



AN ALTERNATIVE APPROACH OF INPUT-OUTPUT TABLES TO DYNAMIC STRUCTURE CHANGES IN KOREAN IT INDUSTRIES

Byung-Sun CHO^a, Sang Sup CHO^b, Jungmann LEE^c

^a*Electronics and Telecommunications Research Institute (ETRI),
161 Gajeong-dong, Yuseong-gu, 305-350 Daejeon, South Korea*

^{b, c}*Hoseo University, Anseodong 268 Cheonan city, Chungnam Province, South Korea*

Received 02 March 2011; accepted 08 January 2012

Abstract. The structure of the IT industry has always evolved in line with technological progresses and changes in consumer preferences, as well as with regulatory trends. This is why, when assessing the effect that a new technology or industry policy may have on the national economy, companies and policy-makers need to consider dynamic structural changes affecting the IT industry. One of the most popular existing methods for economic impact analysis is based on a traditional input-output table, and is conducted over a period between the current time and a given time in the future. In this study, we compare the accuracy of RAS and Cross Entropy (CE), the two most widely employed methods for updating input-output (IO) tables, by applying them to Korean IT industries. The main results of this study are the following. In terms of the accuracy of input coefficient estimates, we have found that both the RAS and CE methods have a tendency to overestimate or underestimate them. When the Korean industry was first divided into fourteen sectors, and the RAS and CE methods were applied to each of the fourteen industries, it was difficult to discern a consistent trend for the two methods concerning their accuracy in estimation of input coefficients. Secondly, when used to update an IO table in which the IT industry is subdivided into IT equipment and services, neither the CE nor RAS method proved distinctly superior to the other. Third, in light of the above two findings, we concluded that updating IO tables is best done through a hybrid method combining the CE and RAS methods. This paper proposes a procedure consisting of two steps: IO tables are first updated using the two methods, which are once again updated by employing the OLS average approach through the use of optimal weights.

Keywords: IO tables, RAS, cross entropy (CE), mallows criterion, optimal weights, information technology, Korea.

Reference to this paper should be made as follows: Cho, B.-S.; Cho, S. S.; Lee, J. 2013. An alternative approach of input-output tables to dynamic structure changes in Korean IT industries, *Technological and Economic Development of Economy* 19(2): 257–271.

JEL classification: C11, C63, C67, D57.

Corresponding author Jungmann Lee
E-mail: mam@hoseo.edu

Introduction

IO tables are an important means of detecting changes affecting the structure of a country's industry (Acemoglu *et al.* 2011; Atalay *et al.* 2011; Elliott *et al.* 2011; Jones 2011). They are also an effective tool for assessing the economic effect of a policy undertaking and the cost-benefit relationship of any alternatives that are considered. Kim (Kim 2000) can be referred for a discussion on basic analyses related to input-output tables and their areas of application. The accuracy of IO tables is of paramount importance because a correct structural understanding of industry sectors, as well as the validity of economic analysis, depends upon it. Rim (Rim *et al.* 2004) are a fine example of studies investigating the accumulation of R&D capital and externalities of the IT capital, using IO tables. Wan (Xing *et al.* 2011) investigates the convergence hypothesis using IO tables for telecommunications industry. In Korea, IO tables are compiled and published by the Bank of Korea. However, due to the complexity and cost of the data collection process, the Bank of Korea currently issues five-year benchmark tables instead of annual tables, publishing, in some special circumstances where unusual changes occur to the economy, extended versions based on these tables. There is a call for compiling new IO tables on a yearly basis. However, the argument by the proponents of this idea that significant and measurable changes occur to the structure of the industry over a period of one year is of questionable validity, and the practical usefulness of publishing new IO tables annually remains doubtful. Therefore, to perform an economic analysis using IO tables of the Korean economy, they need to be appropriately updated and re-adjusted. There are various methods for updating an IO table.

The goal of this study is to contribute to the methodology for updating IO tables by comparing RAS, the most widely used method for this purpose, and Cross Entropy (CE), a comparatively recent method, in order to determine their relative accuracy. This study has both theoretical and practical significance. It is also significant for its timeliness, given that the Bank of Korea (Bank of Korea 2003, 2008, 2009, 2010, 2011) has recently adopted a modified RAS approach for estimating input-output tables. Meanwhile, with regard to the industry sector in which structural changes take place at a rapid pace, the tremendous contribution of IT to Korea's national economy gives further significance to this study comparing the accuracy of IO estimates using these two methods.

To compare the two methods for updating IO tables for their accuracy and explore a new updating method based on the results of this comparison, in this study, we have used IO tables of the Korean economy for the years 2000 and 2005. The two updating methods were evaluated by applying them to Korea's fourteen major industries, with a particular focus on IT, to determine their relative superiority based on the value of the minimum error. We first estimated input coefficients based on the 2000 and 2005 transaction tables of domestic goods and services by industry using the two methods. These estimates were then compared with actual input coefficients for the year 2005. Finally, based on the results from the above steps, we propose a new method for updating IT IO tables, combining the two above methods through an OLS average approach.

