investment portfolios consisting of some specific stocks.

4.3. Investment Portfolio Performance. The main objective of patent analysis and prediction is to understand how consideration of patent informatics while investing can be commercially beneficial. In this section, because patent prediction equations for financial indicators are established, stock selection criteria based on these patent prediction equations and the performance of investment portfolios are discussed. Since four quarters, that is, Lag = 4, are applied in patent prediction equations, the stock price return rates of all investment portfolios are compared at annual intervals.

Because the following four predictive modeling periods are used in constructing the patent prediction equations, the investment portfolio performance is also observed over these four predicting modeling periods:

Period I: 2016Q4 to 2017Q4.

Period II: 2017Q1 to 2018Q1.

Period III: 2017Q2 to 2018Q2.

Period IV: 2017Q3 to 2018Q3.

Then, we make an objective comparison between the performance of all the A-shares, which represents the market trend and performance of all the effective samples of stocks over the same periods.

4.3.1. Investment Strategy I. Top 100, top 200, and top 300 stocks to be included in the investment portfolios are selected by the higher predictive values of financial indicators.

Patent prediction equations are used to generate the predictive values of the selected stocks for each financial indicator in the first quarter of each period. These predictive values are then compared with the averages of the rates on annual return on the actual stock prices.

The above comparisons are shown in Table 8. Because of the decline in the overall economic environment, stock price annual return rates of the A-shares from period I to period IV are negative. Especially, due to the impact of the China-US trade conflict, the decline in the periods III and IV is steeper. However, the performance of all effective samples in each period is better on average by 1.14% than the performance of all A-shares together. It means that the average performance of stocks with patents is higher than the stocks without patents. Thus, patent-based stock selection results in better returns.

Among the investment portfolios selected by the higher predictive stock price, in the four periods, PRICE100 has the best performance in two periods (periods III and IV) and PRICE300 has the best performance in two periods (periods I and II). Of these two, PRICE100 is preferable, going by its better average during the four periods. However, PRICE100 has still 4.59% less average than the A-shares' average and 5.73% less average than the effective samples' average. Therefore, selecting the stocks by the higher predictive stock price seems not to be a good stock selection criterion.

As for the selection of stocks for investment portfolios by the higher predictive ROA, among four periods, ROA100 has the best performance in three periods (periods I, II, and III) and ROA200 has the best performance in one period (period IV). Compared to the all of A-shares and all effective samples, each of the ROA100, ROA200, and ROA300 is better. Among these, ROA100 is preferable for its better average over the four periods and has 15.50% higher average than A-shares' average.

Among the investment portfolios selected by the higher predictive ROE, ROE100 has the best performance in three periods (periods I, II, and III) and ROE300 has the best performance in one period (period IV). Compared with the

TABLE 8: Performance comparison of investment strategy I.

Investment portfolio	Actual stock price annual return rate				Average
	Period I 2016Q4→2017Q4	Period II 2017Q1→2018Q1	Period III 2017Q2→2018Q2	Period IV 2017Q3→2018Q3	
A-shares	-19.42%	-22.77%	-27.85%	-36.86%	-26.73%
All effective samples	-18.12%	-21.51%	-26.94%	-35.79%	-25.59%
PRICE100	-43.16%	-32.62%	-18.29%	-31.19%	-31.32%
PRICE200	-40.08%	-33.18%	-24.06%	-33.86%	-32.80%
PRICE300	-35.92%	-30.16%	-25.06%	-34.72%	-31.47%
ROA100	-1.45%	-8.31%	-10.14%	-24.99%	-11.22%
ROA200	-9.37%	-13.52%	-13.59%	-24.55%	-15.26%
ROA300	-11.45%	-12.74%	-18.02%	-28.16%	-17.59%
ROE100	11.21%	-2.25%	-11.89%	-26.77%	-7.43%
ROE200	2.93%	-4.02%	-14.52%	-27.22%	-10.71%
ROE300	-1.46%	-7.84%	-15.76%	-26.55%	-12.90%

PRICE100, PRICE200, and PRICE300 stand for investment portfolios of top 100, top 200, and top 300 stocks selected by the higher predictive stock price. ROA100, ROA200, and ROA300 stand for top stocks selected by the higher predictive ROA. ROE100, ROE200, and ROE300 stand for top stocks selected by the higher predictive ROE.

