International Journal of Business, Human and Social Sciences
ISSN: 2517-9411
Vol:6, No:11, 2012

Estimating Cost of R&D Activities for Feasibility Study of Public R&D Investment

Ie-jung Choi

***Abstract*—**Since the feasibility study of R&D programs have been initiated for efficient public R&D investments, year 2008, feasibility studies have improved in terms of precision. Although experience related to these studies of R&D programs have increased to a certain point, still methodological improvement is required. The feasibility studies of R&D programs are consisted of various viewpoints, such as technology, policy, and economics. This research is to provide improvement methods to the economic perspective; especially the cost estimation process of R&D activities. First of all, the fundamental concept of cost estimation is reviewed. After the review, a statistical and econometric analysis method is applied as empirical analysis. Conclusively, limitations and further research directions are provided.

***Keywords*—**Cost Estimation, R&D Program, Feasibility Analysis Study.

I. Introduction

As public R&D investment continuously increases in Korea, the Korean government introduced a feasibility study which takes place before the actual investment. This feasibility study for public R&D investment has been implemented as a system since the year 2008, and since then, this study delivered great effort to determine the feasibility of each R&D program. Actually, the feasibility study was initially introduced in 1999. This preliminary research was conducted for the purpose of improving the effectiveness of all types of public finance programs, which include public transportation infrastructure such as roads, railways, harbor, etc. In 2008, feasibility study has expanded to fields and sectors of technology research & development, limited to programs with investment size over approximately $50 million where over approximately $30 Million is funded by government. A total of roughly 60 feasibility studies on public R&D investment programs have been conducted since [1].

In order to increase the effectiveness of public R&D investment through feasibility studies, various evaluation studies have been performed. Systematic evaluation for on-going national R&D programs conducted within the ministries, periodical program evaluations like high-level assessment and in-depth assessment, performance evaluation at completion and prediction of the potential results at implementation [2]. A variety of analysis methods, such as experimental design and statistical review, have been utilized to examine a causal relationship among the results, effects, outputs and inputs, for the purpose of obtaining a more systematic analysis of government R&D programs[3]. An attempt to measure the benefits and effects along with the appropriate cost of government R&D programs is the part of such endeavors.

Despite the effort, a current cost estimation procedure of public R&D investment has its limitations. The objective of this research is to provide an opportunity to investigate the current cost estimation procedure and to provide an improvement solution to the current situation.

The following of this research is divided into five sub-sections. First this research observes the present concept of cost estimation in feasibility studies and general concept of cost estimation. Then an improvement method is introduced followed by an empirical analysis. Conclusively, discussion of empirical analysis results is provided with the conclusion at the end of this research.

II. Concept of Cost Estimation

A. Present Estimation Concept

Currently, the general R&D cost estimation is done with similar historical samples. In other words, the feasibility studies concerning appropriate cost of public R&D are executed via the similar programs which have been initiated in the past. In addition, based on the prior R&D programs the specific cost of R&D program, including R&D activities, R&D related facilities, R&D equipments, etc., being studied are determined.

However, identifying the appropriate corresponding historical data itself is difficult and deriving the appropriate level of cost for the specific R&D activities could be biased. This is due to a few following reasons. First of all, difficulties on identifying the proper ‘so called’ similar R&D investments are an important cause. Also, difficulties on matching the current R&D activities with initiated programs due to differentiated size, focus, length, etc. is another essential factor which make the cost estimation difficult. Last but not least, especially difficult due to the fact that most public R&D investments are initiated in fields where not much previous investments have been provided, is a primary concern.

Conclusively, due to these issues, the R&D cost estimation on public R&D investments have limitations in terms of preciseness and accuracy.

B. General Cost Estimation Concepts

Before developing or modeling improvement schemes for cost estimation, general cost estimation concepts are overviewed. The concept of cost estimation has been an issue in

Ie-jung Choi is with the Korea Institute of Science & Technology Evaluation and Planning, Seoul, Korea (phone: 02-589-2283; fax: 02-589-5231; e-mail: matthew@ kistep.re.kr).

many different fields of study. Cost estimation itself can be defined as a mathematical algorithm or parametric process to estimate the cost of a specific subject. It is executed to forecast the cost of something that has not been taken in place yet. In similar terms the U.S. Government Accountability Office(GAO) defines cost estimation as, "the summation of individual cost elements, using established methods and valid data, to estimate the future costs of a program, based on what is known today" [4].

