CSCanada
Progress in Applied Mathematics
Vol. 4, No. 1, 2012, pp. [22–31]
DOI: 10.3968/j.pam.1925252820120401.Z0896

ISSN 1925-251X [Print]
ISSN 1925-2528 [Online]
www.cscanada.net
www.cscanada.org

A Study on Causes of Errors of Enterprise R&D Statistics & Accounting Data

SHI Ping[a],* ZHU Chengliang[b] and WANG Yurong[c]

[a] Professor, School of Economics and Management, Northwest University, China.
[b] Ph.D. Student, School of Economics and Management, Northwest University, China.
[c] Ph.D. Student in Northwest University, lecturer in Xi'an Finance and Economics University, China.

* Corresponding author.
Address: School of Economics and Management, Northwest University, Xi'an, Shaanxi 710127, China; E-Mail: pshi@nwu.edu.cn

Supported by Key project of soft science of shaanxi province "R&D statistical indicators research" (2011 KRZ01); Shaanxi enterprise management in key discipline construction projects.

Received: March 20, 2012/ Accept: June 15, 2012/ Published: July 31, 2012

Abstract: The level of the R&D statistics quality is directly related to the accuracy of the S&T policy research and making, so to explore the causes of the R&D statistical error is great significance for improving the R&D statistics quality. On the basis of the results of previous studies and seek expert advice, the paper summarizes the possible causes of the R&D statistical error. Through field research to get first-hand data of Shaanxi province, the paper identifies the key causes of the R&D statistical error based on principal component analysis. We found that: the key causes of the R&D statistical error are followed by inconsistent of R&D statistical and accounting port, lack of training to the R&D fill staff, not allowed to grasp the connotation and extension of the R&D indicators and more difficult to obtain data. Therefore, the R&D statistical work should: on the one hand to strengthen among the various departments exchange and communication to clarify the port of R&D statistical and accounting; the other hand should be carried out theoretical studies to solve technical problems in the accounting and define the connotation and extension of the R&D indicators; Moreover, it should be do R&D statistics training to improve professional quality of R&D fill staff.

Key words: Principal component analysis; R&D statistics; Statistical error

Shi, P., Zhu, C., & Wang, Y. (2012). A Study on Causes of Errors of Enterprise R&D Statistics & Accounting Data. *Progress in Applied Mathematics, 4* (1), 22–31. Available from http://www.cscanada.net/index.php/pam/article/view/j.pam.1925252820120401.Z0896 DOI: 10.3968/j.pam.1925252820120401.Z0896

1. INTRODUCTION

The scientific activity includes the research and development activity (R&D), the application activity of R&D as well as the science and technology service activity, among which R&D activity is the most creative part of the science and technology activity, so the statistical index of R&D activities is the comparative index in the world technology activities. Along with the quickening progress of science, the countries all over the world pay more attention to the statistics of R&D activities. The science and technology statistical work of our country began in the 1970s, and gradually introduced the international science and technology statistical standards and norms in the 80s. First national R&D input sampling survey drew back the curtain of the China R&D statistics in 1988. Since the 90s, our country gradually established R&D statistical system relying on the statistical system of science and technology. China's first national R&D resources inventory work was held in 2000, so far the R&D statistical work in China stepped into the fast development period [10]. In 2009 six departments, such as the national bureau of statistics, the ministry of science and technology, the national development and reform commission and so on, jointly carried out "the second national inventory of R&D resources". At present, The statistical data information provided by R&D statistics system of our country can basically meet the needs of people working in all fields. However, China's R&D statistics focused more on the macro level, while the medium industry level and the micro enterprise level are lack of R&D index.

In order to smoothly develop the enterprise R&D input statistical work, scientifically defining of the target of statistics and the statistical range, and setting up of statistical classification standard and a statistical index system, strengthening the construction of statistical system level are needed. The discretion of the statistical quality R&D is related to the accuracy of the research and set of science and technology policy, and it is significant to explore the causes of deviation to improve the quality of R&D statistical work. This paper taking Shaanxi province as a case, analyzes the defects and the insufficiency of the enterprise statistics work on R&D by field investigation and research, and finds out the key factors that influence the enterprise R&D input data deviation in Shaanxi province, and puts forward countermeasures for Shaanxi province to establish the perfect enterprise statistics system, develop standard enterprise R&D statistics work, and design scientific enterprise R&D statistical index, in the hope of improving the quality of enterprise R&D statistics work of Shaanxi province, and laying the solid foundation for the government departments to comprehensively accurately grasp the enterprise R&D situation.

