

RESEARCH

Open Access



# Relation of R&D expense to turnover and number of listed companies in all industrial fields

Jun-Hwan Park<sup>1</sup>, Bangrae Lee<sup>1</sup>, Yeong-Ho Moon<sup>1</sup>, GyuSeok Kim<sup>2</sup> and Lee-Nam Kwon<sup>1\*</sup> 

\* Correspondence: [yнкwon@kisti.re.kr](mailto:yнкwon@kisti.re.kr)

<sup>1</sup>Korea Institute of Science and Technology Information (KISTI), 66 Hoegi-ro, Dongdaemun-gu, Seoul 02456, Republic of Korea  
Full list of author information is available at the end of the article

## Abstract

In this research, we studied the relation of research and development (R&D) investment to turnover and number of listed companies by using the financial information of publicly listed enterprises in all industrial fields of the world from 2007 to 2015. First of all, the present condition (as of 2017) of number and distribution of publicly listed enterprises was investigated. Secondly, the industrial areas having top 10 average turnovers and R&D expenses during 9 years (2007 ~ 2015) were analyzed by using their descriptive statistics and CAGR values. Finally, the analyses of correlation and linear regression were performed by using average R&D expense (independent variable) and average turnover or the number of listed enterprises (dependent variables). In other words, two models with different combination of independent and dependent variables (Model A: R&D expense and turnover, Model B: R&D expense and number of listed firms) were developed for the statistical analyses. As a result, it was confirmed that both the turnover and the number of listed companies were influenced by the R&D investment because the coefficients of determination for Model A and Model B were 0.686 and 0.612, respectively (both  $p$ -values  $< 2.2 \times 10^{-16}$ ). From the results of this study, it is expected that the unlisted firms (e.g., start-up companies) can build the basis of their growth and innovation when they invest in R&D higher inducing the increases in (1) turnover and (2) probability of becoming a listed firm. Thus, the financial information of enterprises can be utilized effectively as the quantitative evidence in order to develop the research model and methodology related to their growth and innovation.

**Keywords:** Correlation, Linear regression analysis, R&D investment, Turnover, Publicly listed enterprises, US SIC primary code, ORBIS database

## Introduction

Since industrial trends are being changed very quickly under globalization of world economy in recent years, it is very important to find the opportunities of growth and innovation through research for trends of various industrial areas and government support systems (Patra and Krishna, 2015; Kim et al., 2016; Jeong and Ko, 2016; Yun et al., 2016). In order to search for the promising industrial fields, the financial information of enterprise has achieved a great deal of attention recently (Carp and Mironiuc, 2015; Marilena and Alina, 2015). It is because the worth of business and the tendency of industry are very closely related to the financial data of companies including

turnover, research and development (R&D) investment, R&D intensity, and so on. The R&D intensity is the ratio of R&D investment to turnover (Montresor and Vezzani, 2015). Recently, its level of enterprise or country and effect on the innovation performance is being studied (Goschin, 2014; Karahan, 2015; Doruk, 2015; Savrul and Incekara, 2015). It indicates that the study for relationship between R&D expenditure and turnover has received high research interests. Also, the relation between the business ability and the R&D expenditure is being investigated by many researchers (Bočková and Meluzín, 2016; Ugur et al., 2016a, b). In other words, the effect of R&D investment on the growth and the innovation of companies has been studied and it is very important. The financial information of enterprises in the whole world utilized in this study was obtained from the [ORBIS database of Bureau van Dijk](#). The ORBIS database is being widely used to gain the enterprise data (Martin et al., 2014; Kwon et al., 2016; Faria and Andersen, 2017). Moreover, the forecast of market and technology opportunities/trends in the future based on the financial data of companies as the quantitative evidence is very significant because the enterprises are able to search for their proper industrial areas by utilizing the result of market/technology prediction.

In our previous work, the method for selecting industrial areas suitable to small and medium-sized enterprises (SMEs) in Korea was investigated (Park et al., 2016). The aim of present work is to study the relation of R&D expense to turnover and number of listed enterprises in all industrial fields of the world by using the financial information of enterprises as the quantitative evidence. We expect that the result of this study can be used for the unlisted companies (e.g., start-up firms) to build the basis for their growth and innovation accompanied by increasing in turnover and becoming a listed enterprise from higher investment in R&D. To achieve this goal, number, turnover, and R&D investment of publicly listed companies in the world by the US Standard Industrial Classification (US SIC) primary code from 2007 to 2015 were utilized as the company data. Then, the relationship of R&D investment to turnover and number of listed enterprises in all industries was studied by using the analyses of correlation and linear regression.

### **Literature review**

Socca (2012) studied the role of financial data for decision-making by managers of companies. From this previous work, it was found that the financial information has the following effects on the managers: (a) assisting them in knowing the past and the present situations of their enterprises, (b) providing a quantitative overview of their companies, (c) preparing the activities and decisions for their future, and so on. Babkin et al. (2015) investigated the influence of innovation strategies and R&D investment on the performance of information technology (IT) enterprises. They discovered the correlation between innovation strategies and performance of enterprises. In addition, the result of their study showed that increase in R&D expense makes larger revenue for companies. Pilinkienė (2015) analyzed the effect of R&D expenditure on the competitiveness of companies in the Baltic States. This previous study obtained the correlation matrix between the chosen indicators including competitiveness index, GDP growth rate, and R&D investment in particular Baltic States. Also, they suggested with a policy perspective that the success in economic growth and the increase in country's competitiveness for the Baltic States can be achieved by using R&D expense more efficiently.

