Applied Mathematics and Nonlinear Sciences

<https://www.sciendo.com>

Analysis of enterprise management technology and innovation based on multilinear regression model

Fangfang Du^{1*}, Adel A. Bahaddad², Radwan Kharabsheh³¹Business School of Henan University, Henan, 475000, China²Department of Information System, Faculty of Computing and Information Technology,
King Abdulaziz University, Jeddah, Saudi Arabia³Applied Science University, Al Eker, Kingdom of Bahrain

Submission Info

Communicated by Juan Luis García Guirao

Received June 16th 2021

Accepted September 24th 2021

Available online December 15th 2021

Abstract

With the continuous innovation of scientific research, technological level and social and economic development, the technological innovation of enterprises in the continuous management is also facing a new breakthrough and upgrade. Especially for enterprises in the era of big data, in order to better respond to the new demands of the development of the new era, enterprise managers should make effective technological innovation from the global economic growth trend on the basis of clarifying their own development advantages. Therefore, on the basis of understanding the multiple linear regression model and its construction conditions, this paper analyses the influencing variables in the actual development according to the current management technology and innovation of enterprises, and obtains the clear results of the model research.

Keywords: multivariate linearity, regression model, enterprise management, technology, innovation, F test, T test.**AMS 2010 codes:** 93A30

1 Model establishment and conditions

1.1 Model

In order to better study the real relationship between enterprise management technology and innovation under the influence of multiple factors, this paper will make a comprehensive exploration based on multiple linear regression model.

Definition 1. It is assumed that enterprise management technology and innovation as the dependent variable Y_1 will be affected by multiple factors. K , and y_1 not being a linear function of XTJ , are the parameters β_0 and β_i ,

*Corresponding author.

Email address: www2534527609@163.com

$I = 1, 2, 3, \dots$ is a linear function of k and the random error term is U_1 . In this case, the formula of multiple linear regression model is:

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_k x_{tk} + u_t, (t = 1, 2, \dots, n)$$

In this case, $E(y_t) = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_k x_{tk}$ represents the population multiple linear regression equation.

Theorem 1. Where k represents the number of explanatory variables, β_0 represents the intercept term, $\beta_1, \beta_2, \dots$. The beta K is the population regression coefficient. $\beta_I, I = 1, 2, 3, \dots, K$ represents the average amount of change in dependent variable Y caused by the change of one unit of independent variable X_{tj} under the condition that other independent variables remain unchanged, which is also called partial regression coefficient.

Proof. In clear samples ($Y_t, X_{t1}, X_{t2}, \dots, X_{tk}$), $t = 1, 2, \dots$. Under the condition of n , the above model can be expressed as:

Proposition 2. Under this condition, Y_t and X_{t1} are clear, but β_i and U_1 are unknown, so the corresponding matrix expression formula is:

$$\begin{cases} y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_k x_{1k} + u_1 \\ y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_k x_{2k} + u_2 \\ y_3 = \beta_0 + \beta_1 x_{31} + \beta_2 x_{32} + \dots + \beta_k x_{3k} + u_3 \\ \dots \\ y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_k x_{tk} + u_t \end{cases}$$

Lemma 3. Under this condition, Y_t and X_{t1} are clear, but β_i and U_1 are unknown, so the corresponding matrix expression formula is:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_T \end{pmatrix}_{(T \times 1)} = \begin{pmatrix} 1x_{11} \dots x_{1j} \dots x_{1k} \\ 1x_{21} \dots x_{2j} \dots x_{2k} \\ 1x_{31} \dots x_{3j} \dots x_{3k} \\ 1x_{T1} \dots x_{Tj} \dots x_{Tk} \end{pmatrix}_{(T \times k)} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \dots \\ \beta_k \end{pmatrix}_{(k \times 1)} + \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \dots \\ u_T \end{pmatrix}_{(T \times 1)}$$

The above formula can be simplified to $Y = X\beta + u$, which is also a simple form of actual regression model analysis.