The two IO table updating methods evaluated in this study are the RAS and CE methods proposed by Stone (Stone 1961) and Golan (Golan *et al.* 1996), respectively. They are the two most widely used methods since the methodology for updating IO tables was first discussed decades ago, and are used both in academic and industry research organizations. In recent years, Robinson *et al.* (2001) in their discussion of applicability of the CE and RAS methods, used the CE method for updating the SAM (Social Accounting Matrices) of emerging countries. Meanwhile, Ahmed (Ahmed *et al.* 2007) compared RAS and CE by applying them to Korean IO tables and found that CE was a more accurate updating method than RAS in regard to Korean tables. Ahmed (Ahmed *et al.* 2007) is distinct from this study insofar as the former analyzed the input-output relations of the overall Korean industry, instead of comparing the accuracy of RAS and CE by applying them to specific industry sectors.

The comparative analysis of the two methods based on Korean IO tables resulted in the following findings. Firstly, when the input coefficients estimated using the RAS and CE methods were compared with the actual input coefficients, both methods revealed a tendency for overestimation or underestimation (estimated input coefficients either exceed or fall short of actual input coefficients), with the extent of overestimation or underestimation varying depending on the industry sector. However, the gap between the sum of input coefficients from the actual IO tables and the sum of estimated input coefficients was smaller in the case of the IT industry than for other industries. These results suggest that errors resulting from updating an IO table are comparatively minor for the IT industry. They also suggest that relying on any single method for updating IO tables in the context of an analysis of the future industry structure or economic impact of a policy undertaking is likely to lead to estimation bias. Therefore, a method that appropriately combines RAS and CE could be a better alternative ensuring greater accuracy of estimation. As a way of combining estimates from the two methods, using the average of coefficient estimates obtained through an OLS estimator, this study proposes calculating optimal weights for the two estimates, employing the Mallows criterion, and updating the IO tables. Although there are many different methods for combining estimates from the two methods, the most widely employed method with econometric relevance uses least squares estimates. This study will not discuss other methods for combining coefficient estimates, as this lies outside its scope. The combination of existing methods to create a composite estimation method, as proposed in this study, has been widely practiced in recent times to predict economic variables with large fluctuations in time series structure (Stock, Watson 2004; Rapach, Strauss 2010). There is a large econometrics literature on choice or combination of forecasting estimators reviewed by Geweke and Amisano (2011) and Geweke (2010). Smith and Wallis (2009) offer the basic insight that forecast combination can generate complex puzzles as well.

The rest of this paper is organized as follows. We begin by briefly reviewing the theoretical background of the RAS and CE methods. Next, we apply them to actual IO tables of the Korean economy to compare their respective estimates with regard to the corresponding input coefficients matrix. For the IT industry, we use an optimal weight for each of the two methods. In the final section of the paper, we review the results of the comparative analysis and suggest some implications for applying the mentioned updating method to the IT industry.

1. Literature review and method

1.1. Literature review

Attempts to compare RAS and CE estimates have thus far led to varying results, depending on the analytical setup. (Robinson *et al.* 2001), for instance, found that the CE method was superior to the RAS method in terms of accuracy of update estimates for an IO table. On the other hand, (Jackson, Murray 2004; Junius, Oosterhaven 2003) maintained that RAS is a simpler and more economical method, likely to continue to remain in use, while stressing the need for developing a new and more robust method for updating IO tables. They demonstrated that relative estimation accuracy among RAS and other RAS-based updating methods can significantly vary depending on the criteria used for their comparison and also depending on the intended use of the IO data.

McDougall (McDougall 1999), describing CE and RAS as two methods that are similar to each other, pointed out that RAS can be expanded, using CE, to form a more practical and useful method. In his evaluation of the two updating methods, he stated the following: First of all, the RAS method can be included among generalized CE methods. Secondly, the CE method can be useful not only for updating IO tables, but also for updating other types of matrices. Based on these considerations, he argued that the RAS method can be expanded, using the theoretical CE model, into a more efficient updating technique.

(Gilchrist *et al.* 1999) demonstrated that using additional information from an IO table (information in cells) can considerably improve the updating estimation efficiency of a classical RAS-based approach. They also showed, using Canadian data that by incorporating information from IO tables related to the national-level economy in the analytical data can help increase the accuracy of updating IO tables for local industries quite significantly.

The Bank of Korea (2003, 2008, 2009) recently extended the IO tables of the Korean economy, using a new method. The 2005 benchmark IO data was extended to provide input-output relations for the year 2006, using a new sum of rows and columns. The updating approach used for this extension is a modified RAS method. Using this estimation method, the Bank of Korea (2003, 2008, 2009) derived a variety of inducement coefficients for 403 sub-industry sectors. Finally, (Huang *et al.* 2008) proposed IGRAS and INSD as techniques for expanding the classical RAS method in a manner to allow the updating of negative matrix elements, while keeping the negative sign untouched.

From the above discussion of the existing literature, we draw the following implications: First, the choice of a method for updating IO tables may differ depending on the intended use of the input-output data. Second, this choice can also depend on how much preliminary information is available for the evaluated industry. Third, IO tables have been mostly updated using a single method, while an econometric approach, appropriately combining RAS and CE (the two most widely used methods) can be a fine alternative that is potentially more efficient. Finally, a tacit presupposition in existing studies is that, as an updating estimate is only one of the many equally valid estimates, updating input-output transaction matrices necessarily involves the subjective judgment of the researcher who performs it. Given the involvement of subjective judgment, a researcher's knowledge about an evaluated industry sector plays a role in the estimation, and may influence the precise way a given estimation method is used as well.