Cost estimating methods may vary by type and class. In the perspectives of cost engineering, three different general methods could be introduced.

TABLE I
GENERAL CONCEPTS OF COST ESTIMATION METHODS

Methods	Strength	Weakness	Application
Analogy	Possible with less data Based on actual data Relatively quick	Relatively subjective Uncertain similarity Differentiation unapplied	Little feasible data Approx. estimation
Bottom-UP	Possible to track Procurement cost	Necessary detail plan Complex	SW development Negotiation
Parametric	Relatively objective Identifying cost factor	Lack of concreteness Needs to development model Needs of understanding model	Budget estimation Comparison

As displayed in table above, each cost estimation methods has its own pros and cons. Analogy is a method were similar cases are applied as standard. This method uses similar samples to estimate the subject of interest.

On the other hand, Bottom-Up methodological concept is rather a structural process. This method shares the fundamental core of WBS (work break down structure). A WBS is a structural decomposition of elements and components of projects or programs [5]. The Bottom-Up method, similar with WBS, defines and groups discrete elements to organize, establish and determine the whole program.

Parametric method is applying modeling to the cost estimating procedure. This method usually starts by identifying the cost factor. After identifying the cost factor, a model which explains the effect of the cost factor on total cost is analyzed. This method generally applies econometric analysis or regression analysis to recognize the relation between the cost factor and total cost.

III. IMPROVING COST ESTIMATION

Due to the large budget size and complex characteristic of the R&D programs, cost estimation for public R&D investment has been quite a challenge for both the government and research institutes. Presently, cost estimation within feasibility studies are based on either Analogy methods, Bottom-Up methods or both combined.

Although slight alternations exists, depended on the related research team which conducts the feasibility study, the basic cost estimation structure in feasibility studies for public R&D investment is divided into components such as R&D activity, R&D equipment, R&D facility, etc. Additionally, these investment of R&D activity, equipment, facility, etc. are classified and separately estimation before added up for final estimation, which is similar to the Bottom-Up method.

Furthermore, cost components such as R&D equipments and facilities are divided into smaller scaled components. Then, these smaller scaled components are estimated by unit cost based on past data. Cost estimation of these smaller scaled components and even the whole R&D equipments and facilities are determined to be quite accurate.

Unfortunately, costs of R&D activities are still roughly estimated. Commonly, as cost estimation of R&D activities for the feasibility study, Analogy is applied. To put it concretely, when estimating cost of R&D activities, as the first step, few or at least one similar program which has been launched in the pass are investigated. After similar programs have been identified, direct comparison among the individual projects of past similar programs and individual projects of presently studied program is performed. Based on this comparison, the appropriateness of the individual projects consisting the subjective program is determined. This rough estimation has various limitations. Among these limitation a serious limitation is that Analogy method applied in the current estimation process is cost estimation without considerations of each programs' properties, sizes, etc.

In this research, an improvement solution of cost estimation is introduced. The improvement solution is to apply both Analogy and Parametric estimation method to the R&D cost estimation. In detail, Analogy method is applied in terms of collecting and classifying historical data which show similarity with the subjective program. On the other hand, Parametric method is applied as procedure of identifying the cost factor and effect estimation of the cost factor on the total cost. Conclusively, the cost estimation equation can be derived.

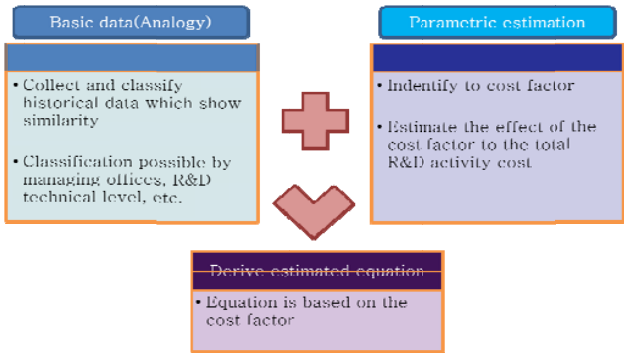

Fig. 1 Improvement Solution for Cost Estimation

Despite the fact that the introduced cost estimation procedure is expected to improve the current cost estimation procedure it should be empirically carried out to test the possibility of

applying to actual circumstances. Therefore, in this research, empirical analysis is provided to display the actual application process. Also, the empirical analysis is predicted to display specific limitations, difficulties and important issues when actual application takes place.

