Mining Patent Big Data to Forecast Enterprise Performance

Yu-Jing Chiu[(✉)]

Chung Yuan Christian University, 200 Chung Pei Road, Chung Li District,
Taoyuan City 32023, Taiwan, R.O.C.
`yujing@cycu.edu.tw`

Abstract. With intellectual assets are more important, the enterprise pay higher attention. And the patent is an important indicator to measure the R&D results of the company. As China gradually became the leading national competition, this study will be based on Chinese listed companies for the study. Using Granger causality test to identify which patents indicators has a leading enterprise performance. Based on the difficulty of data collection, the study is only with four patent indicators and three enterprise performance indicators for verification of indicators. Patent index are including patent growth rate, patent approval rate, patent number and patent activities and enterprise performance index are net profit, return on assets (ROA) and return on net worth (ROE). Results from this study show that patents leading indicator of different industries have a great deal of difference, which also similar situations and economic leading indicator, therefore, the study recommends more industrial application of different patents leading indicator.

Keywords: Patent indicator · Lead index · Patent-based lead index
Enterprise performance index · Granger causality test

1 Introduction

The competitive advantage that brought to an enterprise by intangible assets would even be more important than tangible assets. Hence, patent rights also have decisive influence in market value of an enterprise. To the competition in an industry, patents have those strengths like enhancing bargaining power of our side, increasing entry barriers, establishing image of our side and striking at the other side. Furthermore, patent rights have a monopolizing character to change the situation of competition in an industry. As the importance of intangible assets on industrial competitive strategies growing with each passing day, every country raises the investment rate of intangible assets. According to [1] estimation, the United States invests about one trillion USD in intangible assets and the investment amount is almost equal to plants and equipment. Even though the rate that intangible assets accounting for net worth of enterprises would be affected by economic fluctuations, empirical data has proved that the long swings of enterprises' market-to-book ratio grow from small to big if we take the difference between market and book as the evaluated value for enterprises' intangible assets. And it can also prove the trend that intangible assets have become the core value

of enterprises [2, 3]. In fact, tangible assets accounted market value for more than 80% and intangible assets accounted for less than 20% in enterprises in the United States as S&P500 showed in 1975. However, the ratio of intangible assets has risen to 81% as Fig. 1 showed in 2009.
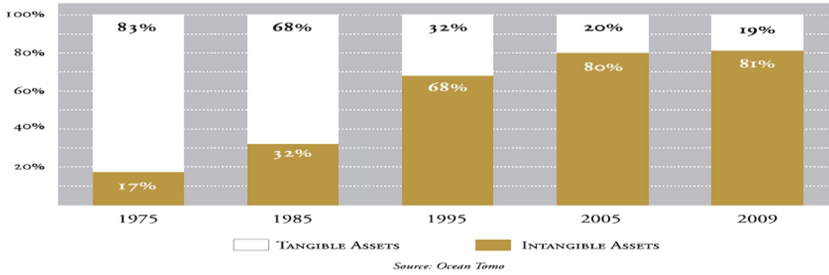
Fig. 1. Market values' component parts of S&P 500 index Source: WWW.OCEANTOMOIN DEXES.COM

To industry policies, patents are a means to improve industrial technique competitiveness; to firms, patents were regarded as monopoly in law so that it plays an important role in market strategy. In recent years of Taiwan, the quality and quantity of patents have greatly increased since government advocates it continually and research institutes, academic units and lots of enterprises fling themselves into it positively. After accumulating a larger quantity of intellectual assets, the needs of asset activation and circulation such as authorization, transaction, transfer, margin trading and securitization must derive from the follow-up of these intellectual assets. It shows that we can get royalties income through patent licensing. Besides, transactions of patent technology has been in a period of vigorous development and become an unmistakable trend. Furthermore, multi-national governments and privately owned enterprises also institute platforms for transactions of patent technology so that they can be a bridge of patent-circulation to fulfill the needs of patent transactions between patent suppliers and demanders.

In the past, the financial instruments of the United States were so highly developed, and they were also applied to IP. For instance, American Ocean Tomo cooperated with American Stock Exchange to announce an IP-based index that was called Ocean Tomo 300[TM] Patent Index (Ocean Tomo as following) so that people can participate in IP transaction market without investing and managing IP directly.

