The Influencing Factors of Industrial Technological Innovation Ability--Using Bayesian Model Averaging

Mengni Zhang
Shanghai University
(School of Economics)
Shanghai,China
melodyyimeng@163.com

***Abstract*—Capacity for independent innovation is an important factor to measure the economic strength of a country or a region. Taking industrial enterprises as the research object, this paper analyzes the influencing factors of technological innovative ability. Previous studies only focus on a single field and lack of considering the model uncertainty, leading to an unfaithful conclusion. By abandoning the limitations of single model and synthesizing the different theories about the technological innovation ability at home and abroad, this paper treats the influencing factors as random variables themselves and uses the Bayesian model averaging to measure the uncertainty of model selection. This study found that there are six variables, R&D investment, R&D personnel input, value of import and export, export trade dependence, GDP growth and government S&T input in the 15 possible influence factor, had great influence on the technology innovation ability of industrial enterprises.**

Keywords—Bayesian Model Averaging (BMA); Technology innovation ability; Model uncertainty; Industrial enterprises; Learning-by-exporting hypothesis

I. INTRODUCTION

Technical innovation embodies comprehensive national strength and leads enterprises to faster and sustainable development. Industry plays an important role of national economy. It is the main source of China's state revenue and it is the guarantee of the national economic, political independence, and defense modernization. It is meaningful to study the influence factors of industrial company's' technology innovation.

Previous studies of the influencing factors of technological innovative ability only focus on a single field. As a result, different research fields have put forward different theories. For instance, "endogenous growth theory" focuses on the relationship between R&D input and output. They usually study the relationship between innovation capacity and R&D input including R&D personnel input and R&D investment, e.g. Cohen(1989), Yeaple (2004) and Xia Zheng (2014). The "regional economic theory" studies the technological externalities arising from industrial agglomeration. Their research indicators are industrial structure and industrial competition, e.g. Lederman and Maloney (2003), Jefferson et al. (2006) and Jianchu Shang (2005). "International trade theory" studies the impact of import trade, export trade and trade dependence rate on technological innovation capabilities. They support “learning-by-exporting effect”, e.g. Griffith (2000) and Peiyuan Xu (2012). What’s more, the "FDI spillover effect" emphasizes that foreign investment will promote the technology transfer and technology diffusion, thus it can improve the ability of technology innovation, e.g. Hu and Jeffersom (2001), Cheung and Lin (2004). There are also other theories, for example, "institutional economics" e.g. Aghion and Angeletos (2005), "intellectual property theory" e.g. Lederman and Maloney (2003) .

Obviously, most of these researches only consider one or several categories of influence factors and lack of considering the model uncertainty. They are prone to get one-sided and biased conclusions. Artificially determining variables and models in advance may lead to systematic deviations. In order to solve the problem of model uncertainty, this paper uses the Bayesian model averaging to improve the accuracy of the model.

II. MODEL UNCERTAINTY AND BMA

A. Model uncertainty

The idea of the Bayesian model averaging method (BMA) was proposed by Draper (1995) to solve the problem of model uncertainty. In the actual process of modeling, due to the complexity and multidisciplinary nature of research issues, the problem of model uncertainty is often encountered. For example, for the established multiple regression model,

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

Different researchers focus on different variables x_k, but researchers actually cannot accurately judge which explanatory variables x_k should be added to the multiple linear regression models. The econometric models established in this way may reach erroneous conclusions and have serious errors in the process of analysis, prediction, and policy evaluation, etc. This is a typical model uncertainty problem. The Bayesian model average problem is a tool for solving such problems.

  DOI: 10.25236/emcs.2018.055

B. The Bayesian model averaging

The Bayesian model averaging treats the influencing factors as random variables themselves. In other words, the variables themselves are stochastic and follow some kinds of probability distribution. The target model is uncertain. It is derived from the model space $\mathrm{M} = \{\mathrm{M}_1, \mathrm{M}_2, \ldots \mathrm{M}_\mathrm{K}\}$ which contains all the possible model M_K. Here M_K represents the Kth model, and we have,

$$P(\theta|X) = \sum_{j=1}^{k} P(M_j|X)P(\theta|M_j, X),$$

where θ is the parameter vector to be evaluated.