Malichová and Ďurišová (2015) estimated the financial performance of companies operating in the sector of IT using the setting of financial indicators such as return on assets (ROA), return on equity (ROE), return on sales (ROS), and so on. By utilizing those financial indicators, they investigated the operating results of enterprises and suggested the identification of important features to accomplish high performance with continuation of finding their changes of financial performance in the sector of IT. Lee and Choi (2015) studied the effect of the financial structure of pharmaceutical businesses on the R&D expense for making a profit source in the next generation or develop the cost-effective medicines for improving the value of company. From their work, the results for the influences of various indicators on the R&D expenditure were as follows: (a) a positive effect of the current ratio, (b) a negative influence of the debt ratio, and (c) little effect of the net sales growth rate. Kretschmer and co-workers (2016) used the financial data such as sales and assets of enterprises from the ORBIS database in order to study the cloud adaptiveness within the sectors of industrial fields by merging them with the technology information offered from the Harte Hanks technology database for 13 countries in Europe from 2000 to 2007. Castellani et al. (2017) investigated the influences of multinationality on the productivity of business by using the ORBIS database. In addition, they found a positive effect of multinationality on the worth of R&D intensity defined in their study as the ratio of R&D expense of company to the number of its workers. Braganza and co-workers (2017) utilized the ORBIS database for enterprises in the world as one of various data sources about businesses, venture capital transactions, academic papers, patents, and so on. They developed an archetype business procedure for the initiatives of big data and the parts needed to handle the resource of big data effectively.

As above-mentioned, many researchers are using the financial performance of enterprises in order to investigate its various effects on the activities of companies such as productivity and R&D. Also, some works mentioned in this section studied the correlation among the financial data of companies (e.g., profitability and R&D). However, in the most previous works, the influence of financial information for specific industries such as pharmacy and IT were studied. On the other hand, in this research, we analyzed the financial data for all industries of the world from a macro perspective and then the correlation and linear regression analyses were performed by developing two models with different combination of independent and dependent variables (1. R&D expense and turnover; 2. R&D investment and number of listed enterprise) to investigate the influence of R&D expense on turnover and number of listed companies.

## **Data and methodology**

### **Data**

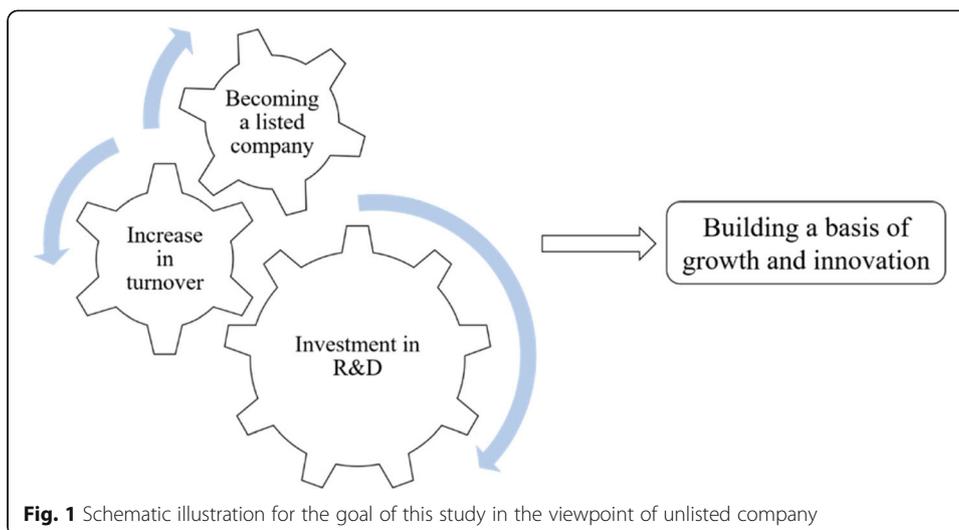
The ORBIS database of Bureau van Dijk updated on 11 May 2017 was utilized in this study. It contains the business data such as financial information, trade description, products, and services in the world (about 200 million companies). In this research, the financial information of publicly listed enterprises was mainly focused for analyzing the relation of R&D investment to turnover and their number in all industrial fields of the world. The US Standard Industrial Classification (US SIC) is the systematic classification assigned by US government in order to identify the industrial activities of