A-shares, each effective sample of the ROE100, ROE200, and ROE300 is better. Among the effective samples, ROE100 is preferable for the average of four periods, which has 19.30% higher average than the A-shares' average.

Figure 3 shows the stock performance of all effective samples based on higher predictive values and the investment portfolios preferable over the A-shares' average, wherein E. samples stands for all effective samples. For clear comparison, the performance of A-shares is set at zero on the horizontal axis. The positive values show performance better than the market trend (performance of A-shares), and the negative values show performance worse than the market trend. In Figure 3, the portfolios selected by the higher predictive stock price have bad performance, but the portfolios selected by the higher predictive ROA and the higher predictive ROE have better performance, and ROE100 is the best. ROE100 has 19.30% higher average than the A-shares' average and 18.16% higher average than all effective samples' average.

Let us focus on the periods III and IV in Table 8, though the stock price annual return rates are deeply in the negative due to the impact of the China-US trade conflict. By either of the criteria, ROA100 to ROA300 and ROE100 to ROE300 demonstrate higher performance than the A-shares and among the effective samples. Using the higher predictive ROA and the higher predictive ROE is a good investment strategy for choosing good investment portfolios during the China-US trade conflict.

Meanwhile, for the number of stocks in the investment portfolio, the top 100 is usually the best; as the number of stocks increases, the performance of the investment portfolio tends to decrease. Therefore, it seems advisable to limit the number of stocks in the investment portfolio to less than 200.

4.3.2. *Investment Strategy II.* For investment strategy II, the stocks in investment portfolios are selected by the higher predictive growth rates of the financial indicators. The patent prediction equations are applied to generate predictive values of the stock price, ROA, and ROE in the first quarter

of each period. When compared with real values of the stock price, ROA, and ROE in the first quarter of each period, the corresponding predictive growth rates of the stock price, ROA, and ROE are arrived at.

Top 100, top 200, and top 300 stocks selected by the higher predictive growth rates are set as the investment portfolios, and then the averages of actual stock price annual return rates are examined. The comparison is shown in Table 9.

Table 9 shows that, of the investment portfolios selected by the higher predictive stock price growth rate in the four periods, PRICE100R has the best performance in two periods (periods I and IV) and PRICE200R has the best performance in two periods (periods II and III). Compared with the A-shares, each of the PRICE100R, PRICE200R, and PRICE300R is better. Among these, PRICE200R is preferable for the average of four quarters and has 18.17% better average than the A-shares' average.

Among the investment portfolios selected by the higher predictive ROA growth rate, ROA300R has the best performance in three periods (periods II, III, and IV) and ROA100R has the best performance in one period (period I). Compared with the A-shares, all three, ROA100R, ROA200R, and ROA300R, are better, and among the three, ROA300R is preferable for the better average of four quarters and has 2.66% better average than the A-shares' average.

Among the investment portfolios selected by the higher predictive ROE growth rate, ROE200R has the best performance in two periods (periods I and III), ROE100R performed best in period IV, and ROE.300 has the best performance in one period (period II). Compared with the A-shares, each of the three effective samples is better than the A-shares' average. Among these, ROE200R is preferable for the best average of four quarters and has 1.98% better average than the A-shares' average.