X is the data sample observed.

M_j is jth model from model space M.

K is the number of models in the model space.

This formula shows that the posterior distribution of the parameter vector θ, conditional on X, is the weighted average of the posterior distribution of the parameter vector θ, conditional on M_j and X. The weight is the posterior probability $P(M_j|X)$.

$$P(M_j|X) = \frac{P(X|M_j)P(M_j)}{\sum_{j=1}^{K} P(X|M_j)P(M_j)}$$

Where $P(X|M_j)$ represents the integral of M_j's likelihood function. It can be calculated using following formula

$$P(X|M_j) = \int P(X|\theta_j, M_j)P(\theta_j|M_j)d\theta_j$$

Where θ_j represents the parameter vector corresponding to model M_j

$P(\theta_j|M_j)$ is the prior distribution of θ_j corresponding to model M_j

$P(X|\theta_j, M_j)$ is the likelihood function of M_j

$P(M_j)$ is the prior distribution of M_j

We can obtain the posterior mean and posterior variance of parameter vector θ,

$$E(\theta|X) = \int_{-\infty}^{+\infty} \theta \left[\sum_{j=1}^{K} p(M_j|X)p(\theta|M_j, X)\right] d\theta = \sum_{j=1}^{K} \hat{\theta}_i \, p(M_j|X)$$

where $\hat{\theta}_i$ is $E(\theta|X, M_j)$. Similarly, we can get posterior variance,

$$Var[\theta|X] = \int_{-\infty}^{+\infty} [\theta - E[\theta|X]]^2 [\sum_{j=1}^{K} p(M_j|X)p(\theta|M_j, X)]d\theta$$

$$= \sum_{j=1}^{K} p(M_j|X)\left(\hat{\theta}_i - \sum_{j=1}^{K} p(M_j|X)\,\hat{\theta}_i\right)^2 + \sum_{j=1}^{K} p(M_j|X)Var^2(y|X, M_j)$$

In conclusion, the key point of the Bayesian model averaging is to take a weighted average of the possible single model and the weight is the posterior probability. Then the posterior probability is used as an objective criterion for selecting explanatory variables. Through this process we can solve the problem of model uncertainty in the modeling process. On the other hand, another advantage of the Bayesian model averaging is that it can easily combine subjective information with model information through prior distribution. By contrast, the Bayesian model averaging has higher accuracy than other model prediction methods. Based on these advantages, this paper applies the Bayesian model averaging to the study of technological innovation capabilities.

III. DATA

A. Measure of technological innovation capability

For the measurement of technological innovation capability, scholars have proposed different scales. Liu Linlin et al. (2002) believe that innovation capability is the ability to convert knowledge into new products, new processes, and new services. They use the number of patents to measure the city's innovative capability. The studies of Chen Guanghan et al. (2007) and Wang Qingyuan et al. (2010) find that there is a positive correlation between R&D investment and patent output, and they are believed that the number of patents could be used to measure the capacity of innovation output. Huang Zhiyong (2013) proposed that the number of patents only reflects the quantity of innovation and it needs a new variable to reflect the economic value of innovation. He proposed that it should be measured by the sales of new products. This theory has received wide

attention. Cao Yong et al. (2013) selected six indicators including the number of patent applications, the number of patents granted, new product sales， and high-tech industry output value et al. to build a factor model through factor analysis.

Based on these studies, this paper uses the sales revenue of new products as the measure of technological innovation capability. New products refer to the products that have been developed and produced using new technologies, new design concepts, or have significantly improved their structure, materials, and processes in comparison with the original products. Compared to the number of patents, new product sales revenue can reflect the value of technological innovation capabilities instead of quantity.

B. Influencing factors of technological innovation ability

As mentioned before, different researchers have proposed different theories about the influencing factors of large and medium-sized industrial enterprises' technological innovation ability. The variables selected in this article are summarized in table I. Among them, L&M indicates large and medium-sized industrial enterprises.

TABLE I: SUMMARY OF SELECTED VARIABLES.