1.2 Conditions

In order to obtain the best estimator in the model research, the model constructed must meet the following conditions: first, the mean value of the random error term is 0, and the variance σ_2 it has consistent and finite characteristics. Second, the random error terms are independent of each other and have no correlation. Third, the explanatory variable X_{tj} , $j = 1, 2, 3, \dots$. There is no linear relationship between k and k . Fourth, in the explanatory variable X_{tj} , $j = 1, 2, 3, \dots, K$ is the explicit variable, and the error terms are independent of each other. Fifth, the explanatory variable belongs to the non-random variable, and under the condition of $T \rightarrow \infty$, $T^{-1}X'X \rightarrow Q$, where Q refers to a finite number of non-singular matrix. Sixth, the random error term obeys the normal distribution. Seventh, the regression model is properly designed.

1.3 Calculation and analysis

Corollary 4. First, the least square method. According to this principle, the above regression model can be transformed into a sample regression model with the specific formula of $Y = X\hat{\beta} + \hat{u}$, where $\hat{\beta} = (\hat{\beta}_0 \hat{\beta}_1 \dots \hat{\beta}_k)$ refers to the column vector of the estimated value of β , while $\hat{u} = (Y - X\hat{\beta})$ refers to the column vector of residual. Since $\hat{u} = Y - X\hat{\beta}$, $\hat{u}$ is also a linear combination of Y .

Second, minimum sample size. In the research of enterprise management technology and innovation based on multiple linear regression model, the sample size compiled must exceed the number of explanatory variables of the model; in other words, it must conform to this formula.

Third, basic requirements. Generally speaking, only under the condition of $n \geq 30$ or $n \geq 3(k+1)$ can the demand estimated by the expected model be reached.

1.4 Adjustment of multivariate determinability coefficient

In the practice test analysis, the statistical quantity is the determination coefficient, and the model constructed at this time is the multiple regression, so the sample determination coefficient R^2 is called the multiple determination coefficient. Based on the analysis of multiple linear regression model, the decomposition formula of the sum of squares of the total deviation of a single linear regression model is still valid, as will be discussed in the succeeding sections of the article. In this case, the degree of freedom of TSS is $n-1$, where n represents the sample size.

Conjecture 5. *The degree of freedom of ESS is k , k represents the number of independent variables and the degree of freedom of RSS is $n-k-1$. Finally, the following formula can be obtained:*

$$TSS = ESS + RSS$$

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

Combined with the model application analysis, it can be seen that if a new explanatory variable is added to the model, R^2 will continue to rise, which proves that the square of residual will decrease with the increase of explanatory variable. At the same time, since the goodness of fit of the multiple linear regression model needs to be studied, R^2 must be adjusted.

On the basis of clear sample size, increasing explanatory variables will inevitably reduce the degree of freedom. In this case, the adjustment operation should make the sum of squares of residuals and total deviations in their own degrees of freedom, respectively, so as to take advantage of the influence of the number of variables corresponding to the goodness of fit. At this point, the formula of multivariate determination coefficient is as follows:

$$\bar{R}^2 = 1 - \frac{RSS/(n-k-1)}{TSS/(n-1)} = 1 - \frac{n-1}{n-k-1} (1 - R^2)$$

As the independent variable in the model study continues to rise, if $RSS/(n-k-1)$ decreases, then R^2 will increase. At this time, it can be considered that the independent variable has an impact on the dependent variable and needs to be put into the model; otherwise, it needs to be discarded.