companies. It is comprised of total 11 divisions and they are divided into 83 major groups (2-digit level), 416 industry groups (3-digit level), and 1005 industries (4-digit level). Furthermore, the US SIC primary code of each company is decided by its first line of business (e.g., its business area generating the highest profit). The raw data from the ORBIS based on the list of individual enterprises in the 4-digit level of the US SIC primary code was merged into its 2-digit level to gain the figures related to (a) number of listed businesses (2017), (b) turnover (2007 ~ 2015), and (c) R&D expenditure (2007 ~ 2015). The turnover used in this study means the total operating revenues including net sales, other operating revenues, and stock variations. Also, it does not include value added tax (ORBIS database). The definition of R&D investment is the total amount of expenses on R&D activities (ORBIS database). Moreover, thousand US dollar (USD) was used as a currency unit in this study. The overall procedure to analyze the company-related data is composed of total 4 steps. First of all, the number and the ratio of listed businesses in each US SIC code were obtained in order to observe their distribution in all industrial fields. Secondly, the top 10 average turnovers with their compound annual growth rate (CAGR) figures and descriptive statistics for each US SIC primary code were gained to analyze the industrial trends in the viewpoint of market circumstance. Thirdly, the top 10 average R&D expenditures with their CAGR values and descriptive statistics were calculated to comprehend the trends of industry about research activities of listed enterprises. Finally, investigation into relation of R&D expense to turnover and number of listed companies in all industries was performed through correlation and linear regression analyses. In this step, among 83 major groups (2-digit level), total 6 codes were excluded in the viewpoint of independent variable (R&D expenditure from 2007 to 2015) by the following two reasons: (1) No R&D investment (43 - United States postal service, 81 - Legal services, 88 - Private households, 91 - Executive, legislative, and general government, except finance, and 99 - Nonclassifiable establishments) and (2) Very small average R&D expense (93 - Public finance, taxation, and monetary policy, 0.52 thousand USD).

### Methodology

The scheme for the goal of this work to study the relation of higher R&D investment to increase in turnover and becoming a listed enterprise in the viewpoint of unlisted company (e.g., start-up enterprise) is illustrated in Fig. 1. If there is a positive linear relation of R&D expense to turnover and number of listed enterprises in all industrial areas, it is expected that the unlisted firms are able to build a basis of their growth and innovation through investment in R&D highly for enhancement of turnover and becoming a listed firm. Thus, the expense in R&D was selected as an independent variable for analyses of correlation and linear regression in this study.

Among all of listed and unlisted businesses in the whole world, the listed enterprises were chosen as the target of information analysis in this research. This is because their financial data with more stable financial position than the unlisted firms can be utilized as the representative for each industrial field. In Table 1, the search strategy to gain the number of publicly listed companies in the world by utilizing the ORBIS database is presented. The search condition and result (number of enterprises) of each step are as in the following: (1) all businesses (204,523,868), (2) active companies (163,191,236), (3)



publicly listed enterprises (65,639), (4) public announcement of financial statement based on the consolidation code (62,592), (5) accounting practice by International Financial Reporting Standards (IFRS) or Local Generally Accepted Accounting Principles (Local GAAP) (62,592), (6) accounting template of industrial companies excluding their branches (59,573), and (7) presence of US SIC primary code for the listed companies (59,430). The number of publicly listed enterprises in the whole world by the US SIC primary code (2-digit level) was obtained to study their distribution in all industries. Also, their turnover and R&D expense based on the US SIC primary code during 9 years (2007 ~ 2015) were calculated and utilized as the financial indicators for analyzing the trends of industry.

Also, correlation and linear regression analyses were performed by using R (an open source programming language, R Core Team, 2017) to study the relationship of R&D expense (one independent variable) to turnovers and number of listed companies (two dependent variables) in all industrial areas of the world. For statistical analysis, natural log values of average R&D expense (thousand USD), average turnover (thousand USD), and number of listed enterprises were used. Above all, we tried to confirm the

**Table 1** Search strategy in order to achieve the number of publicly listed enterprises in the world

No.	Search strategy for each step in the ORBIS	Search result (Number of enterprises)
1	All companies	204,523,868
2	Active companies	163,191,236
3	Publicly listed companies	65,639
4	Consolidation code: C1 (companies with consolidated accounts only), C2/U2 (companies with both types of accounts), U1 (companies with unconsolidated accounts only)	62,592
5	Accounting practice: International Financial Reporting Standards (IFRS), Local Generally Accepted Accounting Principles (Local GAAP)	62,592
6	Accounting template: Industrial companies excluding branches	59,573
7	Publicly listed companies with US SIC primary code	59,430

correlation between (1) R&D investment and turnover and (2) R&D expense and number of listed enterprises. Then, the effects of change in the independent variable on two dependent variables were studied by using the linear regression analysis.

## **Results and discussion**

### **Study for number and distribution of publicly listed enterprises**

As one of dependent variables, the number of listed companies in all industrial fields was investigated. Table 2 presents the number and the ratio of listed companies in the world based on the US SIC primary code (2-digit level). Among 83 major groups, the largest number of listed businesses was observed in the area of business services (4524 enterprises, 7.61%). In the next ranking, the industrial fields of (a) chemicals and allied products (3889 companies, 6.54%), (b) electronic and other electrical equipment and components except for computer equipment (3597 businesses, 6.05%), and (c) non-depository credit institutions (3270 enterprises, 5.50%) had the high ratio of distribution. On the other hand, the listed companies were not observed in two US SIC primary codes (88 - private households and 99 - nonclassifiable establishments). Also, we found very low ratios of the publicly listed enterprises in the fields of (a) museums, art galleries, and botanical and zoological gardens (7 companies), (b) United States postal service (6 enterprises), and executive, legislative, and general government except for finance (4 businesses). From these results, we expect that the distribution based on the US SIC primary code of listed enterprises in the whole world can provide the chance for the unlisted firms to observe the industrial areas appropriate for their business situation.