On the 13th September, 2006, Ocean Tomo activated the patent index, and this is the first stock index that took company's IP value as the base. This patent index stands for the combination of 300 companies' shares, and these 300 companies which owned the most patent value were assessed at the book value by Ocean Tomo. American Stock Exchange commented that the index is the first important Broad-Based Index after Dow Jones Industrial Average in 1896, Standard & Poor's 500 in 1957 and Nasdaq Composite Index in 1971. Ocean Tomo patent index uses the software of Patent Ratings System to assess every company's patent portfolios to calculate the patents registered at

United States Patent and Trademark Office, and that was recognized as relative attraction in IP industry. Besides, the process of screen stocks is in a way of regulation and quantification. Hence, the unique and innovation of Ocean Tomo patent index in financial market give people out of institutional investor a chance to use the important type of assets under knowledge economy and to offer a recognized index of assets management under knowledge economy.

Recently, Taiwan makes efforts in researches and rule-making for relative issues of intellectual property rights, but it still comes in significantly below America or other advanced countries. We can find that the index using patents as the base is feasible. Nevertheless, patent is Territoriality Principle. What's more, even though some scholars focus on intellectual assets securitization, most of the topics are about the difficult of intellectual assets securitization, interrelated law system, and the evaluation model of intellectual assets. As you can see from above-mentioned, it's important to set up the accountability and evaluation mechanisms before establishing the market of systematic intellectual assets transactions. Hence, Ocean Tomo 300^{TM} index may be taken as a reference pattern to gradually establish intellectual assets transaction mechanisms for Taiwan's exclusive. However, in these numerous indexes, which indexes have the characteristic of lead index? For example, Baltic Dry Index (BDI) showed a premonition before financial crisis of 2008. Therefore, if these indexes with leading-characteristics can be extracted from numerous indexes, it can improve intellectual assets transaction market mechanisms. And that's the reason why the issue was worth studying. The research purpose of this study is to discuss what patent indicators have characteristics of lead index?

2   Literature Review

At first, the related literature of definition of patent will be mentioned in this part. Next, related literature of patent indicator will be discussed here.

2.1   Definition of Patent

World Intellectual Property Organization (WIPO) mentioned that a patent is an exclusive right granted for an invention, which is a product or a process that provides, in general, a new way of doing something, or offers a new technical solution to a problem. To get a patent, technical information about the invention must be disclosed to the public in a patent application. In another word, inventors create innovations on techniques and are willing to make it public to develop technologies so that laws give them exclusive rights for a period. According to our article 2 of Patent Act, The term "patent" referred to in this Act is classified into the following three categories: (1) Invention patents; (2) Utility model patents; and (3) Design patents. In U.S. patent classification, Utility model patents were called Little Invention patents, and Invention patents include Utility model patents and are collectively called Utility Patents. In these patents, Invention patents are the hardest one since they have to fit in with all the patentability of inventions like industrially applicable, **novelty and** non-obviousness; Utility model patents mean that the shape, structure and installation of an object are

useful and improved; and Design patents indicate the shape, pattern and color of an object appearance are esthetic.

2.2   Patent Indicators

The values of patent rights include two sides: one is technique value and another one is right value. CHI Research Inc. develop a patent indicator system to evaluate the quality of patent right and analyze technique strength of a company through these seven Quantitative *index as following:* Number of patents, Patent growth percent in area, Cites per patent, Current impact index, Technology strength, Technology cycle time, Science linkage, and Science strength [4].

[5] investigated German patent values extensively in 1996 and took Patent Scope, Backward Citations, Forward Citations, Family Size and Oppositions against Patents as evaluation indicators for patent values according to the research results of this field. [6] took patent approved number as an indicator to discuss the relationship between patent approved number, the growing of sales and revenue rate. Also, [7] took patent application number as an indicator to discuss the relationship between patent application, approved number and sales. Besides, [8] mentioned that patent disclosure was taken as an indicator to discuss the influence of patent disclosure level on patent value. [9] indicated that it's not exactly to evaluate patent value by patent count directly since the levels of techniques in patents are not equal. Hence, it's better to evaluate patent value by Citation Frequency.