Conjecture 5. *At the same time, when the sample size is within a certain range, $\bar{R}^2$ has the following characteristics: on the one hand, in the case of $k \geq 1$, $\bar{R}^2 \leq R^2$; on the other hand, $\bar{R}^2$ may have negative values, assuming that $T=10$, $K=2$, $R^2=0.1$, $\bar{R}^2=-0.157$. At this point, the goodness of fit at a negative value does not have any meaning, so it can be regarded as 0.*

In practical research and analysis, it is found that no matter whether the $\bar{R}^2$ or R^2 continues to increase, the actual fitting effect will be better and better. However, this standard is not the only one to evaluate the merits of the model, so it cannot be analysed and judged only according to the size of or R^2 . Therefore, the judgement can be made by combining Akaike Information Criteria (AIC) and Schwartz Information Criteria (SC), as follows:

$$AIC = \ln\left(\frac{ee}{n}\right) + \frac{2(k+1)}{n}$$

$$SC = \ln\left(\frac{ee}{n}\right) + \frac{k}{n} \ln(n)$$

1.5 Basic assumptions of multiple linear regression model

In order to conveniently estimate the parameters of the model, the following basic assumptions are made for Model 1. Explanatory variable x, x , is a deterministic variable, not a random variable, and $\text{rank}(x)=p+1$.

Note 7. *The random error term has a mean of zero and a variance of sigma 2; namely,*

$$\begin{cases} E(\varepsilon_i) = 0, i = 1, 2, \dots, n \\ \text{cov}(\varepsilon_i, \varepsilon_j) = \begin{cases} \sigma^2, i = j \\ 0, i \neq j, j = 1, 2, \dots, n \end{cases} \end{cases}$$

Open Problem 8. This assumption is called the Gauss Markov condition. $E(\varepsilon)=0$, that is, it is assumed that the observed values have no system. The mean of the error, the random error ε , is zero. Random error term ε is zero, indicating a random error; the difference term is not correlated between different samples. It is independent under the normal assumption. There is a sequence correlation and they have the same precision. The assumed conditions of normal distribution are as follows:

$$\begin{cases} \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n \text{ are independent of each other} \\ \varepsilon_i \sim (0, \sigma^2) \quad i = 1, 2, \dots, n \end{cases}$$

For the matrix form of multiple linear regression, this condition can be expressed as:

$$\varepsilon \sim N(0, \sigma^2 I_n)$$

From the above assumption and the properties of multivariate normal distribution, it can be known that the random vector y still obeys the n -dimensional normal distribution.

Expectation vector of regression model (2-3):

$$\begin{aligned} E(y) &= X\beta \\ \text{var}(y) &= \sigma^2 I_n \\ y &\sim N(X\beta, \sigma^2 I_n) \end{aligned}$$

2 Data source

According to the new development trend of social and economic development and technological innovation in recent years, it is necessary to start with the factors as shown in the figure for research and exploration when studying the technological innovation of enterprise management [1]. In front of the building model analysis in this paper, in order to better grasp the enterprise management technology and the development trend and hotspot in the research of the atmosphere of ‘innovation’ and the frontier, with Chinese knowledge resource pool hownet CNKI (China) as the literature information retrieval platform, bibliometrics analysis was carried out with literature data from Chinese academic periodical CNKI database. The retrieval conditions of the journal were as follows: the retrieval of the subject word ‘innovation atmosphere’ OR ‘innovation atmosphere’, the time span from 1999 to 2020, the publication time from 1 January 1999 to 31 December 2020; all journals were selected by source category and a total of 1873 literatures were obtained through retrieval. A total of 576 articles were published in core journals and CSSCI journals. We filtered the required data according to the analysis of the literature in order to ensure the authenticity and reliability of data sources; this paper only analyses the core journal and journal literature from CSSCI by browsing through the paper title, abstract and key words on the initial literature with artificial screening and washing sample data, to eliminate repetition and invalid sample documents. A total of 452 search results were obtained. Through the data transformation of Citesspacev (5.7R2 version) software, the time span was set as 1999–2020, the time slice was set as 1 year and the time slice threshold was set as 50. The correlation analysis was conducted on the distribution of research institutions, the distribution of research authors, the research hot spots and the research frontier trends. The required variables for this study are shown in Table 1, and the expected efficiency relationship between actual enterprise management and technological innovation is shown in Table 2.