With the higher predictive financial indicator growth rate as a stock selection criterion, Figure 4 shows the stock performance of all effective samples and the preferable investment portfolios selected by the higher predictive financial indicator growth rate when compared with the

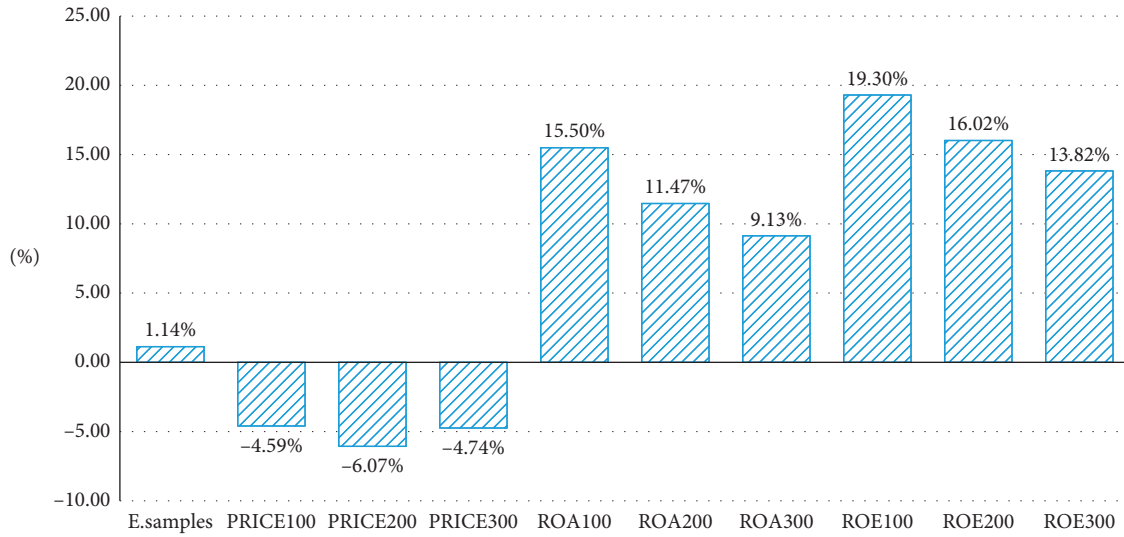


FIGURE 3: Performance comparison of investment strategy I.

TABLE 9: Performance comparison of investment strategy II.

Investment portfolio	Actual stock price return rate				Average
	Period I 2016Q4→2017Q4	Period II 2017Q1→2018Q1	Period III 2017Q2→2018Q2	Period IV 2017Q3→2018Q3	
A-shares	-19.42%	-22.77%	-27.85%	-36.86%	-26.73%
All effective samples	-18.12%	-21.51%	-26.94%	-35.79%	-25.59%
PRICE100R	5.02%	-7.23%	-15.93%	-19.97%	-9.53%
PRICE200R	4.16%	-1.98%	-14.30%	-22.09%	-8.55%
PRICE300R	-0.26%	-6.08%	-15.04%	-24.35%	-11.43%
ROA100R	-10.90%	-21.28%	-30.89%	-37.14%	-25.05%
ROA200R	-13.98%	-22.35%	-27.84%	-36.06%	-25.06%
ROA300R	-13.64%	-20.03%	-27.26%	-35.35%	-24.07%
ROE100R	-16.01%	-23.31%	-27.94%	-36.29%	-25.89%
ROE200R	-12.79%	-21.16%	-27.46%	-37.58%	-24.75%
ROE300R	-15.83%	-21.00%	-28.04%	-36.88%	-25.44%

PRICE100R, PRICE200R, and PRICE300R stand for investment portfolios of top 100, top 200, and top 300 stocks selected by the higher predictive stock price growth rate. ROA100R, ROA200R, and ROA300R stand for top stocks selected by the higher predictive ROA growth rate. ROE100R, ROE200R, and ROE300R stand for top stocks selected by the higher predictive ROE growth rate.