[10] mentioned that Family Size indicator can evaluate patent value. Furthermore, [11] integrated lots of patent indicators indicated before to develop many patent indicators for management, and these indicators include Patent Quality (PQ), Patent Strength (PS), Technology share, Relative Technology Share, Citation Frequency, International Scope, Technology Scope, Share of granted patents, Co-operation intensity, R&D Emphasis and Patent Activity (PA). [12] thought that previous studies about patent value almost focused on two parts: reference information (Foreign patents references, non-patent references and the total number of cited) and Classification Hierarchy (International Patent Classification, IPC and United States Patent Classification, USPC). Hence, they integrated some experts' opinions in this field and generalized 17 parameters that are relative to patent nature and can be used for quantitative statistics, and these parameters as patent evaluation indicators include Number of applicants, The number of inventors, Total number of items, The total number of independent items, US patents references, Foreign patents references, non-patent references, The total number of citations when filing a lawsuit, IPC classification number, USPC classification number, Numbers of global family size, Number of U.S. patent families, Office action times, The number of respondents/amendments, Patent examination time, Graphic number, and Prosecution time (from patent application to infringement prosecution time).

In the part of patent number analysis, [6] studied the relationship of patents granted to residents, the growing of sales and revenue rate with Patents Granted to Residents indexes. And the result shows that the patents made by a great deal of R&D activities can improve the turnover and profit of a company. [13] viewed patent information as an important indicator. And this scholar also viewed patent as an indicator for the output

of invention activity and economic growth rate of America. Besides, [7] indicated Patents Granted to Residents index to discuss the relationship between patent application, patents granted to residents and sales. In their study, they took pharmaceutical firms as an example, and the result showed that a firm producing more patents by investing in more R&D activities will get better performances relatively.

[14] indicated that previous study results showed the patent citation will improve the performance of the firm (productivity or market value). However, these references studied the influence of science linkage in patent citation on performance of a firm in a system rarely. Therefore, they use Taiwan listed electronics companies' information about the granted patens in USPTO to evaluate the improvement of productivity form patent citation, and the result showed that the number of scientific publications quoted by manufacturers' patents has a negative impact on total factor productivity. Such finding can be explained that the paten with scientific patent references includes more complicated and originality knowledge [15–17] and there is still a gap between the application of the knowledge to produce and market [18]. Hence, it can't react to improve productivity directly and immediately. Otherwise, they also control the variables of patent quality (the speed of being cited of the patent) and other firms. They found that the higher patent quality of a firm will improve more productivity of it. In another word, patent quality of a firm can effect on the performance significantly. Also, scientific patent references will affect productivity of a firm besides other kind of citation. For example, scientific non-patent references and productivity of a firm are significantly positive, it means that scientific non-patent references is beneficial for performance of a firm; to cite the patents in USPTO or the patents in non-USPTO (including patents approved in any other country) makes different influence on performance of a firm (statistically non-significant relation between citing patents in USPTO and productivity; statistically significant negative relation between citing patents in non-USPTO and productivity.

As the above shows that different kinds of patent indicators have different purposes and meanings. In a conclusion, the numbers and the quality of a patent play an important role in the value and the effect of it; a patent indicator also plays an important role in the relative competitiveness of companies, industries and countries.

3   Methodology

The study uses Granger causality test testing the causality between the patent indicators and financial indicators. The Granger causality test is introduced as follows.

3.1   Granger Causality Test

Granger argues that when lack or incomplete of theoretical support, may seek positive complementary theoretical shortcomings and made after actual collection, reflecting information to the data itself, and its predicted effects to demonstrate the causal relationships between variables [19].

$$F(X_t|I_{t-1}) = F(X_t|I_{t-1}-), t = 1, 2 \tag{1}$$

$\{X_t\}$ and $\{Y_t\}$ are two bivariate linear stochastic process generated by stationary time series. $\{I_{t-1}\}$ is the message of the t−1 period collection. $F(X_t|I_{t-1})$ means the given conditional probability situations distribution and collection of composition $\{I_{t-1}\}$ means Y_t length L_y space to have a collection of messages, if (1) does not hold, on behalf of Y between X Granger causality and historical data help to predict the direction of Y (Liu 2012).