### **Investigation into industrial areas of top 10 average turnovers and their descriptive statistics from 2007 to 2015**

To study for the industrial trends by utilizing the dependent variable of statistical analyses, we obtained the average turnovers and their descriptive statistics of all US SIC major groups from 2007 to 2015. This is because the turnover of enterprise is a basic indicator to determine its status of finance and business activities (Limanlı, 2015; Ugur et al., 2016a, b). As listed in Table 3, the highest average turnover was observed in the area of transportation equipment (3,414,020,116 thousand USD). The industrial field of oil and gas extraction was in the second place for the average turnover (2,816,148,067 thousand USD). However, the CAGR figures of US SIC primary codes showing the top 2 average turnovers (1.88% and 0.52%) were not large compared to those of other areas (2.07% ~ 7.44%) among the top 10 list in Table 3 except for the fields of petroleum refining and related industries (- 1.82%). On the other hand, the industrial areas of (a) food and kindred products and (b) business services having the ninth and the tenth places for the average turnovers had relatively high CAGR (5.69% and 7.44%, respectively). Also, the difference between average turnover and its median was smaller than 7.8% in all top 10 industrial fields of Table 3. In addition, the largest and the smallest values of standard deviation were found in the areas related to oil and gas extraction (602,844,769 thousand USD) and electronic and other electrical equipment and components, except computer equipment (194,836,027 thousand USD), respectively. Moreover, the greatest span between the maximum and the minimum turnovers was

**Table 2** Number and distribution of listed companies based on the US SIC primary code

US SIC primary code - Text description	Number of listed companies	Ratio (%)
01 - Agricultural production-crops	521	0.88
02 - Agricultural production-livestock and animal specialities	234	0.39
07 - Agricultural services	296	0.50
08 - Forestry	63	0.11
09 - Fishing, hunting and trapping	78	0.13
10 - Metal mining	2475	4.16
12 - Coal mining	203	0.34
13 - Oil and gas extraction	1336	2.25
14 - Mining and quarrying of nonmetallic minerals, except fuels	475	0.80
15 - Building construction-general contractors and operative builders	840	1.41
16 - Heavy construction other than building construction-contractors	668	1.12
17 - Construction-special trade contractors	352	0.59
20 - Food and kindred products	2267	3.81
21 - Tobacco products	66	0.11
22 - Textile mill products	1114	1.87
23 - Apparel and other finished products made from fabrics and similar materials	536	0.90
24 - Lumber and wood products, except furniture	279	0.47
25 - Furniture and fixtures	208	0.35
26 - Paper and allied products	565	0.95
27 - Printing, publishing and allied industries	518	0.87
28 - Chemicals and allied products	3889	6.54
29 - Petroleum refining and related industries	223	0.38
30 - Rubber and miscellaneous plastics products	766	1.29
31 - Leather and leather products	157	0.26
32 - Stone, clay, glass and concrete products	1077	1.81
33 - Primary metal industries	1226	2.06
34 - Fabricated metal products, except machinery and transportation equipment	1019	1.71
35 - Industrial and commercial machinery and computer equipment	2519	4.24
36 - Electronic and other electrical equipment and components, except computer equipment	3597	6.05
37 - Transportation equipment	1166	1.96
38 - Measuring, analyzing and controlling instruments; photographic, medical and optical goods; watches and clocks	1222	2.06
39 - Miscellaneous manufacturing industries	557	0.94
40 - Railroad transportation	75	0.13
41 - Local and suburban transit and interurban highway passenger transportation	163	0.27

**Table 2** Number and distribution of listed companies based on the US SIC primary code  
(Continued)

US SIC primary code - Text description	Number of listed companies	Ratio (%)
42 - Motor freight transportation and warehousing	330	0.56
43 - United States postal service	6	0.01
44 - Water transportation	432	0.73
45 - Transportation by air	204	0.34
46 - Pipelines, except natural gas	57	0.10
47 - Transportation services	511	0.86
48 - Communications	1219	2.05
49 - Electric, gas and sanitary services	1483	2.50
50 - Wholesale trade, durable goods	2146	3.61
51 - Wholesale trade, nondurable goods	1031	1.73
52 - Building materials, hardware, garden supply, and mobile home dealers	44	0.07
53 - General merchandise stores	293	0.49
54 - Food stores	302	0.51
55 - Automotive dealers and gasoline service stations	155	0.26
56 - Apparel and accessory stores	167	0.28
57 - Home furniture, furnishings, and equipment stores	206	0.35
58 - Eating and drinking places	389	0.65
59 - Miscellaneous retail	518	0.87
60 - Depository institutions	1170	1.97
61 - Non-depository credit institutions	3270	5.50
62 - Security and commodity brokers, dealers, exchanges and services	856	1.44
63 - Insurance carriers	137	0.23
64 - Insurance agents, brokers, and service	100	0.17
65 - Real estate	2113	3.56
67 - Holding and other investment offices	2761	4.65
70 - Hotels, rooming houses, camps, and other lodging places	670	1.13
72 - Personal services	121	0.20
73 - Business services	4524	7.61
75 - Automotive repair, services, and parking	150	0.25
76 - Miscellaneous repair services	38	0.06
78 - Motion pictures	222	0.37
79 - Amusement and recreation services	371	0.62
80 - Health services	479	0.81
81 - Legal services	14	0.02
82 - Educational services	192	0.32
83 - Social services	42	0.07
84 - Museums, art galleries, and botanical and zoological gardens	7	0.01
86 - Membership organizations	15	0.03