1.2. Updating IO tables using the RAS technique

Even if one is unaware of the exact flow of goods and services between industries within an IO table, there are various ways of updating this table, as long as the values of all rows and columns are known. Data, such as output by an industry or gross national income, annually published by official sources, is estimated, and benchmark IO tables are only compiled once every five years. Therefore, to estimate the production flows between different industries, it is necessary to add new data to existing IO tables to create updated ones. In other words, if the input coefficient matrix of a new input-output table is A^* , an updated IO table can be estimated by choosing an input coefficient matrix that is the closest to the existing input coefficient matrix. Supposing that T^* is the transaction matrix expressing the input/output flow between two industries, one can write the following equations:

$$t_{ij}^* = a_{ij}^* y_j^*, \text{ and} \tag{1}$$

$$\sum_j t_{ij}^* = \sum_j t_{ji}^* = y_i^*. \tag{2}$$

Here, y^* is the estimate corresponding to the sum of the new rows and columns.

A common and intuitive method for estimating A^* , the new input coefficient matrix, which is the closest to \bar{A} , the existing input coefficient matrix, is proportional estimation, which takes into consideration the sum of rows and columns proportionally. That is,

$$a_{ij}^* = \bar{r}_i \bar{a}_{ij} \bar{s}_j. \tag{3}$$

Now, the above matrix can be simplified into the following formula:

$$A^* = \hat{r} \bar{A} \hat{s}. \tag{4}$$

The values of \hat{r} and \hat{s} , standing for weights for the input coefficient of the existing IO table, can be obtained through simple and repeated computational operations.

1.3. Updating IO tables using the cross entropy (CE) technique

When in an inter-industry transaction table the industry is classified into n number of industry sectors, the number of input coefficients to be estimated is n^2 . However, the information that is already known is independent from $2n - 1$, which corresponds to the sum of new rows and columns. In this case, under a RAS approach, in which there is a restrictive condition of proportionality, the estimation of the input coefficients is done by determining the values of \hat{r} and \hat{s} for $2n - 1$. However, with the CE technique, which involves the estimation of a fair set of parameters from a small quantity of available information, the approach to estimation is quite distinct.

(Golan *et al.* 1994), in their discussion of techniques for estimating a large number of parameters using a small amount of data, proposed a method for re-estimating the input coefficients of an IO table, using the sums of new rows and columns of a new industrial output.

With the CE method, input coefficients are obtained in the following manner. Let us first suppose that the probability for each of the n events, E_1, E_2, \dots, E_n , to occur is q_1, q_2, \dots, q_n .

However, new information resulting from a change in economic conditions can change the above a priori probabilities to a posteriori probabilities of $p_1, p_2 \dots p_n$. In this case, according to the theory advanced by Shannon (Shannon 1948), the probability that a certain amount of information is implicit in a probability is $-\ln p_i$. Each of the events, in other words, E_i , can have additional information from q_i , its a priori probability, and p_i , its a posteriori probability, which leads to the following formula:

$$-\ln \frac{p_i}{q_i} = -[\ln p_i - \ln q_i]. \tag{5}$$

The expected value of each element of information can be calculated by

$$-I(p; q) = -\sum_{i=1}^n p_i \ln \frac{p_i}{q_i}. \tag{6}$$

Here, $I(p; q)$ is the measure of the cross entropy distance between two probability distributions. The new input coefficient is estimated using the CE method, basically by determining the set of a posteriori probabilities under given restrictive conditions, which minimizes the cross entropy distance between a priori and a posteriori probability distributions.

Using the above-described method, Golan *et al.* (1994) employed a CE technique for estimating input coefficients from an IO table. The formula proposed by Golan *et al.* (1994) and Robinson *et al.* (2001) for SAM updating may be modified in a manner appropriate for use with updating of an IO table (see Ahmed *et al.* 2007). This study distinguishes itself from Ahmed *et al.* (2007) in that rather than comparing the two updating methods, it shows how the combination of the two is more efficient than either of them on their own and presents a single technique combining the two. Accounting to Golan *et al.* (1996), to minimize Eq. (6) subject to data consistency, normalization, and new information constraints, following equations are necessary.

$$y_{ij} = Xp_{ij}; \tag{7}$$

$$p_{ij}'1 = 1. \tag{8}$$

In concrete terms, they wrote the following formula to identify a_{ij} , a new input coefficient that minimizes the cross entropy distance from the existing input coefficient:

$$\min \left[\sum_i \sum_j p_{ij} \ln \left(\frac{p_{ij}}{q_{ij}} \right) \right]. \tag{9}$$

s. t.:

$$\begin{aligned} \sum_{i=1} p_{ij} &= 1 \\ \sum_j p_{ij} X_j &= Y_i \\ \sum_i p_{ij} X_j &= X_j. \end{aligned} \tag{10}$$

Here, X_j, Y_i are the column and row totals, respectively of the IO table that is being updated.