In a multivariate linear Gaussian system, if all the variables are observed, the change in the conditional distribution of a variable can be measured via linear regression. The regression formulation of Granger causality states that a variable X is the cause of another variable Y if the past values of X are helpful in predicting the future values of Y. Consider the following two regressions:

$$X_t = \sum_{j=1}^{L} a_j X_{t-j} + \sum_{j=1}^{L} b_j Y_{t-j} + e_t \tag{2}$$

$$Y_t = \sum_{j=1}^{L} c_j X_{t-j} + \sum_{j=1}^{L} d_j Y_{t-j} + f_t \tag{3}$$

where L is the maximal time lag. If Eq. (3) is a significantly better model than Eq. (2), we determine that time series X Granger causes time series Y. To this end, usually the noise variance in two models is compared with each other using a statistical significance test such as the Likelihood Ratio Test. Null hypothesis is as follows:

$$H_0 : b_1 = b_2 = \ldots = b_L = 0$$
$$H_1 : c_1 = c_2 = \ldots = c_L = 0 \quad \text{there is at least one.}$$

The reported *F-statistic*s are the Wald statistics for the joint hypothesis:

$$\beta_1 = \beta_2 = \ldots = \beta_l = 0$$

for each equation. The null hypothesis is that x does not Granger-cause y in the first regression and that y does not Granger-cause x in the second regression.

3.2 Data Delimitation and Limitation

The research scope in this study is in Chinese market since the rise of Chinese market should not be underestimated. Especially in 2010, the amount of invention patent application in Chinese was more than 391,000 and ranked second in the world. At the same time, there were more than 3,000 enterprises applying 12,000 international patent applications through access to the treaty of patent cooperation and the amount of it increased more 56.2% than it in 2009. Besides, in 2010, the amount of effective invention patent applications within the territory of China was about 258,000 and the increment reached to 43%; the percentage of plant invention patent rose from 70.1% of 2006 to 81.3% of 2010. Furthermore, the advantage of owning patents are being turned to the advantage of economy transition in some developed economy regions of China,

and lots of vantage enterprises with core technique patents and international competitiveness spring up, too. Now, both of GDP and the amount of PCT applications in China have ranked second in the world. Hence, intellectual asset rights can improve economy transition. For this reason, Chinese market was taken as the research scope in this study so that to understand the situation of China. Therefore, this study takes about 4138 listed companies in Shanghai Stock Exchange, Shenzhen Stock Exchange and Hong Kong Exchanges and Clearing Limited as the research objects to analyze present situation and the performance of patent indicator through all kinds of patent indicators.

3.3 Data Sampling

These samples of the study is distributed in the different industries, as well as shown in the following Table 1. Industries choose to patent more than 50 filter. Where industrial filters is the number of patents over 50 businesses. In relation to the Shenzhen Stock Exchange, companies' patents in Shanghai Stock Exchange are more widely distributed, while the Shenzhen Stock Exchange of patents is mainly concentrated in the manufacturing sector. These data screening processes are as shown in Fig. 2.

Table 1. Samples of the study

Stock exchange	Industry category	Number of companies
Shanghai Stock Exchange	All industries	261
	Petrifaction, chemistry & plastics	20
	Machinery, equipment & instrument	54
	Pharmaceutical & biotechnology	19
	Metal and nonmetal	22
	Information industry	15
	Manufacturing	106
	Others	25
Shenzhen Stock Exchange	All industries	257
	Manufacturing	255
	Others	2

3.4 Indicators

3.4.1 Patent Indicators

Because the number of patent indicators and the State Intellectual Property Office of the People's Republic of China's patent data can only be part of the patent search methods to do the search. After consider the time of data collection, the degree of difficulty and the data that can be collected, most of the indicators cannot be collected. In finally, the four patent indicators were selected as follows.

a. Patent Growth Percent in Area (GRA)

GRA refers to a business that year compared to percentage increases or decreases in the number of patents (CHI Research Company, 1995).

b. Share of Granted Patents (SGP)

SGP means approved by the company in its field is divided by the number of patents patent activities (Ernst, 1995).

c. Patent Activity (PA)

PA refers to the company in certain technical field of patent applications (Ernst, 1995).

d. Number of Patents (NP)

NP means a company within a certain period in the number of granted patents.

Stock Exchange	Hong Kong Stock Exchange	**Shanghai Stock Exchange**	**Shenzhen Stock Exchange**
Number of Companies	delete	1044	1540

Screening of conditions: More than 50 patents		
Stock Exchange	Shanghai Stock Exchane	Shenzhen Stock Exchange
Number of Companies	1044	1540

Screening of conditions: More than 50 patents and More than five years of financial information		
Stock Exchange	Shanghai Stock Exchange	Shenzhen Stock Exchange
Number of Companies	261	255

Fig. 2. The process of data screening

3.4.2 Indicators of Enterprise Value

In assessing the business value, most people use return on shareholders' equity (ROE), return on assets (ROA) and Net Profit from preparing financial statements, the present use these indicators, identifying patent indicators is influential.