**Table 2** Number and distribution of listed companies based on the US SIC primary code  
(Continued)

US SIC primary code - Text description	Number of listed companies	Ratio (%)
87 - Engineering, accounting, research, management, and related services	1591	2.68
88 - Private households	0	0.00
89 - Services not elsewhere classified	199	0.33
91 - Executive, legislative, and general government, except finance	4	0.01
92 - Justice, public order, and safety	13	0.02
93 - Public finance, taxation, and monetary policy	40	0.07
94 - Administration of human resource programs	8	0.01
95 - Administration of environmental quality and housing programs	18	0.03
96 - Administration of economic programs	54	0.09
97 - National security and international affairs	8	0.01
99 - Nonclassifiable establishments	0	0.00
Total	59,430	100.00

observed in the field of oil and gas extraction (1,488,479,264 thousand USD). From these results, it was found that the turnover of Oil and gas extraction code showed the largest variation among its top 10 industrial areas from 2007 to 2015.

#### Research on industrial fields of top 10 average R&D investments and their descriptive statistics from 2007 to 2015

We computed the average R&D expenses as an independent variable for statistical analysis and their descriptive statistics from 2007 to 2015 for studying the degrees of research activities of publicly listed businesses based on the US SIC primary codes (Table 4). The figures of average R&D investments and their CAGR can be used to study the industrial trends in the viewpoint of research activities/programs of enterprises (Sandu and Ciocanel, 2014; Jin et al., 2016; Elitok et al., 2016; Kang and Cho, 2016). As shown in Table 4, the largest figure of average R&D expenditures was observed in the industrial field related to chemicals and allied products (151,884,109 thousand USD). In addition, among top 10 US SIC primary codes of average R&D investments, the industrial area of business services showed the highest CAGR value (11.19%). However, the amount of R&D expense decreased in the field of communications (CAGR value: -1.05%). Interestingly, total 9 US SIC codes of Table 4 were identical to those of Table 3 although their rankings were different each other. From these results, it is expected that the company having high turnover makes an investment in the R&D expenditure with large value. In addition, the span between average R&D expenditure and its median was smaller than 10.1% in the list of Table 4. Moreover, the greatest and the least standard deviation figures were observed in the codes of Chemicals and allied products (16,847,820 thousand USD) and Communications (1,080,100 thousand USD), respectively. Also, the largest difference between the maximum and the minimum R&D expenses was found in the code of Chemicals and allied products (51,369,100 thousand USD). Therefore, it was confirmed that the largest variation of the R&D investment among codes in Table 4 was observed in the code of Chemicals and allied products.

**Table 3** Descriptive statistics of industrial areas having top 10 average turnovers (2007 ~ 2015)

Ranking	US SIC primary code - Text description	Average turnover (CAGR, %)	Median	Standard deviation	Minimum	Maximum
1	37 - Transportation equipment	3,414,020,116 (1.88)	3425,822,911	339,938,866	2,916,833,631	3,808,847,394
2	13 - Oil and gas extraction	2,816,148,067 (0.52)	2,612,309,352	602,844,769	2,082,624,354	3,571,103,618
3	28 - Chemicals and allied products	2,647,946,798 (4.48)	2,854,291,561	406,439,457	2,051,240,171	3,083,769,070
4	49 - Electric, gas and sanitary services	2,337,933,540 (3.34)	2,317,151,073	343,131,652	1,782,134,510	2,837,146,655
5	29 - Petroleum refining and related industries	2,213,020,301 (-1.82)	2,381,386,062	418,156,592	1,609,563,019	2,639,850,252
6	36 - Electronic and other electrical equipment and components, except computer equipment	2,178,651,683 (2.07)	2,255,989,804	194,836,027	1,898,757,985	2,405,865,824
7	48 - Communications	2,054,883,098 (2.88)	2,169,861,470	195,042,643	1,759,189,019	2,240,045,754
8	35 - Industrial and commercial machinery and computer equipment	1,881,408,977 (4.11)	2,024,095,853	276,782,773	1,505,594,628	2,184,508,210
9	20 - Food and kindred products	1,696,477,390 (5.69)	1,797,364,959	283,201,167	1,204,070,732	2,024,704,606
10	73 - Business services	1,439,930,594 (7.44)	1,375,628,373	311,250,947	1,074,033,567	1,907,609,209

Currency unit: thousand USD

**Table 4** Descriptive statistics of industrial fields with top 10 average R&D investments (2007 ~ 2015)

Ranking	US SIC primary code - Text description	Average R&D investment (CAGR, %)	Median	Standard deviation	Minimum	Maximum
1	28 - Chemicals and allied products	151,884,109 (4.39)	154,019,882	16,847,820	125,179,857	176,548,957
2	36 - Electronic and other electrical equipment and components, except computer equipment	115,438,474 (3.15)	124,361,807	15,109,790	93,528,903	133,242,357
3	37 - Transportation equipment	108,119,248 (2.06)	112,322,852	10,067,455	92,306,185	119,505,243
4	73 - Business services	56,528,208 (11.19)	52,848,383	16,663,724	37,147,573	86,762,354
5	35 - Industrial and commercial machinery and computer equipment	53,895,319 (4.74)	55,258,316	6,953,883	43,909,638	63,622,179
6	38 - Measuring, analyzing and controlling instruments; photographic, medical and optical goods; watches and clocks	25,523,664 (4.05)	26,861,865	2,715,436	21,417,885	29,430,019
7	13 - Oil and gas extraction	16,497,156 (7.23)	18,152,769	3,401,947	11,045,634	19,951,396
8	48 - Communications	12,563,872 (-1.05)	12,504,317	1,080,100	10,715,087	14,265,049
9	29 - Petroleum refining and related industries	10,699,576 (8.03)	10,232,185	1,979,904	7,725,321	14,332,164
10	20 - Food and kindred products	9,180,171 (3.95)	9,860,196	1,257,921	7,231,432	10,202,145