The comparison of the two estimation methods reveals, in sum, that RAS offers the advantage of ease of calculation. CE, on the other hand, provides estimates that are more consistent and have maximum likelihood properties under a given set of assumptions on probability distributions. Golan (Golan *et al.* 1996) and Robinson (Robinson *et al.* 2001) give more details for discussion on this topic. Meanwhile, unlike in RAS-based estimation, which requires the sums of all rows and columns to be known, with the CE technique, all needed parameters can be estimated as long as there is some information about row and column sums.

Ahmed *et al.* (2007) compared estimation errors between CE and RAS using Korean input-output data and found that estimation errors were smaller with the former than with the latter technique. However, the RAS technique is currently more widely used than the CE approach (Israilevich *et al.* 1997), (Okuyama *et al.* 2002). Therefore, the choice between these two methods depends on the goal and object of estimation and the amount of effort required for calculating a given estimate.

2. Comparison of RAS and CE techniques

2.1. Method of comparison

To compare the relative accuracy of the methods in updating an input-output matrix, we used the inter-industry transaction table for the year 2000, which was re-organized into fourteen sectors. Next, using the 2000 table of transactions of domestic goods and services, we measured the input coefficients of the fourteen industries. Then, using the transaction table of domestic goods and services for 2005, we calculated the sum of all intermediate demands between industries and combined the total input and value-added of all industry sectors. In other words, the sums of the rows and columns of the industrial input structure were derived from the 2005 IO table. Finally, based on the 2000 inter-industry transaction table, we estimated the input coefficients of the fourteen industry sectors for the 2005 transaction table of domestic goods and services, using the RAS and CE techniques. Finally, the estimated transaction table of domestic goods and services was compared to the actual 2005 transaction table to measure its accuracy.

Table 1 is the Korean industry classification table used in this study, and Fig. 1 provides a diagram of steps in the research process for this study explained in the preceding paragraph. As can be noted from Fig. 1, updating an IO table using the RAS and CE techniques

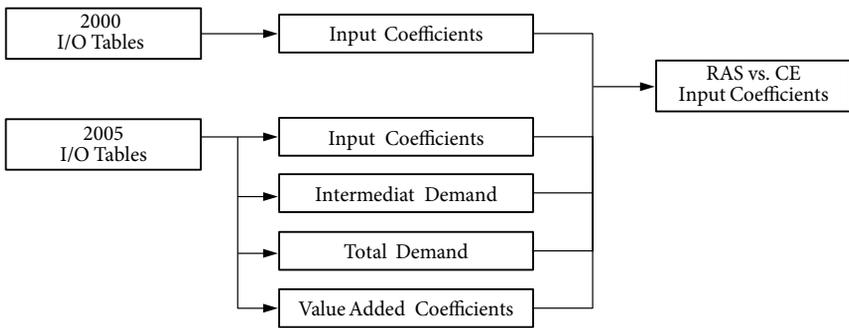


Fig. 1. Steps in comparison of RAS and CE methods

requires information including input coefficients of an existing IO table, the GDP, total output, total intermediate input, and the value-added coefficient of the target year. However, GDP data of the target year is only used in an estimation using the CE approach, not the RAS approach. The use and non-use of certain information each have their own advantages and disadvantages in terms of cost and estimation efficiency. Therefore, it is rather difficult to determine the relative superiority of one method over the other on the basis of use or non-use of information alone.

2.2. Comparative results

The results of estimating the input coefficients for the fourteen industries in Table 1 using the RAS and CE methods are shown in the graphs in Fig. 2. The three graphs show the sum of the actual input coefficients of the fourteen industries evaluated for the year 2005, the sum of the input coefficients estimated using the RAS technique, and the sum of the input coefficients estimated using the CE technique. As seen in these graphs, there is no clear pattern for either the RAS or CE estimates. The estimation results under these two approaches vary quite randomly for the fourteen industries evaluated, with neither of them consistently closer to the actual input coefficients; therefore, it is not possible to judge which of the two is more accurate than or superior to the other. Although the RAS approach appears superior in its

Table 1. Fourteen-category classification of Korean industries

	Industry	2005 revisions
1	Agriculture, fisheries and mining	
2	Manufacturing of textiles, textile products, leather, and leather products	
3	Manufacturing of chemicals	
4	Manufacturing of steel and steel products	
5	Manufacturing of automobiles and auto parts	
6	Shipbuilding	
7	Other manufacturing	
8	Electricity, gas, and water supply	
9	Construction	
10	Wholesale and retail, restaurants and hotels, transportation, storage, and postal services	
11	Finance, insurance, real estate, and real estate development services	
12	Public administration/defense, education, public health, and other services	
13	Manufacturing of IT equipment	Terminals and systems added
14	IT services	Information services and CATV services added

Note: The Korean Standard Industrial Classification Table, consisting of 404 categories in 2000, was revised in 2005, to include 403 categories.