- a. ROE = Net Income/Shareholder Equity
- b. ROA = Net Income/Total Asset
- c. Net profit = Sales revenue − Total costs.

4 Results and Conclusions

According to the result of Granger causality test, we can know different industry has to refer to the implied information of different Leading Index. For example, petrochemical, plastic and plastics industries can use Patents Growth Rate to predict that whether net profit, return on assets ratio and return on net worth will rise up in the future. Furthermore, the sample size of manufacturing turns into 361 companies in total after the combination of two stock exchanges. Such combination shows that it has the highest reference value to manufacturing and the patent-based lead indexes are Number of Patents and Comprehensive Index. In addition, there is no suitable leading index for medical and bio-industry under the situation of the chosen four leading indexes. Hence, it can be inferred the chosen four leading indexes are not appropriate for both medical and bio-industry.

Companies would like to improve the value of the company through the information from patent indexes; investors also want to get some available information from them. However, there are too many numbers of patent indexes to choose. Consequently, the patent-based lead index that can predict the future, rise the value of a firm and are influential will be chosen from this study; moreover, Panel Data Model will be used to deal with cross-sectional data and time series analysis at the same time for the company from different kind of industry to analyze and compare with the result. Even though there were only 4 patent indexes chosen in this research, some of the 4 indexes are not suitable to predict the value of a firm during 2008 to 2012 as the result showed.

As Table 2 shows, the four chosen patent-based lead indexes in this research cannot offer a reference to all industries since the products and the technology cycle time of the companies in all industries are different, and it leads to the time of every company to rise value are different from each other. Also, it's worth taking it as a reference that to choose the patent index according to industrial characteristics. For instance, petrochemical, plastic and plastics industry are more appropriate to take Patents Growth Rate as the patent-based lead indexes of the industry according to the result of this study. Hence, these companies need to pay attention to whether the numbers of patents rise, besides, the numbers of growth will also affect the improvement of the firm's value in the future. On the other hand, mechanical, equipment, instrumental industry and information technology industry need to take more patent-based lead indexes as a reference. In addition, it was found that manufacturing can take Number of Patents as a

reference for patent-based lead indexes after combing Shanghai Stock Exchange with Shenzhen Stock Exchange. Consequentially, all of the results prove that every patent index will not suit to all companies.

Table 2. Patent leading index in every industry

Stock Exchange	Industrial Type	The Value of an Enterprise								
		Net Profit			Return on Assets Ratio			Return on Net Worth		
		Patent Leading Index	Leading Period (Year)	P-Value	Patent Leading Index	Leading Period (Year)	P-Value	Patent Leading Index	Leading Period (Year)	P-Value
	All Companies									
Shanghai	Petrochemical, Plastic and Plastics Industries	Patents Growth Rate	4	0.0364**	Patents Growth Rate	4	0.0334**	Patents Growth Rate	4	0.0252**
		Comprehensive Index	1	0.0000***						
	Mechanical, Equipment, Instrumental Industry	Share of Granted Patent	3	0.0482**				Share of Granted Patent	1	0.0001***
		Patent Activity	3	0.0000***				Patent Activity	1	0.0000***
		Number of Patents	3	0.0013***						
		Comprehensive Index	1	0.0000***						
	Medical, Bioindustry									
	Metal, Non-metal	Comprehensive Index	1	0.0712*						
	Information Technology Industry	Patents Growth Rate	4	0.0006***				Patent Activity	2	0.0995*
		Number of Patents	4	0.0231**						
	Manufacturing	Comprehensive Index	3	0.0000***						
Shenzhen	Manufacturing									
	All Companies									
After Combination	Combined Manufacturing	Number of Patents	4	0.0000***	Number of Patents	4	0.0345**			
		Comprehensive Index	1	0.0000***	Comprehensive Index	1	0.0000***	Comprehensive Index	1	0.0132**

Remarks: 1.*** Significant at the 1% level ; ** Significant at the 5% level ; * Significant at the 10% level. 2. Comprehensive Index were component of Number of Patents, Patent Activity, Patents Growth Rate and Share of Granted Patent.