No.	Variables	Variable name	Theory	Remarks
1	R&D investment	RD_K	Endogenous growth theory	Capital influence of L&M
2	R&D personnel input	RD_L		Human resources influence of L&M
3	GDP	GDP	Macro-economic situation	Market scale
4	GDP growth rate	RGDP		Economic growth rate
5	Rate of interest	IRATE	The level of finance development	The monetary cost of innovation investment
6	Exchange rate	ERATE		
7	The proportion of secondary industry in GDP	Industry1	Regional economic theory	Reflect the influence of industrial structure
8	The proportion of the tertiary industry in GDP	Industry2		
9	Enterprise scale	SIZE	Company characteristics	The company number of L&M
10	Foreign investment	FDI	FDI spillover effect	Foreign capital of L&M
11	Import amount	IM	International trade theory	“Learning-by-exporting effect” hypothesis
12	Imports trade dependence	IM/GDP		
13	Export amount	EX		
14	Export trade dependence	EX/GDP		
15	Government S&T input	Subsidy	Government policy support	The impact of government policy

Notes: These fifteen variables represent different theories of the influencing factors of industrial technological innovation ability. The fourth column lists the corresponding theory. The fifth column shows additional information.

IV. EMPIRICAL STRATEGY

A. Empirical model

If we have p explanatory variables $\{x_1, x_2, \dots, x_P\}$, and X_k is the subset of the whole set which contains one or several factors of $\{x_1, x_2, \dots, x_P\}$. Then we can construct the model:

$$M_k: y = X_k\beta_k + \varepsilon$$

Where X_k is the matrix of the explanatory variable we consider

β_k is the matrix of corresponding regression coefficient

ε is the error matrix, generally assumed as $\varepsilon \sim N_n(0, \sigma^2 I)$

B. Prior Distributions

To obtain the posterior distribution, we first need to make assumptions about prior distribution of the models and parameters. Firstly, for the prior probability of the model, this paper sets it as a uniform probability distribution. In other words, this means that there is no prejudice between every models and the prior probability of a single model is $P(M_i) = 1/K$.

Next, we need to think about the prior distribution of parameters $p(\theta_k|M_k)$. The most common way to set the prior distribution of parameters is Zellner's g prior. Zellner’s g prior is a common prior for coefficients in the linear regression model due to its computational speed of analytic solutions for posterior.

For the multiple regression model we discussed,

$$M_k: y = X_k\beta_k + \varepsilon$$

Zellner's g prior generally assumes that σ^2 have a noninformative prior distribution, then prior distribution of β_k, conditional on σ^2 is subjected to multivariate normal distribution with mean β^0 and covariance matrix $g\sigma^2(X_k'X_k)^{-1}$.

It can be expressed as follows:

$$p(\beta_k|\sigma^2) = N(\beta^0, g\sigma^2(X_k'X_k)^{-1})$$

$$p(\sigma^2) \propto \frac{1}{\sigma^2}$$

And the joint posterior density is

$$p(\beta_k, \sigma^2|X) \propto p(\beta_k|X, \sigma^2)p(\sigma^2|X)$$

We generally set the mean β^0 to 0, so we only need to specify the hyper-parameter g.

This paper use R Studio to handle details.

C. Computation of marginal likelihood function

For the marginal likelihood function, it usually involves to compute high dimensional complex integrand. There is no doubt that this is a big challenge. This is the major reason why Bayesian statistics have not developed over a long period of time. And the proposition of MCMC (Markov Chain Monte Carlo) has been a key step in making it possible to compute large hierarchical models that require integrations over hundreds or even thousands of unknown parameters.

The MCMC method is to construct an appropriate Markov chain for sampling and use the Monte Carlo integration method to do integral operation. By constructing a Markov chain that has the desired distribution as its equilibrium distribution, that is to say, the equilibrium distribution is the model posterior distribution $g(M) = P(\sigma|M, X)$. It means that one can obtain a sample of the desired distribution by observing the chain after a number of steps. The more steps there are, the more closely the distribution of the sample matches the actual desired distribution.

