DOI: 10.20472/IAC.2018.040.013

**TZU-CHIANG CHIANG**

**Department Of Information Management, Tunghai University, Taiwan**

# ANALYSIS ON THE INFLUENCE OF GOVERNMENT EFFICIENCY ON TAIWAN'S SOCIETY, ECONOMY, AND ENVIRONMENT USING DATA MINING: A CASE STUDY ON OPEN GOVERNMENT DATA

## Abstract:

This study uses DEA-SBM to obtain the target variables and the decision tree model to produce important indicators. The government can promote Taiwan's development in all aspects by improving social, economic, and environmental efficiency, while ensuring the sustainable development of environmental protection. The 15 social, economic, and environmental variables are used in the decision tree to obtain 3 repeatable important indicators that can affect Taiwan's social change, economic development, operational capability, environmental pollution, and change. K-means divides the 22 counties and cities into 1 capital city, 5 municipalities, and 16 other counties and cities, in order to understand the influence of the economy on social and environmental development during the period of 2013-2015 in Taiwan.

## Keywords:

Open Government Data, k-means, DEA-SBM, Decision tree, Data Mining

**JEL Classification:** C89

# 1 Introduction

This study selects open government data from 2013 to 2015 for data cleaning, and uses DEA-SBM to analyze the inputs and outputs of society, economy, and environment, and identify the target variables; Next, the decision tree is used to analyze the target variables and determine the key data fields, in order to understand the relationship between the variables and government policies. Finally, this study evaluates the effectiveness of government policies and reforms, economic benefits, and environment destruction brought by industrial and commercial enterprises, with the aim to effectively utilize the outputs and inputs of the 3 dimensions, such as environment, society and economy through DEA-SBM.

# 2 Literature Review

Sustainable development has been applied to many fields, such as agriculture, economy, ecology, and communities. As the fields are different, the views regarding their dimensions of sustainable development are different. Some experts consider that sustainable development involves three dimensions, society, economy, and environment. Munasinghe (1993) analyzed the experiences and processes of the economic development of various states, and developed a sustainable integrated framework (Liao, 2004). Munasinghe also considered that the development of a country shall have 3 main goals: economic growth, social equity, and ecological stability and coordination, which form 3 different frameworks. The correlation between economic growth and social equity lies in equity within generations, i.e. income distribution and social class in the same time period.

Equal distribution within generations can be promoted by eliminating poverty and increasing employment; on the other hand, the correlation between the 2 goals of economic growth and ecological stability coordination lies in the assessment of the influence of economic policies on the environment, internalization, and external costs, that is reducing the load of resource utilization on ecological environments. Finally, the correlation between the 2 goals of social equity and ecological stability coordination lies in the equality of generations, meaning that the rights to enjoy and use resources and the environment shall be reserved for following generations. In short, sustainable development relies on economy, society, and environment to improve our environment and quality of life. Holistic government can establish a window service mechanism through information technology, and use people's life events to reorganize government business divisions to provide 1 integrated service, rather than separate departments and units. Thus, this study aims to supervise government's sustainable operation capabilities in combination with holistic governance and open government data.

DEA uses Pareto Optimality as the main concept to judge efficiency, and all the major data of the decision making units (DMU) are used as input and output. This study uses

operation capability as the input and profitability as the output, and DEA is used to calculate values with significant influence in terms of the variables of the financial statements of companies. The DEA input-output concept originated from the Deterministic Non-parametric Approach, as developed by Farrel, which has been incorporated into Mathematical Programming. An efficiency frontier is established for the evaluated DMUs. The DMUs that fall on the frontier are the most efficient, while those that fail to fall on the frontier have no efficiency, thus, efficiency is a concept of relative efficiency.

## 3 Analysis design of government operation capability

Local open government data are used first, and 15 variables related to social, economic, and environmental issues between 2013 and 2015 are selected; the data are divided into inputs and outputs for DEA, in order to calculate the target variables of government efficiency, and then, the decision tree and random forest models are used to obtain the indicators that affect governmental efficiency. The 22 counties/cities are clustered to observe the correlation between the counties and cities; finally, the results of this study are analyzed to facilitate cities with lower efficiency learning from cities with better efficiency.