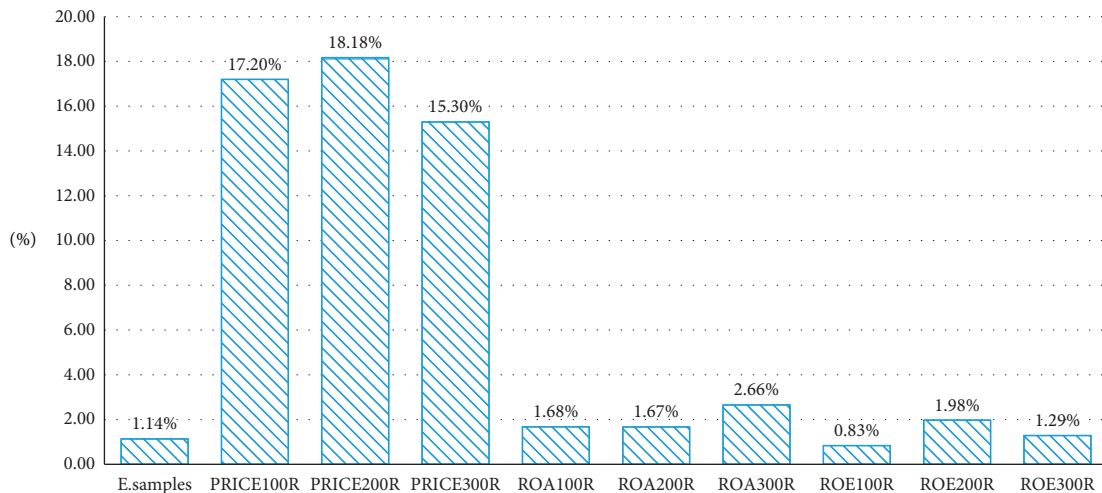


FIGURE 4: Performance comparison of investment strategy II.

A-shares' average. For a clear comparison, the performance of A-shares is placed at zero on the horizontal axis. The positive value means better performance than the market trend, and the negative value means performance worse than the market trend.

In Figure 4, all investment portfolios are seen to have better performance than the market trend. Especially, PRICE100R, PRICE200R, and PRICE300R are outstanding, and PRICE200R is the best. PRICE200R is 18.18% higher than the A-shares' average and 17.04% higher than all effective samples' average.

Let us focus on periods III and IV in Table 9 above, though the stock price annual return rates are seriously negative due to the impact of the China-US trade conflict. Any of PRICE100R to PRICE300R has higher performance than the A-shares and all effective samples. The higher predictive stock price growth rate is also a good investment strategy and useful in identifying good stocks for inclusion in investment portfolios during the China-US trade conflict. However, when compared with Figure 3, we find PRICE200R is 1.12% less than ROE100; therefore, ROE100 is the best investment portfolio in this study.

In Figure 4, going by the number of stocks in the investment portfolio, the top 200 stocks have shown better performance by two of the three criteria; though as the number of stocks increases, the performance of the investment portfolio tends to decrease.

5. Conclusion and Recommendation

Based on patent data, the stock price, ROA, and ROE of China A-shares from 2016Q4 to 2018Q3, we constructed algorithms for finding PLIs, PKIs, patent prediction equations, and strategies for choosing stocks for investment. We arrived at the following conclusions:

- (1) While Chen et al. proposed the use of PLIs for the stock price prediction [3], this study found that PLIs are useful in predicting ROA and ROE.
- (2) The number of PLIs was the largest for Lag = 1, that is, one-quarter lead. As the lag increased, the number of PLIs tended to decrease.
- (3) This study also found that PKIs are useful in predicting for all lead periods (Lag = 1 to 4). The number of PKIs for ROA was the largest at 92 and the least at 29 for the stock price.
- (4) PKIs for different financial indicators were not all the same, but among them, "total backward nonpatent citation counts for valid invention grants (PAi60)" showed up most frequently. It indicated that patents resulting from the fundamental research might have a strong relationship with a company's financial performance.
- (5) This study also constructed patent prediction equations for predicting the stock price, ROA, and ROE. Among them, the ROA patent prediction equation had the most PKIs and the stock price patent prediction equation had preferable goodness

of fit with the adjusted $R^2 = 0.6568$. The number of PKIs in the patent prediction equation did not show relevance to the goodness of fit.