Monte Carlo integration refers to that we can compute an approximation of the integration by the law of large numbers, which can be made arbitrarily close to the integral we want. Suppose we have been able to obtain a sample $\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(M)}$ from $p(\theta|y)$. The value

$$\widehat{g_M(\theta)} = \frac{1}{M}\sum g(\theta^{(m)})$$

can be shown to converge to $E(g(\theta) \mid X)$ as M goes to infinity.

D. Model selection

As we assume there are 15 influencing factors, we can obtain at most 2^{15} models. It is computationally impossible to calculate every posterior probability and marginal likelihood on every model, so effective screening is very important. We can select the effective models to constitute a subset of the model space, and then do Bayesian model averaging on this subset. There are two main approaches at present. The first is the Occam's window method and the second is MC^3(Markov Chain Monte Carlo Model Composition). We use the Occam's window method here.

Occam's window method is proposed by Madigan and Raftery. Although the number of models is large, there are only a few models supported by the data actually. Occam's window method argues that when the number of variable K and the number of models are large, it is true that the posterior probability of most models is much smaller than that of the best model. Therefore, only those models whose posterior probability exceeds a fixed value need to be considered, and other models with low posterior probability could be ignored, improving efficiency. We can write:

$$M^* = \{M_j: \frac{\max_I \Pr(M_I|X)}{\Pr(M_j|X)} \leq C\}$$

Here C is a threshold chosen by the investigator and M^* represents the model space selected. $\max_I \Pr(M_I|X)$ represents the model with maximum posterior probability. If the ratio of $\max_I \Pr(M_I|X)$ to $\Pr(M_j|X)$ less than threshold C, we accept this model and not conversely.

V. Results

A. Result of postprob, postmean and postsd

As we mentioned above, faced with the complexity of the enterprise systems, the influencing factors of technological innovation ability are very extensive. We should not ignore the influence of any field on innovation ability in the study. Most of the previous researches only consider one or several categories of influence factors. However, the use of Bayesian average model method can effectively deal with this uncertainty.

Taking large and medium-sized industrial enterprises as the samples during 1990-2017, this article chooses the sales revenue of new products as the explanatory variable and the fifteen influencing factors as the explained variables. The research data during 1990-2017 comes from the National Statistical Yearbook and the National Statistical Yearbook of Science and Technology.

The results analyzed by R studio are as follows. All variables have been standardized.

TABLE II: CALCULATION RESULT OF POSTPROB, POSTMEAN, AND POSTSD.

Variables	Postprob	Postmean	Postsd	Rank
Intercept	100	0.0000	0.0139	
R&D investment	100	2.8250	0.8785	1
R&D personnel input	100	-2.9480	0.4141	1
Import amount	99.4	-2.6630	0.6522	3
Export amount	99.4	2.6680	0.6370	3
Export trade dependence	94.8	0.1932	0.0720	5
Government S&T input	81.8	1.0540	0.7598	6
GDP	54.2	0.0207	0.0262	7
Rate of interest	46.6	0.0228	0.0328	8
Imports trade dependence	39.9	-0.0351	0.0637	9
Enterprise scale	23.4	-0.0078	0.0208	10
The proportion of the tertiary industry in GDP	19.6	0.0156	0.0503	11
Exchange rate	16.6	0.0030	0.0197	12
The proportion of secondary industry in GDP	13.5	0.0025	0.0159	13
Foreign investment	11.5	-0.0050	0.0310	14
GDP growth rate	10.5	-0.0105	0.2293	15

Notes: This table presents the calculation result of postprob, postmean, and postsd. Postprob is the posterior probabilities of the variables selected. Postmean is the posterior mean of each coefficient from model averaging. Postsd is the posterior standard deviation of each coefficient from model averaging.

The result of posterior probability, posterior mean and posterior standard deviation of calculated are given in the table 2. Postprob is the posterior probabilities of the models selected. It shows the explanatory power of the variable to the explained variable. Postmean is the posterior mean of each coefficient from model averaging. Postsd is the posterior standard deviation of each coefficient from model averaging. It is generally believed that only when the posterior probability is more than 60% can the index variable be considered valid.

As can be seen from the charts, There are six variables, R&D investment, R&D personnel input, value of import and export, export trade dependence, and government S&T input in the 15 possible influence factor, have great influence on the technology innovation ability of industrial enterprises. Among them, except for R&D personnel input and imports volumes, the other four variables are positively correlated.