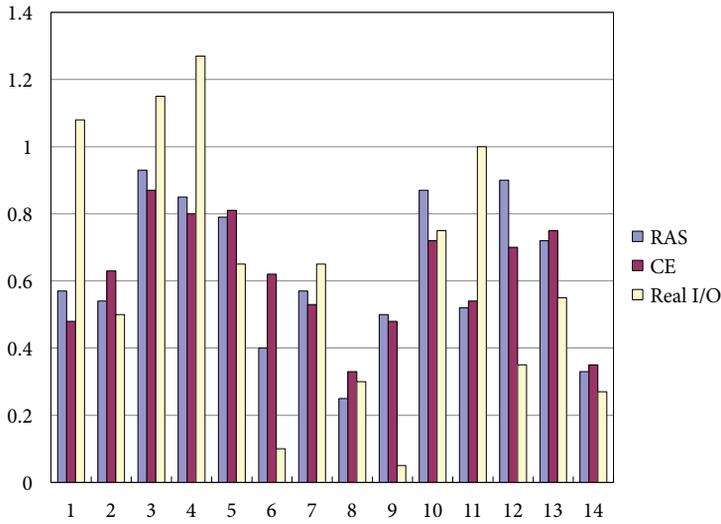


Fig. 2. Comparison of the actual and estimated sums of input coefficients of 14 Korean industries

updating accuracy based on the data in Fig. 2 concerning ‘Manufacturing of IT devices’, when the criteria of comparison proposed by Oosterhaven (2005) are applied, as shown in Fig. 3, the CE technique appears superior to the RAS technique in updating accuracy.

Here, JM is calculated by $IG = \sum_j^{14} q_j \ln(q_j / a_j)$, and JO by $JO = \sum_{j=1}^{14} |q_i| \ln(q_j / a_j)$. The smaller the value, the more efficient the estimation method is deemed. It should be noted, however, that the results of this analysis, comparing the relative accuracy of the two methods in estimating input coefficients for the year 2005 from the 2000 IO data by checking them against the actual coefficients for 2005, may be limited in their generalizability and may not prove valid for other years or other industry classification systems.

One of the main problems in updating IO tables is, as has just been pointed out, the difficulty of precisely determining, which method is better and more accurate than others, because one method can be highly accurate for one industry, while missing the mark with another industry. Therefore, the results of comparison can vary widely depending on the criteria used, particularly when a large number of updating methods are considered.

In this study, we employed one of the techniques designed to enable maximum use of information obtained from updating methods. To resolve the problem posed by estimates that are variably accurate depending on the industry, we estimated the following equation using OLS and calculated the combined input coefficient, which is the average of all the estimated coefficients: For a discussion on the efficiency of an OLS averaging estimator using the Mallows criterion, refer to (Hansen 2007). The simplest method for combining the two methods for estimating input coefficients is using OLS regression having two simple variables as the mediators. However, in this study, we deemed the above-described method to be

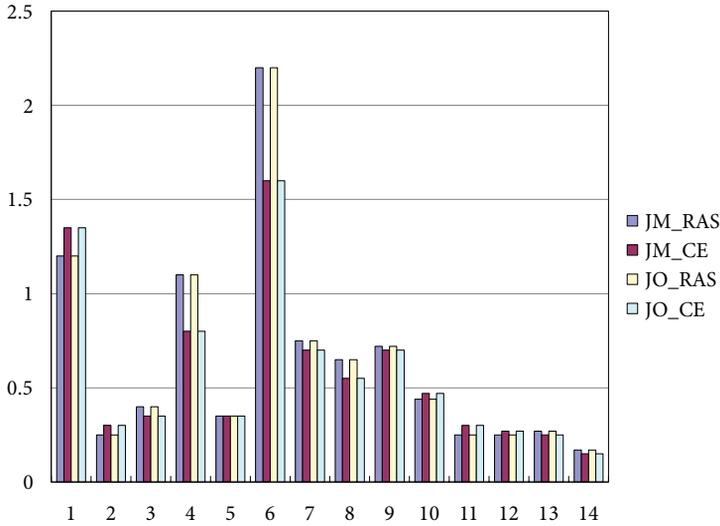


Fig. 3. Comparison of the two estimation methods from the perspective of information acquisition

superior. More importantly, the reviewer requested that the data satisfy homoscedasticity, an assumption guaranteeing efficiency in applying Mallows' criterion. Even if the data fails to satisfy this assumption, the IO table can be, nevertheless, updated using the Jackknife model average; a nonparametric method.

$$y_i = \sum_{j=1}^k \theta_j x_{ji} + b_{mi} + e_i ; \tag{11}$$

$$\hat{\omega} = \arg \min C_n(\omega) . \tag{12}$$

Here, y_i is the sum of the actual input coefficients of IT sectors (sum of the 2005 input coefficients for IT), and x_{ji} is the sum of the IT input coefficients updated using the RAS and CE methods. Meanwhile, b_{mi} stands for the relative error in an approximate value, obtained by averaging the two estimates, and $\hat{\omega}$ is an equation that determines, using the Mallows criterion, the optimal weights for the two updating methods, minimizing the estimation errors. $C_n(\omega)$ is the quadratic form of the error term (e) of least square averaging, which takes into consideration the weight (ω).

The estimated coefficients and weights obtained using the two above equations are given in Table 2 below. Table 2 shows that, in the case of the IT device manufacturing and IT service sectors, the weights assigned to minimize the error resulting from the new input coefficient are significantly larger for CE than for RAS. The estimated average coefficients, calculated through the OLS averaging estimator and applying the above-discussed weights according to Mallows information criterion, proved larger for CE than RAS for both of the IT industry sectors.