2. RELEVANT LITERATURE REVIEW

Through the existing literature we found that the research on R&D statistics is very rare, and the few articles mainly focus on the following several aspects of the study:

One is about the R&D analysis and international comparison of national statistical method and system on the macroscopic level, such as that of Singapore [3], South Korea [8], the United States [7], Britain [12], India [9] and China [11] and etc. (HE Yingming, 1999; WANG Zhigang, 2000; WANG Hui, 2000; ZHANG Song, 2000; WANG Haifeng, 2001, WEI Heqing, 2004).

Another is about the current R&D statistics survey of our government (LIU Shumei, 2002; LI Gao, XING Qiuyu, 2006) and its contribution to economic development, such as CHEN Zhibin, SHI Jianjun (2003) found that R&D activities and arrangement played an important role in the economic development; through its investigation in China, WANG Mengxin (2011) pointed out that China had following problems in its design of R&D statistical reports and statistical work organization: firstly, he statistical coverage of R&D is not comprehensive; secondly, the statistical index specialized in R&D activities is of less; thirdly, part of the index is lack of operation; fourthly, part of the unit does not pay attention to it; fifthly, part of the R&D reporter cant understand the concept and connotation correctly; sixthly, the statistics system of the department is divided; seventhly, the effect of the forehead training is poor; finally, it puts forward the countermeasures and suggestions of R&D statistical system.

Another is about the comparative analysis of R&D in statistical and accounting areas, such as ZHU Ruiping (2008) analyzed the differences of the costs of research and development in statistical and accounting aspects, and took Xuji company as an example to expound its practice of capitalizing the development costs; SHEN Xuezhen (2005) thought R&D index of the department of statistics and accounting, has many differences on the meaning, content and processing method, affected by the accounting characteristics of the department, although in theoretical definition they were rather close. Statistical R&D index mainly reflects a countrys (region) input of the human, material and financial resources in the process of the government at all levels of scientific research and test development activities in the operation of the national economy, and the contribution on one country (region) economic development and economic growth rate brought by the investment. The accounting R&D, part of the management cost, mainly reflects the influence of current expense on current and the long term profit, which reflects the potential ability of the enterprise development. In addition, the collection and processing of statistical R&D data, in accordance with the requirements and characteristics of accounting statistics, is from the basic statistics department according to statistics caliber and submission procedure to form a country (region) R&D index. While accounting data collection and processing is in accordance with the requirements of the accounting system and procedures. It is from the original data-account data to the report data. Under the principle of account data collection, the original data collection respectively goes into “management expenses”, “intangible assets” and other relevant account, and eventually after generalization and comprehension, it will reflect in “the profit and loss statement” and “balance sheet”.

As can be seen, the current statistic index system focuses more on the macro, and lacks of enterprise micro statistical guidance, and of how to establish and perfect the enterprise internal management system of R&D statistics, therefore, modern science and technology cant be applied to the collect and collated work of the statistical data. In addition, we can see that there is no literature of quantitative research on the cause of R&D statistical deviation.

3. STUDY DESIGN, DATA AND METHOD

A. Study Design

Combining with the previous research and the expert advice, based on the investigation, we summarized the possible reasons for R&D statistics deviation, which concludes 10 ones, seeing table 1 below. In this paper, we use the Likert seven marks table to analyze, from "1" to "7" on behalf of the "very disagree" and "disagree", "a little disagree", "normal", "a little agree", "agree" and "very agree". The questionnaire was of 194 copies, and we received 194 responses, among which 158 copies were effective, which counted for 81.44%.

B. Data Description

Table 1 gives the descriptive statistics of all variables. From the mean of the variables, we can find that the enterprise identities to reason 2 (difficult in data acquisition), reason 4 (inconsistent statistics and accounting area), reason 5 (blur the connotation and denotation of R&D) and reason 9 (lack of personnel training) are higher, scored five points above, which shows that enterprise basically agrees that the above four reasons are the main cause of R&D statistical deviation. At the same time, from the mean of all variables we can also find that the enterprise agreement to reason 1 (unimportant index), cause 7 (the devaluation by leadership) and reason 10 (troublesome) is less esteemed, scored below three points, which shows that enterprise basically disagrees that the above three reasons are the main cause of R&D statistical deviation.

Through the statistical analysis, we got some preliminary conclusions: the disagreement on statistics and accounting R&D, lack of personnel training and the blurring of the connotation and denotation of R&D are the main causes of the deviation of R&D statistical data in Shaanxi province. But whether these preliminary conclusions are reliable or not needs to be verified with further quantitative methods.