Not surprisingly, the highest posterior probability of variables is R&D investment. There is no doubt that R&D investment is an important factor of technological innovation ability, because R&D investment is the fundamental guarantee for enterprises to carry out innovation activities. This is consistent with the conclusions of most studies.

The relationship between innovation capacity and the number of R&D personnel is negatively correlated. This may be due to the fact that the excessive input of R&D personnel in China has caused the unreasonable allocation of resources, causing waste of public resources. According to statistics, the number of R&D personnel in China in 2011 was 1.76 times that of the United States, but the number of R&D personnel in actual employment is only 14% of the United States.

At the same time, the factors include both the amount of imports and exports. This shows that the technological innovation capabilities of large and medium-sized industrial enterprises are very relevant to international trade. The results of the export amount and export trade dependence have high posterior probability support the “learning-by-exporting hypothesis”. Competition and exposure to a superior foreign market can speed up technological acquisition and lead to capability promotion.

Government investment in science and technology has a positive impact on technological innovation capability. The government investment in science and technology conveys the information of the country's policy orientation and government support for large and medium-sized industrial enterprises. It guides the active entry of financial capital and social capital and effectively alleviate the financing constraints of the enterprise, reflecting the prospect of industrial development. It also can guide the capital inflow of financial capital and social capital, effectively alleviating the financing constraints of enterprises. In addition, this signaling effect of government investment in science and technology is also conducive to improving the acceptance of products by the market and customers.

The posterior probabilities of other nine factors are relatively low, indicating that these variables have weak explanatory power for technological innovation capabilities. These factors include GDP growth rate, foreign investment, the proportion of the secondary and tertiary industry in GDP, etc. The following figure gives a more direct reflection of the above results

The abscissa represents 112 qualified models selected by Occam’s window method. Blue color represents the negative effect of the variables in the relevant model, red color represents the positive influence in the relevant model, and the white represents no significant influence. The lateral lengths of the variables reflect the value of posterior probability.

As you can see from the figure, the six variables mentioned above have the longest x-coordinate length. It means that these are the variables used most in the selected models, having great influence on the technology innovation ability of industrial enterprises.

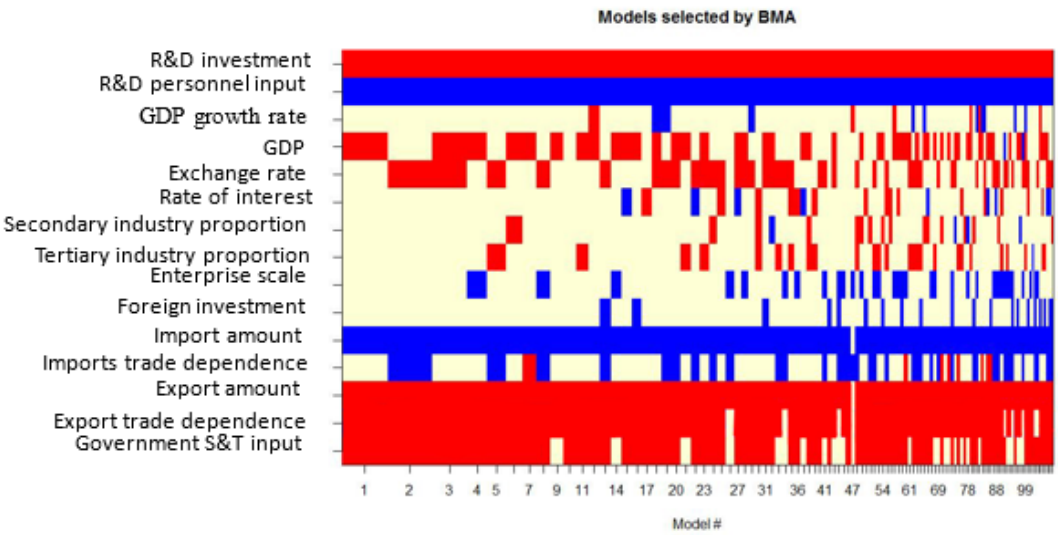

Figure I. Model selected and multi-variable analysis results

B. The best five models

The following table lists the best five models selected, the variables they contain, their BIC, R^2, and the corresponding posterior probability.