Table 2. Estimates of the IT device manufacturing and IT service sectors

Estimates		IT device manufacturing		IT services	
		RAS	CE	RAS	CE
Estimates	Estimated coefficients	0.0198	0.6309	0.0148	0.5951
	Weights	0.0064	0.9935	0.0965	0.9034

Figures 4 and 5 below show \hat{y} , the input coefficients estimated using the results listed in Table 2, corresponding to the updated input coefficients for the IT device manufacturing and IT service sectors. The figures also show the actual input coefficients for 2005, and the input coefficients estimated using the CE methods for the same year. The following implications are derived from the results in Figs 4 and 5. Firstly, the estimated input coefficients for IT device manufacturing (Industry class 13) and IT services (Industry class 14) are very close to the actual input coefficients. Secondly, given that the sum of the input coefficients is used to calculate production inducement, value-added inducement, and employment inducement effects from input data from an IO table, the input coefficient of a given industry is much more important than other input coefficients capturing indirect effects on the rest of the industries. This is why, when calculating an average OLS estimate using the Mallows criterion, a larger weight is assigned to the target industry than the remaining industries. Therefore, updating an IO table using an OLS averaging estimate, calculated with the use of the Mallows criterion proved to be highly efficient and accurate.

To recap the results of analysis thus far discussed, it is difficult to precisely determine which of the two methods for updating IO tables, namely, RAS and CE, is more accurate. When the two methods were applied to the fourteen Korean industries, we found that it was also difficult to judge their relative superiority concerning individual industries. Secondly,

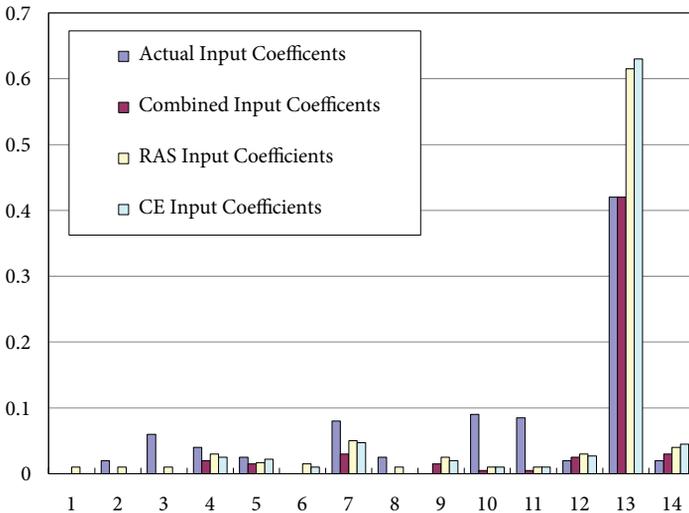


Fig. 4. Comparison of the combined estimated input coefficient and actual input coefficient for the Korean IT device manufacturing sector

we proposed an alternative method that combines the input coefficients estimated through the two updating techniques. An OLS approach was employed to average the RAS and CE estimates using the Mallows criterion. The estimates under the alternative approach proved to be more accurate, in other words, closer to the actual input coefficients for the target year, than those obtained using the RAS or CE technique, demonstrating the validity of the new approach. In other words, using both the RAS and CE techniques and combining the estimates under these two methods according to the alternative technique proposed in this study can effectively help improve accuracy in updating an IO table.

These results are the results of updating input coefficients for only the IT sector. When the input coefficients were updated for all 14 sectors, the outcome was similar; namely, the input coefficient of the combination of the 14 sectors was higher, even if the difference was quite small. As evidence to this effect, we attach the results of calculating optimal weights for the input coefficients of the 14 sectors of the Korean industry, using Mallows' criterion. Another noteworthy detail about the IO table updating method proposed in this study is that it provides the basis for determining which of the methods should be applied a greater weight than the other in order to guarantee that the prediction error is minimal. For example, in Table 3, the differences, visually examined, between the three updating methods are quite minor. But, the results, nevertheless, show clearly which of the two methods, RAS and cross entropy, should be applied a greater weight when updating the input coefficients.

Table 3. Estimated coefficients and optimal weights for 14 sectors

Methods	Input Coefficients of All 14 Sectors Updated	
	RAS	CE
Estimated Coefficients	0.000341	0.99956609
Weights	0.00000029	0.99999971

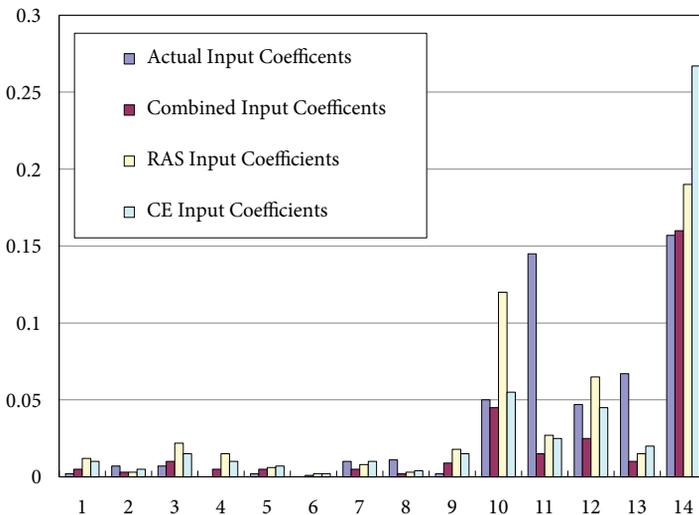


Fig. 5. Comparison of the combined estimated input coefficient and actual input coefficient for the Korean IT service sector