IV. Empirical Analysis

The empirical analysis performed through a few stages. First of all, data is collected and classified through the investigation of historical public R&D programs. Secondly, cost factor is to be identified based on the evaluation of the data. Thirdly, modeling of the R&D activity cost is practice with the identified cost factor. The following is the detail process of the empirical analysis by stages

A. Data

As mentioned above, collecting and classifying the data is the procedure where the concept of Analogy method is implemented.

In this research, the main source of data is data collect from public institutes which managing public R&D projects. Specifically, data used in this research is data collected from 9 different public R&D related research institutes working for 6 different public offices in collaboration. Correspondently, the collected data are implicated from public R&D investment projects which were completed in the years from 2005 to 2011.

Structurally, R&D cost data is made up of 5 different factors: labor cost, equipment cost, research cost, outsource research cost, and overhead cost.

Although data of completed projects were collected, ex post facto, limitations of the data in terms of detail exist. Information such as participation rate of each researcher is excluded from the collected data. As result, most of the collected data are low in quality due to the difficulties in management of public R&D investment program data.

Under these conditions, it is inefficient to apply all the collected data to the empirical analysis. Thus, one of the data sets which were provided by one of the R&D program management institute, is analyzed as an empirical analysis sample.

B. Cost Factor Identification

As the second step of this empirical analysis, cost factor must be identified. To identify the cost factor two separate stages are implemented in this research. First stage is where the cost factor which represents the total public R&D cost. Then, as for the second stage, the specific cost factor is identified as the parametric elements of the cost model, which is developed in the later section of this empirical analysis.

A simple linear regression model is used to identify the cost factor.

$$Total\ R\&D = \beta_0 + \beta_1 \times Cost\ Factor + \varepsilon \quad (1)$$

where β_0 is a constant and β_1 is the parameter which represents the effect of Cost Factor on Total R&D variable. ε is the stochastic term. The 'Cost Factor' variables are filled in with the 5 different factors (labor cost, equipment cost, research cost, outsource research cost, and overhead cost) which are the elements of the collected data.

The econometric estimation of the simple linear regression is conducted via OLS (Ordinary Least Square) estimation. The estimation results of each cost factor are shown below.

TABLE II
Estimation Results of Cost Factor

Factor	Constant (p-value)	parameter (p-value)	R-square (adj R-square)
labor	17.93 (0.467)	3.135 (0.000)	0.6402 (0.6399)
equipment	285.8 (0.000)	1.438 (0.000)	0.8260 (0.8258)
research	375.0 (0.000)	5.209 (0.000)	0.5474 (0.5470)
outsource research	409.6 (0.000)	6.193 (0.000)	0.5395 (0.5391)
overhead	248.4 (0.000)	11.47 (0.000)	0.3330 (0.3324)

Conclusively, based on the estimation results, the 'equipment cost' factor is shown to display the highest descriptive capability followed by the ‘labor’ factor. Therefore, the 'equipment cost' factor should be selected as the cost factor, originally. However, based on common sense, the 'equipment cost' is greatly differed by programs and projects. In other words, the fluctuation of absolute ‘equipment cost’ size is large inducing bias. Hence, in this research further identification of the cost factor is based on the 'labor cost' factor which shown the second highest descriptive capability in the estimation results.

C. Cost Modeling

Before establishing the actual cost model for analysis, certain assumptions and boundaries must be deployed. Firstly, as sampling for estimation, modeling and estimation is done with specifically focused data: data related to public R&D activities done in Universities. Next, estimation model for 'labor cost' is derived from the number of human resource classified by education level.