C. Research Method

This paper analyzes the cause to R&D statistical deviation by using principal component analysis, trying to find out key reasons. The principal component analysis, introduced in 1901 by Pearson, tries to combine the original variables or index (this article refers to the 10 possible reasons) into a group of new, several irrelevant variables or index. At the same time, choose the several less variables or index according to the actual needs, to reflect the original variables or index information as much as possible. The key operation procedure is as follows:

A. Calculate correlation coefficient matrix. After calculating correlation coefficient matrix, inspect the correlation matrix. If most of the correlation coefficient of the related matrix were less than 0.3, it is not suitable for principal component analysis. In addition, two other statistics provided by SPSS can be used for further observation: Bartlett test of Sphericity and KMO measure.

B. Calculate the eigenvalue and eigenvector of the related matrix, and put the eigenvalue in the order from big to small.

C. Calculate the variance and accumulative total variance contribution rate of principal components. Variance contribution rate of No. k components shows the size of the amount information of the original index, so the more the cumulative variance contribution of the first k main components, the more it contains the original information.

Table 1
Statistics of Descriptive Variables

No.	Detailed description	N	Min	Max	Mean	S.D.
1	unimportant index	158	1.00	2.00	1.42	0.49
2	difficult in data acquisition	158	2.00	6.00	5.06	1.09
3	incomprehensive Statistical coverage	158	1.00	6.00	3.91	1.00
4	inconsistent statistics and accounting area	158	5.00	7.00	6.63	0.51
5	the blurring of the connotation and denotation of R&D	158	3.00	7.00	5.82	0.83
6	the worse operation of Part of index design	158	3.00	6.00	4.84	0.85
7	the devaluation by leadership	158	1.00	6.00	2.92	1.14
8	Lack of R&D index	158	2.00	6.00	4.3	1.22
9	lack of personnel training	158	5.00	7.00	6.22	0.81
10	troublesome	158	1.00	3.00	1.91	0.73

Table 2
Correlation Matrix

				Reason 1	Reason 2	Reason 3	Reason 4	Reason 5	Reason 6	Reason 7	Reason 8
	reason 1	1.00									
	reason 2	0.83	1.00								
	reason 3	-0.73	0.94	1.00							
Correlation coefficient	reason 4	-0.79	0.75	-0.75	1.00						
	reason 5	0.79	-0.79	-0.80	0.75	1.00					
	reason 6	0.73	0.81	0.83	0.78	-0.81	1.00				
	reason 7	-0.81	0.73	0.75	-0.82	-0.95	-0.82	1.00			
	reason 8	-0.72	0.89	-0.72	-0.72	-0.95	0.95	0.91	1.00		
	reason9	0.79	0.73	-0.73	0.80	-0.81	0.19	0.83	0.89	1.00	
	reason10	0.73	0.76	0.88	-0.76	-0.81	-0.73	0.87	0.97	-0.93	1.00

D. Choose the number of the main components. The general method to determine the number of the main component is to make the variance contribution rate of first k main components meet certain requirements; Or to make the first k main components eigenvalue more than or equal to 1.

E. Calculate the main component loading matrix. Sometimes factor rotation is needed to make the index load on the main component have obvious tendency.

4. PRINCIPAL COMPONENT ANALYSIS OF R&D STATISTICAL BIAS CAUSES

This study uses principal component analysis to analyze the main influence factors of R&D input statistical bias of shaanxi enterprise. In this study, the possible reasons which may influence the R&D statistical bias will be regarded as observation index. Since the dimensional problem does not exist in this study, there is no need to standardize the variables. The correlation matrix of index is shown in Table 2.

In correlation matrix, most of the correlation coefficient are greater than 0.7, thus it is suitable to use principal component analysis. It is also be found through the KMO and Bartletts test that the value of KMO is equal to 0.895, and all data pass through the Bartletts test. That is the significant at 1% level reject the null hypothesis that correlation matrix is unit matrix (shown in Table 3), thus in this paper, the observation data is suitable for principal component analysis.

Table 3
The Test of KMO and Bartlett

Sampling of sufficient degree	**Kaiser-Meyer-Olkin measure**	0.895
	Approximate chi-squared	274.373
Bartlett sphericity test	df	45
	Sig.	0.000

From Table 4, this study extracts four principal components whose characteristic value is greater than 1, and the cumulative contribution rate of 80.913% is close to 1. That is the data basically retains the original information, however, the factor reduced from 10 to 4, thus the study extracted 4 principal components. This achieves the purpose of reduction and simplification.

Figure 1 shows the gravel pot of each principal component. From the figure, the extracted 4 principal components mostly included all the original information, and the other 6 only accounted less than 20% of total information. Thus the extracted 4 principal component is reasonable.