Conclusions

This study compared the two methods that are most widely employed for predicting future economic effects of an industry or a policy undertaking or trend, using Korea's 2000 and 2005 IO tables. Assessing the economic effects of a new policy undertaking or the potential economic impact of an industry on the rest of the industry sectors is usually done using the data from IO tables. The most delicate part of this task is that an IO table, pertaining to the current data, must be appropriately updated to project input-output relations at a point in time in the future. Therefore, an accurate method for updating an IO table is crucial for this process.

In this study, we compared the input coefficients of fourteen Korean industries, estimated using the RAS technique, which is widely used both by researchers and policy makers, and the CE technique, which has recently received much attention in theoretical discussions of IO updating. The main results are the following. First of all, it is not easy to precisely determine which of the two methods is superior based only on the comparison of input coefficients estimates. The relative degree of accuracy of the two methods varies widely depending on the industry, and even more so when the estimation involves a large number of industries. Both the RAS and CE estimates on the fourteen Korean industries were variably accurate, thereby not permitting a judgment in favor of either of them. As an alternative to these two methods, which are as equally valid as they are flawed, we proposed another method that appropriately combines input coefficients updated using the two methods. In this study, we used average OLS estimates calculated from the Mallows criterion. The input coefficients estimated using the alternative procedure proved to be closer to the actual input coefficients than the estimates under the RAS or CE approach, favorably demonstrating its accuracy as an updating technique. This method also has the advantage that it is based on the two traditional methods, rather than replacing them, and that it combines the estimates resulting from them, which leads to greater accuracy in updating an IO table.

As a general rule, in order to make a meaningful use of an updated IO table for IT-related industries, the amount of information that may have an impact on the relative superiority of the RAS or CE technique must be controlled in advance by imposing a restrictive condition. This suggests that, to ensure the best updating results, the estimation must be performed using preliminary information available about a target industry. However, this is both a time- and cost-intensive procedure. Therefore, an alternative method that allows estimates resulting from the two methods to be appropriately complemented could be a more practical solution. The simplest way of doing this is combining the results under the RAS and CE techniques using a traditional OLS approach. In addition to general implications, the combined updating method is applied to investigate the predictive power and accuracy of I/O updating table with those of RAS and CE methods for real high-technology industry sectors.

Directions for future research expanding on this study are as follows. The choice of an updating method for an IO table depends to a high degree on how the industry is classified and the degree of consolidation between industry classes. Therefore, when classifying a national industry, future research may improve the validity of results by carefully choosing the number of industry sectors. Also, since both the RAS and CE methods have their own

disadvantages that undercut their respective advantages, an optimal method for combining estimates from these two methods remains an essential topic in IO updating research. Therefore, future research should attempt to develop a more comprehensive econometric method for achieving this goal.

References

- Acemoglu, D.; Carvalho, V.; Ozdaglar, A.; Tahbaz-Salehi, A. 2011. The network origins of aggregate fluctuations, *Econometrica* 80: 1977–2016.
- Ahmed, S. A.; Preckel, P. V. 2007. *A comparison of RAS and entropy methods in updating IO tables*, Purdue University, Working Paper.
- Atalay, E.; Hortacsu, A.; Robert, J.; Syverson, C. 2011. Network structure of production, *PNAS* 108: 5199–5202. <http://dx.doi.org/10.1073/pnas.1015564108>
- Bank of Korea. 2003. *Input-output tables of the Korean economy for 2000*. Bank of Korea Publishing Company.
- Bank of Korea. 2008. *Input-output tables of the Korean economy for 2005*. Bank of Korea Publishing Company.
- Bank of Korea. 2009. *Input-output tables of the Korean economy for 2007*. Bank of Korea Publishing Company.
- Bank of Korea. 2010. *Input-output tables of the Korean economy for 2008*. Bank of Korea Publishing Company.
- Bank of Korea. 2011. *Input-output tables of the Korean economy for 2009*. Bank of Korea Publishing Company.
- Elliott, J.; Foster, I.; Kortum, S.; Munson, T.; Cervantes, F.; Weisbach, D. 2011. Trade and carbon taxes, *American Economic Review; Paper & Proceedings* 100: 465–469. <http://dx.doi.org/10.1257/aer.100.2.465>
- Geweke, J.; Amisano, G. 2011. Optimal prediction pools, *Journal of Econometrics* 164: 130–141. <http://dx.doi.org/10.1016/j.jeconom.2011.02.017>
- Geweke, J. 2010. *Complete and incomplete econometrics and statistics*. Wiley, Hoboken.
- Gilchrist, D. A.; St Louis, L. V. 1999. Completing input-output tables using partial information with an application to Canadian data, *Economic Systems Research* 11: 185–193. <http://dx.doi.org/10.1080/09535319900000013>
- Golan, A.; Judge, G.; Robinson, S. 1994. Recovering information from incomplete or partial multi-sectoral economic data, *Review of Economics and Statistics* 76: 541–549. <http://dx.doi.org/10.2307/2109978>
- Golan, A.; Judge, G.; Miller, D. 1996. *Maximum entropy econometrics*. John Wiley & Sons, Ltd.
- Hansen, B. 2007. Notes and comments: least squares model averaging, *Econometrica* 75: 1175–1189. <http://dx.doi.org/10.1111/j.1468-0262.2007.00785.x>
- Huang, W.; Kobayashi, S.; Tanji, H. 2008. Updating an input-output matrix with sign-preservation, *Economic System Research* 20: 111–123. <http://dx.doi.org/10.1080/09535310801892082>
- Israilevich, P. R.; Hewings, G. J. D.; Sonis, M.; Schindler, G. R. 1997. Forecasting structural change with a regional econometric input-output model, *Journal of Regional Science* 37: 565–590. <http://dx.doi.org/10.1111/0022-4146.00070>
- Jackson, R. W.; Murray, A. T. 2004. Alternative input-output matrix updating formulations, *Economic Systems Research* 16: 135–148. <http://dx.doi.org/10.1080/0953531042000219268>