As we can see from table III, the probability of model 1 is 7.5%, and the probability of model 2 is 7.3%. These two models have the highest probabilities of all 2^{15} models. The cumulative posterior probability of the best five models is 30.1%. The high percentage indicates that the Bayesian model averaging method works well.

R.D investment, R.D personnel input, Import amount, Export amount, Export trade dependence, Government S.T input these six variables are included in all five models, and it indicates that these six variables are the most important variables. This is consistent with the conclusion drawn above.

The results of R^2 and BIC also show that these models fit well. R^2(coefficient of determination) is the goodness-of-fit of a model. It summarizes the discrepancy between observed values and the values expected under the model in question. The closer the goodness-of-fit is to 1, the better the fitting effect of the model. BIC is a criterion for model selection among a finite set of models; the model with the lowest BIC is preferred.

TABLE III: THE BEST FIVE MODELS SELECTED

	model 1	**model 2**	**model 3**	**model 4**	**model 5**
Intercept	0.0000	0.0000	0.0000	0.0000	0.0000
R.D investment	2.7270	2.4880	2.5340	2.7600	2.7070
R.D personnel input	-2.9800	-2.9720	-2.8780	-2.8610	-3.0300
GDP growth rate	.	.	.	.	.
GDP	0.0350	.	0.0373	0.0418	.
Exchange rate	.	0.0620	0.0270	.	0.0684
Rate of interest	.	.	.	.	.
The proportion of secondary industry in GDP	.	.	.	.	.
The proportion of the tertiary industry in GDP	.	.	.	.	0.0728
Enterprise scale	.	.	.	-0.0198	.
Foreign investment	.	.	.	.	.
Import amount	-2.7880	-2.7340	-2.9380	-2.6760	-2.6930
Imports trade dependence	.	-0.0882	.	.	-0.1069
Export amount	2.7780	2.7050	2.9140	2.6780	2.6850
Export trade dependence	0.2178	0.1701	0.2404	0.2098	0.1792
Government S.T input	1.2030	1.4130	1.3180	1.0550	1.1680
nVar	7.0000	8.0000	8.0000	8.0000	9.0000
r2	0.996	0.997	0.997	0.997	0.997
BIC	-128.6000	-128.5000	-128.0000	-126.9000	-126.7000
post prob	0.075	0.073	0.059	0.048	0.046

Notes: Statistical result shows that the cumulative posterior probability of the best five models is 30.10%. The high percentage indicates that the Bayesian model averaging method works well. It also can be seen from the table that the variables with high posterior probability selected as mentioned earlier are mostly included in these models. The results of R^2 and BIC also show that these models fit well.

VI. CONCLUSION

Because of the complexity of the enterprise systems, different researchers have proposed different theories of the influencing factors of technological innovation ability. Researches focusing on a single field tend to have the problem of model uncertainty. By treating the influencing factors as random variables themselves, this paper uses the Bayesian model averaging problem to solve the problem of model uncertainty.

Taking the new product sales revenue as the explanatory variable and fifteen variables which reflect different innovation theories as the explanatory variables, this paper analyzes the real influencing factors of large and medium-sized industrial enterprises' technological innovative abilities.

In empirical research, we set the prior distribution of parameters as Zellner's g prior, and the prior probability of the model was set as a uniform probability distribution. Besides, we use the MCMC method to compute marginal likelihood, which can solve the difficulty of high dimensional complex integrand calculation. Considering that the sample space is too large, we use the Occam's window method to select effective models from the model space, filtering out models that have very small posterior probabilities. The results of variables' posterior mean, posterior variance, and corresponding posterior probability are shown in table 1 and the best five models are listed in Figure 1.

We can see that, there are six variables, R&D investment, R&D personnel input, value of import and export, export trade dependence, and government S&T input Among the 15 possible influence factor, have great influence on the technology innovation ability of industrial enterprises. Except for R&D personnel input and imports volumes, the other four variables are positively correlated. The rest nine factors have no significant impact. These results enlighten our relevant policies.