### Description of main variables

This study examines the production efficiency of local governments, as based on the 3 dimensions of society, economy, and environment. The 15 variables are described in Table 3-1:

**Table 3-1 Description of 15 variables**

| Dimension | Category | Definition |
|---|---|---|
| Social | Population | All citizens registered in a region, whether or not they live in the region, shall be included in the regional population |
| | Welfare of the disadvantaged | Total: long-term care of the elderly, actual number of people in support institutions, children and adolescents in placement and correctional institutions, people receiving emergency family support, and the physically and mentally disabled |

| | | |
|---|---|---|
| | Criminal cases | Number of cases accepted at complaint, prosecution, surrender, or service |
| | Employed population | Number of paid workers aged 15 years old or unpaid family workers who work for more than 15 hours (1000 persons) |
| Economic | Average disposable income per household | Average disposable income per household (NTD) |
| | Sales of profit-seeking business | Declared or approved sales of profit-seeking businesses in accordance with the relevant provisions (thousand NTD) |
| | Annual expenditure | All expenditures in 1 fiscal year (million NTD) |
| | Average tax per capita | Average net tax imposed upon one person<br>Formula: actually imposed net tax/number of persons per year (NTD) |
| Environment | Concentration of PM | Particles suspended in the air (including particles with a size of above 10 microns). The PM concentration of each area is the arithmetic mean from the stations in the area (value obtained from continuous 24 hours monitoring). Before 1992, the data whose monitoring months are less than 8 are not included in the calculation. After 1993, the data whose monitoring frequency is less than 16 times are not included in the calculation. |
| | Volume of cleaned and transported garbage | Volume (MT) of waste transported by waste disposal authorities or a company entrusted by waste disposal authorities to the waste disposal field. |
| | Penalty of | Penalty that are imposed by environment protection units |

| | | |
|---|---|---|
| | industrial wastewater | during inspection upon livestock wastewater, other industrial wastewater, and violations against legal regulations. |
| | Average environment expenses per capita | Average environmental protection expenses per capita: annual expenditure budget and final accounts of all levels of environmental protection units, as provided by the Environment Protection Administration of the Executive Yuan. |
| | Sales of electric power | Lump-sum electric power, LV, HV, and UHV power sold by Taiwan Power Company for users (megawatt-hour). |
| | Land area | Land area of the administrative regions, including reclaimed land and affiliated islands ($km^2$). |
| | Planned city area | Total area of planned cities ($km^2$) that conform to the statutory procedures |

**Design of target variables**

The open data are downloaded from the Directorate General of Budget, Accounting and Statistics, which are classified into social, economic, and environmental categories, and cleaned. Only the data of the counties/cities between 2013 and 2015 are selected, which are converted into input, output, and undesired output, as shown in Tables 3-2 ~ 3-4.

**Table 3-2 Descriptive statistics of input (2013-2015)**

| **Input** | Employed population (1000 persons) | Annual expenditures (million NTD) | Sales of electric power (megawatt-hour) |
|---|---|---|---|
| **Maximum value** | 1,945.000 | 169,580.240 | 22,535.000 |
| **Minimum value** | 3.000 | 3,000.850 | 26.000 |
| **Average value** | 504.697 | 45,293.541 | 6,551.621 |
| **Standard deviation** | 520.909 | 47,184.288 | 7,314.824 |

**Table 3-3 Descriptive statistics of output (2013-2015)**

| **Output** | Welfare of the disadvantaged (persons) | Sales of profit-seeking business (1000 NTD) | Average disposable income per household (NTD) |
|---|---|---|---|
| **Maximum value** | 171,193.000 | 12,810,834,067.000 | 1,314,031.000 |
| **Minimum value** | 451.000 | 4,318,736.000 | 636,162.000 |
| **Average value** | 54,580.470 | 1,786,163,440.848 | 894,040.667 |
| **Standard deviation** | 47,642.092 | 2,745,399,916.730 | 163,718.021 |