This study offers not only a reference to develop strategies for a company through patent-based lead indexes but also the useful information to reduce risks for stakeholders of the company; besides, it also offers useful patent indexes to decrease in weakness and barriers to entry for those companies which have not entered the industry to compete. Furthermore, Share of Granted Patent cannot be a reference to most firms, and heterogeneity exists in different industry according to the result of fixed effect model. That means, the technique or the product life cycle of each industry may be

different from each other so that different industry has to use different patent index. Hence, that's the reason some patent indexes have leading effect to some industries in the result of the study. In addition, it was found that the relation of Patent Activity to net profit is the most significant. For further discussing, the Patent Activity behind 2 periods and 3 periods are the best variable to explain the changes of net profit, but other patent indexes are not significant on it. Such finding also assists the information about patent-based lead indexes and offers more reference value for firms.

Although this research only takes 4 patent indexes as explanatory variables, we can still find that patent indexes need to be sieved to fit in the company if we only choose 3 variables from the indexes of company value. It's needed to sieve out more appropriate patent-based lead indexes from lots of patent indexes for each company and each industry in further researches. Besides, the choice of company value can stand on other indictors in further researches.

References

1. Nakamura, L.: Investing in intangible: is a trillion dollars missing from GDP? Bus. Rev. (Q4), 27–36 (2001)
2. Edvinsson, L., Malone, M.S.: Intellectual Capital: Realizing Your Company's True Value by Finding its Hidden Roots. HarperBusiness, New York (1997)
3. Booth, R.: The measurement of intellectual capital. Manag. Acc. **76**(10), 26–28 (1998)
4. Thomas, P.: A relationship between technology indicators and stock market performance. Scientometircs **51**, 319–333 (2001)
5. Harhoff, D., Scherer, F.M., Vopel, K.: Citations, family size, opposition and the value of patent rights. Res. Policy **32**(8), 1343–1363 (2003)
6. Scherer, F.M.: Corporate inventive output, profits and growth. J. Polit. Econ. **73**(3), 290–297 (1965)
7. Comanor, W.S., Scherer, F.M.: Patent statistics as a measure of technical change. J. Polit. Econ. **77**(3), 392–398 (1969)
8. Green, J.R., Scotchmer, S.: On the division of profit in sequential innovation. RAND J. Econ. **26**(1), 20–33 (1995)
9. Schankerman, M., Pakes, A.: Estimates of the value of patent rights in European countries during the post-1950 period. J. Econ. **96**, 1052–1076 (1986)
10. Putnam, J.: The Value of International Patent Rights. Yale University, New Haven (1996)
11. Ernst, H.: Patent information for strategic technology management. World Pat. Inf. **25**(3), 233–242 (2003)
12. Lai, Y.H., Che, H.C.: Modeling patent legal value by extension neural network. Expert Syst. Appl. **36**, 10520–10528 (2009)
13. Schmookler, J.: Invention and Economic Growth. Harvard University Press, Cambridge (1966)
14. Hall, B.H., Jaffe, A., Trajtenberg, M.: Market value and patent citations. RAND J. Econ. **36**(1), 16–38 (2005)
15. Henderson, R., Orsenigo, L., Pisano, G.: The pharmaceutical industry and the revolution in molecular biology: interactions among scientific, institutional and organizational change. In: Mowery, D., Nelson, R. (eds.) The Sources of Industrial Leadership, pp. 267–311. Cambridge University Press, Cambridge (1999)

16. Mowery, D.C., Nelson, R.R., Sampat, B.N., Ziedonis, A.A.: The growth of patenting and licensing by US universities: an assessment of the effects of the Bayh-Dole act of 1980. Res. Policy **30**(1), 99–119 (2001)
17. Sampat, B.N., Mowery, D.C., Ziedonis, A.A.: Changes in university patent quality after the Bayh-Dole act: a re-examination. Int. J. Ind. Organ. **21**(9), 1371–1390 (2003)
18. Cassiman, B., Veugelers, R., Zuniga, P.: In search of performance effects of (in)direct industry science links. Ind. Corp. Change **17**(4), 611–646 (2008)
19. Granger, C.: Investigating causal relations by econometric models and cross-spectral methods. Econometrica **37**(3), 424–438 (1969)