Government and industrial enterprises should pay attention to the resource utilization efficiency of R&D investment. Especially, we should pay attention to the cultivation of practical talents that can be transformed into professional and innovative talents.

The results that export amount and export trade dependence have high posterior probability support the "learning-by-exporting hypothesis". It indicates that the government needs to improve the trade structure and give attention to the spillover effects of trade technology. For industrial enterprises, on the one hand, they should vigorously develop high-tech export abilities. On the other hand, they need to enhance enterprises' learning ability and make full use of the technology spillovers brought from import trade.

The high posterior probability of government investment in large and medium-sized industrial enterprises reflects that government actions can reflect government support for industrial enterprises. It can guide the active entry of financial capital and social capital, effectively alleviating the financing constraints of industrial enterprise. Government should pay attention to supporting the industry development and give more support to R&D activity of cutting-edge technologies. Meanwhile, it should focus on key technology projects and allocate scientific research funds reasonably, promoting the development of related industries.

REFERENCES

[1] Liang Wang. (2016). The Research and Prospect of Bayesian Model Average Method. Technical Economics and Management Research, 3, 19-23

[2] Xin Yu, Guohua Zhou. (2011). Model averaging method and its application in prediction. Statistical Research , 28(6), 97-103

[3] Feng Liang, Rui Paulo, German Molina, Merlise A Clyde, Jim O Berger. (2008). Mixtures of g priors for bayesian variable selection. Journal of the American Statistical Association, 103(481).

[4] Liu X H, Buck T. (2007). Innovation Performance and Channels for International Technology Spillovers: Evidence from Chinese High-tech Industries. Research Policy ,36(3) , 355-366.

[5] Aitken, A., Hanson, G. and Harrisom, A. E. (1977). Spillovers, Foreign Investment and Export Behaviour. Journal of International Economics ,43, 103-132

[6] EdwardI George, Dean P Foster. (2000). Calibration and empirical bayes variable selection. Biometrika 87(4):731–747.

[7] Falvey R, Foster N, Greenaway D. (2004). Imports, Exports, Knowledge Spillovers and Growth. Economic Letters ,85(2): 209-213.

[8] Feng, F., Bo, W., Yuanyuan Z. Yuneng, D. (2013). A New Internet DEA Structure: Measurement of Chinese R&D Innovation Efficiency in High Technology Industry . International Journal of Business & Management, 8(21) , 32-40.

[9] Fernandez C, Ley E, Steel M. F. J. (2001). Model Uncertainty in Cross-country Growth regressions. Journal of Applied Econometrics, 16(5) , 563 -576.

[10] Hoeting, J.A., Madigan, D., Raftery, A.E., Volinsky, C.T. (1999). Bayesian model averaging: a tutorial. Stat. Sci, 14 (4), 382-401.

[11] Joseph B Kadane , Nicole A Lazar. (2004). Methods and criteria for model selection. Journal of the American Statistical Association 99(465):279-290.

[12] Lee C Y. (2002). Industry R&D intensity distributions: regularities and underlying determinants. Journal of Evolutionary Economics 12(3), 307-341.

[13] Liang Wang, Jinquan Liu. (2010). Analysis of Determinants of China's Economic Growth: Based on the Bayesian Model Averaging (BMA) Approach. Statistics & Information Forum ,25, 3-7

[14] Madigan D, Raftery A E. (1994). Model Selection and Accounting for Model Uncertainty in Graphical Models Using Occam's Window. Journal of the American Statistical Association, 89(428), 1535-1546

[15] MA Efroymson, 1960. Multiple regression analysis. Mathematical methods for digital computers, 1,191–203

[16] Peiyuan Xu, Hua Zhang. (2014). Factors affecting technological innovation ability in Fujian Province. The Southeast Academic, (1), 124-134

[17] Raftery, A.E.; Madigan, D.; Hoeting, J.A. (1997). Bayesian model averaging for linear regression models. J. Am. Stat. Assoc. 92(437), 179-191

[18] Zhiyong Huang. (2013). Impact of R&D, FDI, and International Trade on Innovation Capabilities: An Empirical Analysis Based on China's Industry Data. Research on Industrial Economy, (3), 84-90