3 Statistical tests

3.1 F test

In order to study various influencing factors of enterprise management technology and innovation from an overall perspective, the original hypothesis tested can be clearly identified as $H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$, where

Table 1 Study variables and explanations

Variable type	Variable name	Variable code	Representations of variable
The explained variable	Investment in technological innovation	R&D	R&D investment/business investment
Explanatory variable	Technological innovation efficiency	D-1	Development investment/R&D investment
	Concentration of ownership	Cr3	The sum of the shareholding ratio of the top three shareholders
	Nature of the largest shareholder	ney	State ownership is marked as 1, otherwise marked as 0
	Proportion of independent directors	idr	Number of independent directors/total number of board members
	Chairman and general manager	lst	The combination of two jobs is marked as 1, otherwise as 0
	Two jobs in one		
	The Board of Supervisors holds shares	bss	Number of shares held by members of the Board of Supervisors/total share capital
	Executive equity incentives	mit	Number of shares held by managers/total share capital
	Debt to asset ratio	debt	Total liabilities/total assets
	Suppliers rely on	spe	Top 5 Supplier Product Amount/Total Product Value
Control variables	Company size	size	The natural logarithm of assets in tens of millions
	Corporate performance	roe	Net profit/total assets
	Industry characteristics	ich	For high-tech enterprises, it is 1, otherwise it is 0

Table 2 The efficiency expectation relationship between enterprise management and technological innovation

Corporate governance structure	Investment in technological innovation	Efficiency of technological innovation
Concentration of ownership	+	+
The nature of the state-owned shareholding	-	-
Proportion of independent directors	+	+
Two jobs in one	+	-
The Board of Supervisors holds shares	+	+
Executive equity incentives	+	+
Debt to asset ratio	-	-
Suppliers rely on	-	+

K represents the number of regression coefficients of the equation, also known as the number of independent variables. Assuming that this formula is true, there is no linear relationship between the explained variables and the explanatory variables in the model. However, if not $H_1 : \beta_j$ all are 0, and the hypothesis is valid, then the test measurement formula is:

$$F = \frac{ESS/k}{RSS/(n-k-1)} F(k, n-k-1) \quad (1)$$

Figure the composition of the relationship between enterprise management and technological innovation

This refers to the F distribution with K, N-K-1 degree of freedom. After the pre-provided significant level A is defined, the exact corresponding degree of freedom can be shown by combining the F distribution.

Assuming that the test level is A, the following two test principles need to be met: First, if $F \leq F_a(k, n-k-1)$, then the original hypothesis should be accepted. Second, if $F > F_a(k, n-k-1)$ is used, then another assumption needs to be accepted.

At this point, the formula for the relationship between F and $\bar{R}^2$ is:

$$F = \frac{R^2}{1-R^2} \cdot \frac{n-k-1}{k}$$

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1+kF} \quad (2)$$

Combined with the formula analysis, it can be seen that F and R2 belong to a direct proportional relationship, and the value of F increases with the increase of R2. In this process, the overall F test will also become more significant, and the corresponding R2 value will continue to rise. At this time, the fitting degree of the regression equation will also become better and better. Therefore, this test can be regarded as a test of goodness of fit.

3.2 t test

In order to eliminate the explanatory variables that have no significant influence on the explained variables in the regression model research, it is necessary to build a simpler multiple linear regression model. Assuming that the explanatory variable X_{tj} is not significant for the explained variable, then the corresponding regression coefficient β_j is 0. In this case, we only need to verify whether the regression coefficient β_j of the explanatory variable X_{tj} is 0 to get the answer. First of all, we put forward the hypothesis of $H_0 : \beta_j = 0, j = 1, 2, \dots, k$; secondly, the alternative hypothesis $H_1 : \beta_j \neq 0$ should be proposed. Finally, the judgement criteria should be defined. The original hypothesis should be accepted under the condition of:

$$|t| \leq t_{\frac{\alpha}{2}}(n-k-1) \quad (3)$$

And the alternative hypothesis should be accepted under the condition of By (1) and (3).

$$|t| > t_{\frac{\alpha}{2}}(n-k-1) \quad (4)$$

4 Result analysis

According to the data collected from the integrated research, the descriptive statistical results of management technology and innovation of an enterprise are shown in Table 3, and the Pearson coefficient test results in the actual multicollinearity analysis are shown in Table 4. Therefore, it can be clearly recognised that there is no multicollinearity problem among the variables.