Currency unit: thousand USD

**Analyses of correlation and linear regression between R&D expense and turnover (model a)**

In order to investigate the relation between average R&D investment (independent variable) and average turnover (dependent variable) during 9 years (2007 ~ 2015) in all industrial areas (total 77 major groups, 2-digit level) of the world, the correlation and the linear regression analyses were performed for the Model A as listed in Table 5. The correlation coefficient of the Model A ( $R_A$ ) was 0.828 and it means that there is a strong direct (increasing) linear relationship between two variables. Since the  $p$ -value of the Model A ( $p_A$ ) was very small ( $< 2.2 \times 10^{-16}$ ), it is clear that R&D expense and turnover have the correlation. Also, the linear regression equation between two variables was obtained by using their natural log values (Table 5 and Fig. 2).

Regression equation from the Model A:  $y_{A,n} = 0.583 \times x_n + 11.379$  (n: US SIC code).

In other words,  $x_n$  and  $y_{A,n}$  mean the average R&D expense and the average turnover by the US SIC primary code, respectively. The coefficient of determination for the Model A ( $R_A^2$ ) was 0.686. Thus, we can say that the change in the R&D expense affects the turnover in the level of industrial fields although they may not have causal relationship. For example, the increase in the turnover of specific US SIC code is related to the enhancement in its R&D investment. In addition, the residual standard error of the Model A ( $RSE_A$ ) was 1.229. In this work, the residual figure is the difference based on the natural log between the actual value and the estimated value obtained from the regression equation. The  $RSE_A$  indicates that there is the error of 1.229 when the Model A is used to predict the turnover.

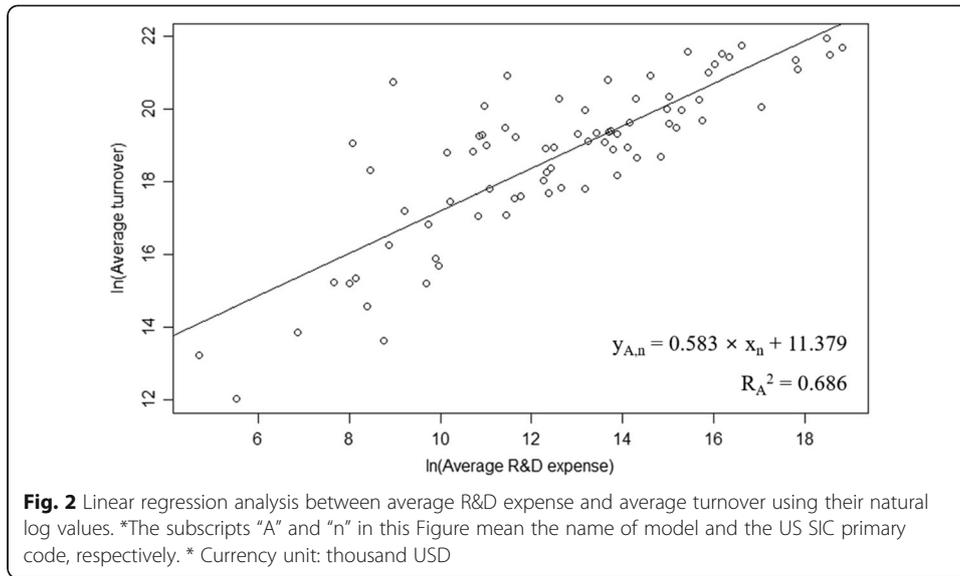
**Analyses of correlation and linear regression between R&D investment and number of listed enterprises (model B)**

The relation of average R&D expenditure (independent variable) to number of listed companies (dependent variable) was analyzed by using their natural log figures (Model B). The scope of analysis such as US SIC primary codes and period was the same as that of Model A. As shown in Table 5, the correlation coefficient for the Model B ( $R_B$ ) was 0.782 and the  $p_B$  was less than  $2.2 \times 10^{-16}$ . From these results, it was found that there is the correlation between the R&D investment and the number of listed companies. In addition, the following regression equation for the Model B could be achieved (Table 5 and Fig. 3).