- (6) The stock portfolios selected by the higher predictive ROA, the higher predictive ROE, and the higher predictive stock price growth rate had higher performance than the market trend, wherein ROE100 was the best. Even under the impact of the China-US trade conflict, any of ROA100 to ROA300, ROE100 to ROE300, and PRICE100R to PRICE300R showed better performance than the A-shares' average.
- (7) Chen et al. [3] proposed an investment strategy to select stocks by the higher predictive stock price growth rate. This study found alternative investment strategies to select stocks by the higher predictive ROA and the higher predictive ROE, wherein ROE100 had the best performance among all the stock portfolios.
- (8) The patent-based prediction algorithm proposed in this study was proved useful for identifying good stocks of inclusion in investment portfolios, despite the overall economic environment being unstable and the worsened decline due to the China-US trade conflict. It is believed that the proposed algorithm is also useful for choosing stocks during the COVID-19 pandemic. However, to test that speculation, more data are required.

This study raises some possible issues for further research. For example, how to appropriately combine three preferable stock strategies, the predictive ROA, the predictive ROE, and the predictive stock price growth rate, to get the best performance? And, can AI be used for modeling patent prediction equations and compiling investment portfolios that would perform better?

Data Availability

The data, which were provided by Shenzhen TekGlory Intellectual Property Data Technologies, Ltd., supporting the conclusions of this study are not available for the readers because of confidentiality and commercial reasons.

Conflicts of Interest

None of the authors have any conflicts of interest.

Acknowledgments

The authors would like to thank Shenzhen TekGlory Intellectual Property Data Technologies, Ltd. for processing all the valid patent indicators of China A-shares.

References

- [1] Z. Deng, B. Lev, and F. Narin, "Science and technology as predictors of stock performance," *Financial Analysts Journal*, vol. 55, no. 3, pp. 20–32, 1999.

- [2] P. Thomas, "A relationship between technology indicators and stock market performance," *Scientometrics*, vol. 51, no. 1, pp. 319–333, 2001.
- [3] T.-M. Chen, C.-C. Wei, and H.-C. Che, "Contribution of patent indicators to China stock performance," in *Proceedings of the 7th International Congress on Advanced Applied Informatics (ICAAI 2018)*, pp. 793–798, Yonago, Japan, July 2018.
- [4] B. Branch, "Research and development activity and profitability: a distributed lag analysis," *Journal of Political Economy*, vol. 82, no. 5, pp. 999–1011, 1974.
- [5] Z. Griliches, "Market value, R&D, and patents," *Economics Letters*, vol. 7, no. 2, pp. 183–187, 1981.
- [6] I. Cockburn and Z. Griliches, "Industry effects and appropriability measures in the stock M," *The American Economic Review*, vol. 78, no. 2, p. 419, 1988.
- [7] B. H. Hall, A. Jaffe, and M. Trajtenberg, "Market value and patent citations," *RAND Journal of Economics*, vol. 36, no. 1, pp. 16–38, 2005.
- [8] B. Branch and C. Chichirau, "Mispricing vs risk premia in R&D-intensive firms," *International Review of Financial Analysis*, vol. 19, no. 5, pp. 358–367, 2010.
- [9] M. M. Crossan and M. Apaydin, "A multi-dimensional framework of organizational innovation: a systematic review of the literature," *Journal of Management Studies*, vol. 47, no. 6, pp. 1154–1191, 2010.
- [10] S. Pandit, C. E. Wasley, and T. Zach, "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*, vol. 26, no. 1, pp. 121–144, 2011.
- [11] S. Fabrizi, S. Lippert, P. Norback, and L. Persson, "Venture capital, patenting, and usefulness of innovations," 2011.
- [12] D. Hirshleifer, P.-H. Hsu, and D. Li, "Innovative efficiency and stock returns," *Journal of Financial Economics*, vol. 107, no. 3, pp. 632–654, 2013.
- [13] T. Caner, O. Bruyaka, and J. E. Prescott, "Flow signals: evidence from patent and alliance portfolios in the US biopharmaceutical industry," *Journal of Management Studies*, vol. 55, no. 2, pp. 232–264, 2018.
- [14] G. J. Yu and K. Hong, "Patents and R&D expenditure in explaining stock price movements," *Finance Research Letters*, vol. 19, pp. 197–203, 2016.
- [15] H. B. Mama, "Innovative efficiency and stock returns: should we care about nonlinearity?" *Finance Research Letters*, vol. 24, pp. 81–89, 2018.