Table 3 Analysis of descriptive statistical results

Variable	The mean	The standard deviation	The minimum	The maximum
R&D investment (R&D)	0.0727	0.0778	0	0.9839
Technological innovation efficiency (D_1)	0.1223	0.4972	0	15.51
Equity concentration (CR3)	0.5154	0.1299	0.0445	0.9412
State-owned Equity (NEY)	0.0423	0.2016	0	1
Independent Director (IDR)	0.3743	0.0561	0.333	0.667
Two Shares in One (LST)	0.4256	0.4946	0	1
Board of Supervisors shareholding (BSS)	0.5052	0.5002	1	1
Executive ownership [MIT]	0.1940	0.1973	0	0.8968

Table 4 Pearson coefficient test results analysis

	Concentration of ownership (CR3)	State-owned Equity (NEY)	Independent Director (IDR)	Two Shares in One (LST)	Board of Supervisors shareholding (BSS)	Executive ownership [MIT]	Debt load ratio	Supplier Dependency (SPE)
Cr3	1.0000	-0.0614	0.0152	0.0813	-0.2384	0.1403	-0.0807	0.0866
ney	-0.0614	1.0000	-0.1152	-0.1637	-0.0838	-0.1689	0.0397	-0.0365
idr	0.0152	-0.1152	1.0000	-0.0015	-0.0323	0.1417	-0.0553	-0.0853
lst	0.0813	-0.1637	-0.0015	1.0000	-0.0054	0.6012	-0.0276	0.0803
bss	-0.2384	-0.0838	-0.0323	-0.0054	1.0000	0.0655	-0.1309	0.0362
mit	0.1403	-0.1689	0.1417	0.6012	0.0655	1.0000	-0.1150	0.0960
debt	-0.0807	0.0397	-0.0553	-0.0276	-0.1309	-0.1150	1.0000	-0.1799
spe	0.0866	-0.0365	-0.0853	0.0803	0.0362	0.0960	-0.1799	1.0000

At the same time, combined with the following Table 5 analysis shows that governance and technical innovation investment between inside and outside the enterprise impact is mainly manifested in the following: first, the enterprise technological innovation under the 5% level and equity concentration showed positive correlation, prove that a certain degree of equity concentration can further strengthen the enterprise's technological innovation, the enterprise will also increase with big shareholder holding status. In the long run, controlling shareholders can consider investment decisions from the perspective of long-term development, and then accelerate the pace of technological innovation of enterprises. Second, leadership structure is negatively correlated with technological innovation at the 5% level, which proves that the management model combining the two powers limits the technological innovation of enterprises. Third, the board of supervisors stake to under 5% level, which proves that the increase of the proportion with the members of the board of enterprise technology innovation investment will be continuously heightened; the use of this form to increase the board of supervisors

and the correlation of corporate interests help in the comprehensive regulatory management; and the board at the same time is empowered to reduce investment risk faced by technology innovation [3].