**Table 5** Statistical values obtained from analyses of correlation and linear regression

Description and statistical values		
Model	A	B
Independent variable ( $x_n$ )	Average R&D expense	Average R&D expense
Dependent variable ( $y_{i,n}$ )	Average Turnover	Number of listed companies
Correlation coefficient ( $R_i$ )	0.828	0.782
Slope ( $a_i$ )	0.583	0.390
Intercept ( $b_i$ )	11.379	0.883
Coefficient of determination ( $R_i^2$ )	0.686	0.612
p-value ( $p_i$ )	$< 2.2 \times 10^{-16}$	$< 2.2 \times 10^{-16}$
Residual standard error ( $RSE_i$ )	1.229	0.968

All statistical figures were computed by using the natural log values of average R&D expenses (thousand USD), average turnovers (thousand USD), and number of listed enterprises  
 The subscripts "i" and "n" in this Table present the model name and the US SIC primary code, respectively

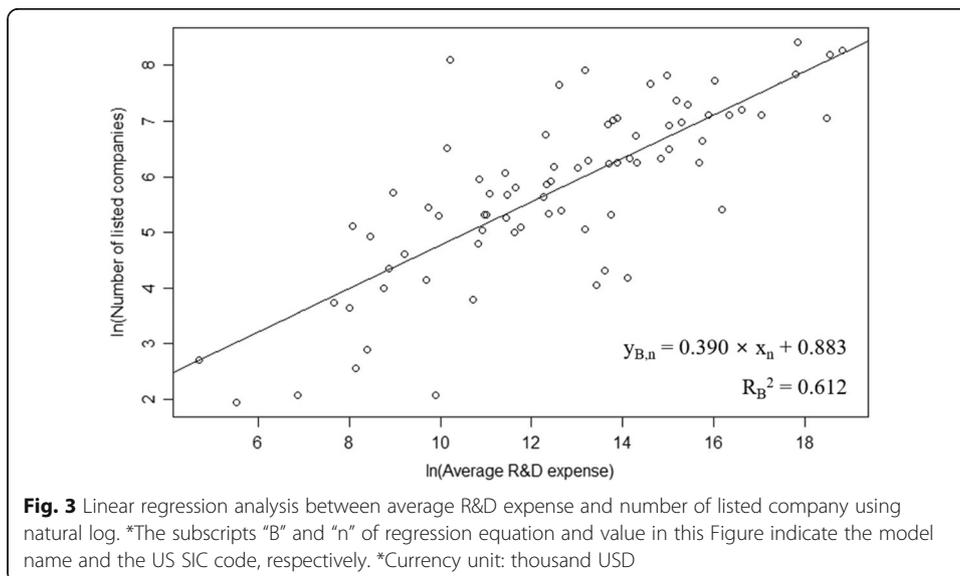


Regression equation from the Model B:  $y_{B,n} = 0.390 \times x_n + 0.883$  (n: US SIC code).

The value of  $R_B^2$  was 0.612 and this indicates that the number of listed enterprises is influenced by the change of R&D expense. In other words, there is the relationship between the higher R&D expenditure of specific industry and its more number of listed companies. Also, the  $RSE_B$  was 0.968 and it is smaller than  $RSE_A$ . This result is caused by the difference in the scale of dependent variables between the Model A (turnover) and the Model B (number of listed companies).

### Conclusions and future work

In conclusion, the relation of R&D expense to the turnover and the number of listed companies in all industrial areas of the world from 2007 to 2015 was studied by using the analyses of correlation and linear regression based on the US SIC primary code.



Through the overall procedure of this work, it was found that both the turnover and the number of listed enterprises were affected by the R&D investment. Therefore, we expect that the unlisted companies (e.g., start-up firms) are able to develop the basis of their growth and innovation as they invest in R&D higher which can induce the increases in (1) turnover and (2) probability of becoming a listed enterprise. However, the main limitations of this work are as in the followings: (1) using only three variables (turnover, R&D expense, and number of listed companies) and (2) performing only simple linear regression analysis. As the future research, it is needed to perform the multiple linear regression analysis using more kinds of financial information (e.g., cost of goods sold, gross profit, net income, and so on) for all listed enterprises in the whole world. In addition, the specific industry areas (3-digit or 4-digit level) included in the major groups of US SIC primary codes (2-digit level) showing the high figures of turnovers and R&D expenditures can be studied. Also, the comparative analysis of financial data among nations in a global sense can be performed. Therefore, it is expected that the above-mentioned future work can be performed effectively by using the results of this research in order to develop more advanced model and methodology based on the financial information of companies as the quantitative evidence.

#### Acknowledgements

This research was supported by Korea Institute of Science and Technology Information (KISTI).

#### Authors' contributions

All authors read and approved the final manuscript.

#### Competing interests

The authors declare that they have no competing interests.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

#### Author details

<sup>1</sup>Korea Institute of Science and Technology Information (KISTI), 66 Hoegi-ro, Dongdaemun-gu, Seoul 02456, Republic of Korea. <sup>2</sup>Korea Institute of Science and Technology Information (KISTI), 245 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea.