Table 5 gives the principal component load matrix before rotation. The table shows that some index does not have obvious tendency in the load of four principal components before rotation, and this is not conductive to the explanation of factor, thus it is necessary to do the factor rotation. For example, the load of principal component 1 is uniform in reason 2, reason 3 and reason 4, and their principal component loading are 0.554, 0.601 and -0.755. It shows that the tendency of principal component 1 is not apparent, thus it is not conductive to the explanation of principal component 1. The load of principal component 2 is uniform in reason 2, reason 5 and reason 9, and their principal component loading are 0.708,0.738

Table 4
Complete Explanatory Variable

P.C.	Initial state			Extracted principal component before rotation			Extracted principal component after rotation		
	C.V	V.C rate%	C.V.C. rate%	C.V	V.C rate%	C.V.C. rate%	C.V	V.C rate%	C.V.C. rate%
1	3.07	30.66	30.66	3.07	30.66	30.66	2.42	24.23	24.23
2	2.08	20.82	51.48	2.08	20.82	51.48	2.06	20.58	44.80
3	1.58	15.82	67.29	1.58	15.82	67.29	1.86	18.56	63.37
4	1.16	13.62	80.91	1.16	13.62	80.91	1.76	17.55	80.91
5	0.52	9.65	90.56						
6	0.21	6.12	96.69						
7	0.19	1.90	98.58						
8	0.10	0.97	99.56						
9	0.02	0.24	99.79						
10	0.02	0.21	100.00						

and 0.843. It shows that the tendency of principal component 2 is not apparent, thus it is not conductive to the explanation of principal component 2. The load of principal component 3 is uniform in reason 4, reason 5 and reason 9, and their principal component loading are 0.665 -0.745 and 0.756. It shows that the tendency of principal component 3 is not apparent, thus it is not conductive to the explanation of principal component 3. The load of principal component 4 is uniform in reason 2, reason 4 and reason 9, and their principal component loading are -0.700, -0.654 and 0.764. It shows that the tendency of principal component 4 is not apparent, thus it is not conductive to the explanation of principal component 4.

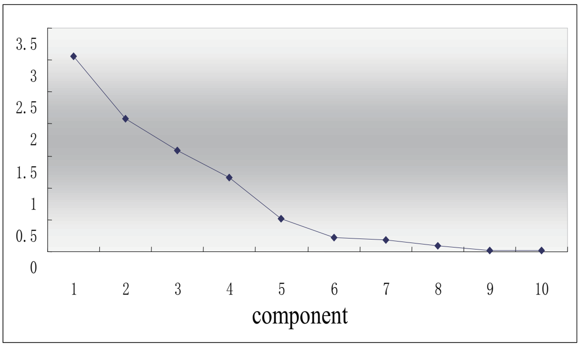

Figure 1
Gravel Plot

The above analysis shows that the information in reason 2 does not only include principal component 1 and principal component 2, but also include principal component 4; the information in reason 4 does not only include principal component 1 and principal component 3, but also include principal component 4; the information in reason 5 does not only include principal component 2 and principal component 3, but also include principal component 4, thus it is necessary to do rotation analysis, and it is contributive to the index explanation of principal component.

Table 5
Principal Component Loading Matrix Before Rotation

	Principal component load			
	1	2	3	4
Reason 1	-0.007	0.145	-0.109	0.379
Reason 2	0.554	0.708	-0.028	-0.700
Reason 3	0.601	0.156	-0.409	-0.436
Reason 4	-0.755	0.048	0.665	-0.654
Reason 5	-0.233	0.738	-0.745	-0.050
Reason 6	0.201	0.274	0.249	-0.232
Reason 7	0.105	-0.278	0.038	0.184
Reason 8	0.153	0.012	0.195	0.322
Reason 9	0.151	0.843	0.756	0.764
Reason 10	0.365	0.048	-0.158	0.125

Principal component 1 reflects the information of reason 4, and the principal component load is 0.863; principal component 2 reflects the information of reason 9, and the principal component load is 0.911; principal component 3 reflects the information of reason 5, and the principal component load is -0.903; principal component 4 reflects the information of reason 2, and the principal component load is 0.898. All these show that four principal components reflect the most information, that is reason 2, reason 4 reason 5, and reason 9 can highly generalize the ten reasons which lead to the bias of R&D statistics, which is also the main reason to the bias in Shaanxi province. The order that the four reasons influence the bias is as follows: reason 4, reason 9, reason 5, reason 2. Among these reasons, the contribution rate of reason 4 is 24.22%, the contribution rate of reason 9 is 20.576%, the contribution rate of reason 5 is 18.564%, and the contribution rate of reason 2 is 17.546%.