From the perspective of technological innovation efficiency, its influence on the internal and external governance of enterprises is mainly reflected in the following points: First, ownership structure. Under the influence of modern enterprise system, the ownership and management rights of enterprises should be adjusted appropriately, and the concentration of ownership should be paid attention to, which can help to reduce the agency cost of enterprise innovation and improve the output efficiency of technological innovation while strengthening supervision. Second, in the development of technology-intensive enterprises, it is very important to put forward a scientific incentive mechanism. In the era of big data, talents are the core resources of enterprise technological innovation, and only a systematic and effective incentive mechanism can fully mobilise employees' enthusiasm for work. Therefore, based on the multiple linear regression model and impact analysis proposed in this paper, in order to fully show the advantages of enterprise technological innovation, it is necessary to build a systematic talent management system, and put forward diversified incentive models and motivation guarantee for their medium and long term development. For example, different material and spiritual incentives can be designed according to talents at different stages of the core layer, backbone layer and potential layer, so as to provide a better development platform for talents at other levels while building a competitive and fair incentive system. Third, the system of the board of supervisors. According to the analysis of agency cost theory and power balance theory, as the basic content of guiding the orderly development of the enterprise board of supervisors system, under the condition of insider control, the decision-making problems of the innovation subject will inevitably affect the benefits obtained from the overall operation and limit the pace of technological innovation and development of the enterprise. On the contrary, the reasonable operation and comprehensive incentive of the Board of Supervisors system can help to further supervise whether the Board of Supervisors can perform tasks in strict accordance with regulations and requirements, and bring about a positive impact on the practice of technological innovation activities [4].

Table 5 Regression analysis

D_i	Symbol anticipates	The regression coefficient	The standard deviation	T value	P values
Cr3	+	0.210***	0.134	2.97	0.003
ney	-	-0.130***	0.086	-2.11	0.010
idr	+	0.447	0.307	0.75	0.410
lst	-	0.033**	0.042	1.77	0.044
bss	+	0.026***	0.035	2.14	0.009
mit	+	0.074	0.109	0.68	0.499
debt	-	-0.111***	0.119	-2.78	0.005
spe	+	-0.036	0.090	-0.4	0.687
size	+	0.001	0.027	0.03	0.974
roe	-	-0.013***	0.022	-4.00	0.000***
ich	-	0.002	0.145	0.11	0.916

Note: *** means significant at 1% level, ** means significant at 5% level, * means significant at 10% level

5 Conclusion

To sum up, therefore, under the background of the era of big data, in order to better cope with the changing market environment, improve the efficiency of enterprise's internal and external management levels and technical innovation, the ownership structure must be optimally constituted, a board of supervisors system featuring talent development with multi-angle in-depth thinking must be ensured and with these must be combined previ-

ous development experience on management optimisation. At the same time, we should start from the current internal and external management environment, and strengthen the supervision and restraint of various departments, so as to reduce the impact of internal and external pressure on technological innovation on the basis of realising the maximum value of the enterprise. In addition, from the perspective of technological innovation, from the enterprise's own development trend and research on scientific research innovation results, the use of multiple linear regression model in-depth discussion of the future development of enterprise management and the main direction of technological innovation, and then combined with practice cases for verification and analysis, finally get clear research results. This can not only provide the basis for enterprise management optimisation, but also help leaders and managers to make clear the importance of technological innovation.

References

- [1] MaaoUa, Ne M, Zouggar S, Krajai G, et al. Modelling industry energy demand using multiple linear regression analysis based on consumed quantity of goods[J]. *Energy*, 2021, 225(4):120270.
- [2] Wen H, JBi, Guo D. Calculation of the thermal conductivities of fine-textured soils based on multiple linear regression and artificial neural networks[J]. *European Journal of Soil Science*, 2020, 71(4).
- [3] Odhiambo O S, Wanyonyi S W, Marangu D M, et al. Analysis of Household Electricity Consumption in Nonresident Rent Halls Using Linear Regression Analysis Model[J]. *British Journal of Biomedical Science*, 2019, 2(3):45779.
- [4] Ghosh S, Chakraborty S. Seismic fragility analysis of structures based on Bayesian linear regression demand models[J]. *Probabilistic Engineering Mechanics*, 2020:103081.
- [5] Li W, Hu Y, Feng J, et al. Research on Armature Structure Optimization of Rail Gun Based on Multiple Linear Regression[J]. *IEEE Transactions on Plasma Science*, 2019, 47(11):5042-5048.
- [6] Muhlroth C, Grottke M. Artificial Intelligence in Innovation: How to Spot Emerging Trends and Technologies[J]. *IEEE Transactions on Engineering Management*, 2020, PP(99):1-18.

This page is intentionally left blank