Received: 28 September 2017 Accepted: 14 March 2018

Published online: 21 March 2018

#### References

- Babkin, A. V., Lipatnikov, V. S., & Muraveva, S. V. (2015). Assessing the impact of innovation strategies and R&D costs on the performance of IT companies. *Procedia - Social and Behavioral Sciences*, 207, 749–758.
- Bočková, N., & Meluzin, T. (2016). Electronics industry: R&D investments as possible factors of firms competitiveness. *Procedia - Social and Behavioral Sciences*, 220, 51–61.
- Braganza, A., Brooks, L., Nepelski, D., Ali, M., & Moro, R. (2017). Resource management in big data initiatives: Processes and dynamic capabilities. *Journal of Business Research*, 70, 328–337.
- Carp, M., & Mironiuc, M. (2015). The impact of reported financial information on the transfer prices of securities. Comparative empirical study. *Procedia Economics and Finance*, 23, 1426–1433.
- Castellani, D., Montresor, S., Schubert, T., & Vezzani, A. (2017). Multinationality, R&D and productivity: Evidence from the top R&D investors worldwide. *International Business Review*, 26, 405–416.
- Doruk, Ö. T. (2015). Intra industry trade and R&D intensity: An empirical assessment for Turkey. *Procedia - Social and Behavioral Sciences*, 210, 52–57.
- Elitok, Y., Karhan, G., & Türk, E. (2016). The comparative analysis of outgoings for R&D activities in Turkey for 2014. *Procedia - Social and Behavioral Sciences*, 229, 307–315.
- Faria, L. G. D., & Andersen, M. M. (2017). Sectoral patterns versus firm-level heterogeneity - The dynamics of eco-innovation strategies in the automotive sector. *Technological Forecasting & Social Change*, 117, 266–281.
- Goschin, Z. (2014). Research and development intensity in Romania. A regional perspective. *Procedia Economics and Finance*, 15, 64–70.
- Haug, K. C., Kretschmer, T., & Strobel, T. (2016). Cloud adaptiveness within industry sectors – Measurement and observations. *Telecommunications Policy*, 40, 291–306.
- Jeong, H. J., & Ko, Y. (2016). Configuring an alliance portfolio for eco-friendly innovation in the car industry: Hyundai and Toyota. *Journal of Open Innovation: Technology, Market, and Complexity*, 2, 24.

- Jin, X., Lei, G., & Yu, J. (2016). Government governance, executive networks and enterprise R&D expenditure. *China Journal of Accounting Research*, 9, 59–81.
- Kang, J., & Cho, S. (2016). A study on the accountability of the regional R&D program: The case of APCTP. *Journal of Open Innovation: Technology, Market, and Complexity*, 2, 3.
- Karahan, Ö. (2015). Intensity of Business Enterprise R&D Expenditure and high-tech specification in European manufacturing sector. *Procedia - Social and Behavioral Sciences*, 195, 806–813.
- Kim, S.-J., Kim, E.-M., Suh, Y., & Zheng, Z. (2016). The effect of service innovation on R&D activities and government support systems: The moderating role of government support systems in Korea. *Journal of Open Innovation: Technology, Market, and Complexity*, 2, 5.
- Kwon, L.-N., Park, J.-H., Moon, Y.-H., & Lee, B.-R. (2016). An analysis of growth engine industries using the ORBIS DB. *Asian Journal of Innovation and Policy*, 5, 275–292.
- Lee, M., & Choi, M. (2015). The determinants of Research and Development Investment in the Pharmaceutical Industry: Focus on financial structures. *Osong Public Health Res Perspect*, 6, 302–309.
- Limanli, Ö. (2015). Determinants of R&D investment decision in Turkey. *Procedia - Social and Behavioral Sciences*, 195, 759–767.
- Malichová, E., & Ďurišová, M. (2015). Evaluation of financial performance of enterprises in IT sector. *Procedia Economics and Finance*, 34, 238–243.
- Marilena, M., & Alina, T. (2015). The significance of financial and non-financial information in insolvency risk detection. *Procedia Economics and Finance*, 26, 750–756.
- Martin, R., Muûls, M., Preux, L. B. d., & Wagner, U. J. (2014). On the empirical content of carbon leakage criteria in the EU Emissions Trading Scheme. *Ecological Economics*, 105, 78–88.
- Montresor, S., & Vezzani, A. (2015). The production function of top R&D investors: Accounting for size and sector heterogeneity with quantile estimations. *Research Policy*, 44, 381–393.
- ORBIS, Bureau van Dijk (<https://orbis.bvdinfo.com>). Accessed on 16 May 2017.
- Park, J.-H., Lee, B., Moon, Y.-H., & Kwon, L.-N. (2016). Study for selection of industrial areas suitable to small and medium-sized enterprises (SMEs) in Korea. *Journal of Open Innovation: Technology, Market, and Complexity*, 2, 19.
- Patra, S. K., & Krishna, V. V. (2015). Globalization of R&D and open innovation: linkages of foreign R&D centers in India. *Journal of Open Innovation: Technology, Market, and Complexity*, 1, 7.
- Pilinkienė, V. (2015). R&D investment and competitiveness in the Baltic States. *Procedia - Social and Behavioral Sciences*, 213, 154–160.
- R Core Team (2017). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Sandu, S., & Ciocanel, B. (2014). Impact of R&D and innovation on high - tech export. *Procedia Economics and Finance*, 15, 80–90.
- Savrul, M., & Incekara, A. (2015). The effect of R&D intensity on innovation performance: A country level evaluation. *Procedia - Social and Behavioral Sciences*, 210, 388–396.
- Socea, A.-D. (2012). Managerial decision-making and financial accounting information. *Procedia - Social and Behavioral Sciences*, 58, 47–55.
- Ugur, M., Trushin, E., & Solomon, E. (2016). A firm-level dataset for analyzing entry, exit, employment and R&D expenditures in the UK: 1997–2012. *Data in Brief*, 8, 153–157.
- Ugur, M., Trushin, E., Solomon, E., & Guidi, F. (2016). R&D and productivity in OECD firms and industries: A hierarchical meta-regression analysis. *Research Policy*, 45, 2069–2086.
- Yun, J. J., Won, D., & Park, K. (2016). Dynamics from open innovation to evolutionary change. *Journal of Open Innovation: Technology, Market, and Complexity*, 2, 7.

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)

---