- Jones, C. 2011. Intermediate goods and weak links in the theory of economic development, *American Economic Journal: Macroeconomics* 3: 1–28. <http://dx.doi.org/10.1257/mac.3.2.1>
- Junius, T.; Oosterhaven, J. 2003. The solution of updating or regionalizing a matrix with both positive and negative entries, *Economic Systems Research* 15: 87–96. <http://dx.doi.org/10.1080/0953531032000056954>
- Kim, K. H. 2000. *Saneop yeongwan bunseongnon* [Industrial input-output analysis]. Seoul: Yeonamsa Publishing Company.
- McDougall, R. 1999. *Entropy theory and RAS are friends*, GTAP Working Paper.
- Okuyama, Y.; Hewings, G. J. D.; Sonis, M.; Israilevich, P. R. 2002. An economic analysis of biproportional properties in an input-output system, *Journal of Regional Science* 42: 361–388. <http://dx.doi.org/10.1111/1467-9787.00263>
- Oosterhaven, J. 2005. GRAS versus minimizing absolute and squared differences: a comment, *Economic Systems Research* 17: 327–331. <http://dx.doi.org/10.1080/09535310500221864>
- Rapach, D.; Strauss, J. 2010. Bagging or combining (or both)? An analysis based on forecasting U.S. employment growth, *Econometric Reviews* 29: 511–533. <http://www.tandfonline.com/doi/abs/10.1080/07474938.2010.481550>
- Rim, M.-H.; Cho, S.-S.; Moon, C.-G. 2004. Measuring economic externalities of IT and R&D, *ETRI Journal* 27: 206–218.
- Robinson, S.; Cattaneo, A.; El-Said, M. 2001. Updating and estimating a social accounting matrix using cross entropy methods, *Economic Systems Research* 13: 47–64. <http://dx.doi.org/10.1080/09535310120026247>
- Shannon, C. 1948. A mathematical theory of communication, *Bell System Technical Journal* 27: 379–423.
- Smith, J.; Wallis, K. 2009. A simple explanation of the forecast combination puzzle, *Oxford Bulletin of Economics and Statistics* 71: 302–355. <http://dx.doi.org/10.1111/j.1468-0084.2008.00541.x>
- Stock, J.; Watson, M. 2004. Combination forecasts of output growth in a seven-country data set, *Journal of Forecasting* 23: 405–430. <http://dx.doi.org/10.1002/for.928>
- Stone, R. 1961. *Input-output accounts and national accounts*. OECD, Paris, France.
- Xing, W.; Ye, X.; Kui, L. 2011. Measuring convergence of China's ICT industry: an input-output analysis, *Telecommunications Policy* 35: 301–313. <http://dx.doi.org/10.1016/j.telpol.2011.02.003>

Byung-Sun CHO obtained his PhD in Economics from the University of Kansas in the USA. He has served as a Senior Researcher at ETRI (Electronics and Telecommunications Research Institute) for 13 years. His main research areas were technology policy and forecasting, especially in IT fields. Nowadays, his research interests are: future technology strategy for national R&D and technology forecasting linked technology with markets.

Sang Sup CHO received his BA degree in economics from Hanyang University, Seoul, Korea in 1985, the MA degree in economics from Hanyang University, Seoul, Korea in 1987, the MS degree in accounting from Missouri, Kansas City, USA, in 1996, and the PhD degree in economics from Saint Louis University, USA, in 1999. His research interests range from non-stationary panel analysis in econometrics to applied macroeconomic analysis, high-tech industries growth and cycle analysis.

Jungmann LEE obtained his BA degree from Korea University, Seoul, Korea in 1986 and PhD degree in Economics from the City University of New York, USA, in 1997. He is an associate professor at the Division of Digital Business at Hoseo University. He has served as an advisor for various projects (IT technology policy and HRD Policy) of the Ministry of Information and Communication, Korea. His research interests are focused on the areas of technology policy, R&D management, and the economics of technology innovation at the Electronics and Telecommunications Research Institute.