The final cost modeling equation is as below.

$$Labor = \beta_0 + \beta_1 \times Bachelor + \beta_2 \times Master + \beta_3 \times Ph.D. + \varepsilon \quad (2)$$

here β_0 is a constant and β_1, β_2 and β_3 are the parameters which each represents the effect of classified human resource by degree of education on Total Labor budget. ε is also the stochastic term for econometric estimation.

The econometric estimation of the multi-variable linear regression is also conducted via OLS (Ordinary Least Square) estimation. The estimation results of each parameter are shown in Table III.

Based on the estimation results, it is possible to confirm that the number of human resource with Ph. D affects the total 'labor cost ' by approximately more than two times than others. This result matches with common sense, which indicate that the estimation result is general buyable.

TABLE III
ESTIMATION RESULTS OF COST MODEL

Estimation mode for labor cost derived from the number of human resource classified by education level			
	Parameter	P-value	R-square (adj R-square)
Ph.D	9.226	0.000	
Master	3.381	0.000	
Bachelor	4.204	0.000	
Others	2.798	0.011	
constant	45.22	0.000	
1000 thousand Korean won per personnel			0.7418 (0.7392)

With both estimation results of the cost factor identification and estimation results of the cost modeling, introduced above, the cost estimation equation for total R&D activities can be derived with the equation below.

$$Total\ R\&D = \beta_{0,0} + \beta_{0,1} \times (\beta_0 + \beta_1 \times Bachelor + \beta_2 \times Master + \beta_3 \times Ph.D.)$$

However, validation of the equation is not applied yet for actual application. Therefore, it seems to be insufficient to provide a finalized cost estimation equation for now.

V. DISCUSSION AND CONCLUSION

In this research, an overview of the current cost estimation in public R&D investment is introduced. Along with the introduction, an improvement solution is provided, followed by an empirical analysis to display the detail procedure.

Through this research it was possible to derive a parametric equation, based on historical samples, which can estimate the cost of public R&D activities. However, limitations exist in terms of the methodology and empirical analysis example.

First of all, the empirical analysis conducted in this research excludes the factor of public R&D investment period. Expressed in a different way, length and size of R&D investment could have different effects with the labor factor, and the introduced cost model does not imply this effect in anyway. Secondly, estimations in this research such as cost factor identification and cost modeling the labor cost with number of human resource classified by education levels are limited to a linear model. Thus they were analyzed with other various models introduced in past studies in similar researches. Thirdly, data handling is another limitation. In this research, outliers in the data have not been excluded from the estimation. To derive more absolute estimations for the parameters, an intense data sorting process should be considered. Last but not least, more data in terms of quantity and quality must be collected and analyzed for more accurate estimations.

Although the introduced improvement solution in this research indicates limitations, it could be hopefully be a significant approach to future research directions related to similar fields of study. In future researches cost elements, cost indexes and other tools such as promote time adjustment of capital costs, following changes in technology, availability of materials and labor, and inflation [6], could be additionally considered in cost estimation.

Feasibility study, in Korea, is applied to demonstrate the feasibility of large-scale, long-term public investment R&D programs and also to enhance fiscal efficiency and productivity. As feasibility study on public R&D program is still at an initial stage, small developments and improvements, as this research, are expected to reinforce the efficiency and productivity of the study itself.

REFERENCES

[1] S. Yim, H. Kang, "Feasibility Analysis Studies on New National R&D Programs in Korea," World Academy of Science, Engineering and Technology 59, 2011.
[2] Y. B. Lee, J. Hwang, "A Study on Correlation between Elements of AHP for Government R&D Programs," Proceedings of the International Symposium on the Analytic Hierarchy Process, 2011.
[3] E. Triantaphyllou, *Multi-Criteria Decision Making Methods: A Comparative Study*, Kluwer Academic Publishers, 2000.
[4] GAO Cost Estimating and Assessment Guide, Best Practices for Developing and Managing Capital Program Costs, GAO-09-3SP, United States Government Accountabity Office, 2009.
[5] G. T. Haugan, The Work Breakdown Structure in Government Contracting, Management Concepts, 2003.
[6] K. K. Humphreys, 2005, "Project and Cost Engineers," Handbook. 4th ed.