To sum up, this study finds, after principal component analysis, that four reasons, which are inconsistent between statistics and accounting, lack in training of reporting staff, uncertainty of connotation and denotation in R&D and the difficulty to obtain the data, are the main reasons which leads to the bias of R&D statistical data.

5. SUMMARY AND CONCLISIONS

The R&D statistical quality affects the accuracy of researching and making science and technology policy, and to explore the causes to deviation is significant to improve R&D statistical quality. Based on the previous research results and the expert advice, we summarized the possible reasons for R&D statistical deviation, and the article used the principal components analysis to find out the key factors of the cause of the R&D statistical deviation. By the principal component analysis, we

Table 6
Principal Component Loading Matrix After Rotation

	Principal component load			
	1	2	3	4
Reason 1	0.142	-0.199	0.269	-0.288
Reason 2	0.170	-0.018	0.235	0.898
Reason 3	0.252	0.066	-0.243	-0.057
Reason 4	0.863	-0.040	-0.112	0.114
Reason 5	-0.131	-0.186	-0.903	-0.029
Reason 6	0.257	0.074	0.347	0.187
Reason 7	-0.043	0.215	0.026	-0.139
Reason 8	0.204	0.374	0.109	-0.081
Reason 9	-0.138	0.911	0.294	0.063
Reason 10	0.217	0.095	-0.058	-0.525

found that inconsistent statistics and accounting area, lack of personnel training, the blurring of the connotation and denotation of R&D, and difficult in data acquisition are four main reasons to R&D statistical deviation. Therefore, in order to solve the problem, those are the four main ways needs to be considered. In particular, (1) Strengthen the exchange and communication between departments, especially that of statistics and accounting. The national bureau of statistics, the ministry of education and the ministry of science and technology shall coordinate with each other. We propose that statistics department was given more support and help on the accounting information disclosure with the existing close and relevant indicators by relevant accounting supervisor and the regulators. (2) Carry out theoretical research to solve the technical problems of accounting. In statistical practice, how to distinguish a sum of money between R&D & non R&D activities, and different types of the R&D activities, and how to calculate the salary of non professional R&D personnel has some difficulties. These problems certainly will affect the quality of statistical data. Therefore, theoretical research should be carried out to find a feasible theory and method to improve the quality of R&D statistics. (3) Get the training work done, and improve personnel's service skill.

REFERENCES

[1] Chen, Zhibin, & Shi, Jianjun (2003). Statistical analysie of regional R&D activities and the research of promoting economic development countermeasure. *Statistics Research*, (2), 16-20.

[2] Guo, Zhigang (2003). *Social statistical analysis method–application of SPSS* (pp. 95-98). Beijing: Renmin University of China Press.

[3] He, Yingming (1999). The characteristics of Singapore science and technology activities from the aspect of R&D statistics survey. *The Outlook of Global Economic, Science and Technology*, (7), 37.

[4] Li, Gao, & Xing, Qiuyu (2006). Present situation and improvement advise of R&D accounting in China. *Statistics and Decision*, (10), 57-59.

[5] Liu, Shumei (2002). The present situation & countermeasure research of government R&D funds in China. *China Science & Technology Forum*, (5), 56-61.

[6] Shen, Xuezhen (2005). Comparison between statistics & accouting of R&D. *Statistics of Education*, (3), 22-24.

[7] Wang, Hui (2000). The method and system of R&D statistics in America. *The Outlook of Global Economic, Science and Technology*, (8), 18-20.

[8] Wang, Zhigang (2000). The method and system of R&D statistics in South Korea. *The Outlook of Global Economic, Science and Technology*, (9), 24-25.

[9] Wang, Haifeng (2001). The index sysytem of R&D statistics in India. *The Outlook of Global Economic, Science and Technology*, (4), 20-21.

[10] Wang, Mengxin (2011). The countermeasures of perfecting Chinas R&D statistics system. *Reformation and Strategy*, (5), 31-33.

[11] Wei, Heqing (2004). The international comparison & development direction of R&D accounting in China. *Statistics and Decision*, (10), 66-67.

[12] Zhang, Song (2000). The method and statistics of R&D in England. *The Outlook of Global Economic, Science and Technology*, (10), 38-39.

[13] Zhu, Ruiping, Zhao, Chenggong, & Zhu, Xiaonan (2008). The division of research and development expenses in statistics and accounting. *Henan Science and Technology*, (10), 32-33.