**Table 3-4 Descriptive statistics of undesired output (2013-2015)**

| **Undesired output** | Concentration of PM | Volume of garbage (metric ton) | Penalty of industrial wastewater | Criminal case (number) |
|---|---|---|---|---|
| **Maximum value** | 88.660 | 427,356.000 | 149,962.000 | 54,387.000 |
| **Minimum value** | 29.700 | 2,112.000 | 19.000 | 58.000 |
| **Average value** | 56.628 | 148624.364 | 13,484.621 | 13,379.167 |
| **Standard deviation** | 13.045 | 130,258.295 | 22,582.567 | 13,909.368 |

This study uses the DEA-SBM model to analyze the target input and output variables of the 3 years; namely, the less the input, the more the output, and the higher the efficiency. Regarding the target variables of undesired output, the corrected model of non-demand output, as proposed by Cooper et al., is used.

The non-demand output model is divided into 2 types: the first type is the undesired output model, where the output is divided into desired output and undesired output, and

there is no relationship between them; the output model is divided into 2 types, the one type is the undesired output model, and the output is divided into desired output and undesired output, and there is no relationship between them. The second type is the non-split model: namely, desired output and undesired output cannot be split. This study selects the split undesired output model.

The undesired output model evaluation system has n SBMs, where each SBM has 3 input and output factors: input, desired output, and undesired output, which are represented by 3 vectors, respectively: $X \in R^m, Y^g \in R^{s1}, Y^b \in R^{s2}$; the three vector matrices X, $Y^g, Y^b$ are defined as: X=$[x_1, \dots, x_n] \in R^{m\times n}$, $Y^g$ =$[y_1^g, \dots, y_n^g] \in R^{S1\times n}$, $Y^b$=$[y_1^b, \dots, y_n^b] \in R^{S2\times n}$, and it is assumed that if X>0, $Y^g > 0, Y^b > 0$, the production possibility set can be defined as:

$$P = \{(x, y^g, y^b) | x \geq X\lambda,\ y^g \leq y^g\lambda, y^b \leq y^b\lambda, \lambda \geq 0\}, \qquad (3.1)$$

According to the definition of Eq. (3.1), the corrected SBM model can be presented in the following equation:

$$\min p^* = \frac{1-\frac{1}{m}\sum_{i=1}^{m}\frac{s_i^{input}}{x_{io}}}{1+\frac{1}{s_1+s_2}\left(\sum_{r=1}^{s_1}\frac{s_r^{output}}{y_{ro}^g}+\sum_{r=1}^{s_2}\frac{s_r^{undesirbed\ output}}{y_{ro}^b}\right)}$$

$$\text{s.t.} \quad x_o = X\lambda + s^{input} \qquad (3.2)$$

$$y_o^g = Y^g\lambda - s^{output}$$

$$y_o^b = Y^b\lambda + s^{undesired\ output}$$

$$\lambda, s^-, s^g, s^b \geq 0$$

In Eq. (3.2), $s^{input}$ is excess input quantity, $s^{undesirbed\ output}$ is excess undesired output quantity, and both shall be reduced; $s^{output}$ is a shortage of desired output, and shall be increased. $0 < p^* \leq 1$, let the above linear programming expression have the optimum value ($\lambda^*, s^{input*}, s^{undesirbed\ output*}, s^{output*}$), then $s^{input*} = 0, s^{undesirbed\ output*}$=0, $s^{output*}$=0, $p^* = 1$, and this means $DMU_o$ in the undesired output model has efficiency.

If $p^* < 1$, $DMU_O$ in the undesired output model has inefficient SBM, thus, in order to ensure that SBM is effective, excessive quantities of input and undesired output must be reduced, while the shortage of described output shall be increased. Eq. (3.3) is the undesired output model, which has no efficient SBM adjustment model, and $(\widehat{x_0}, y_0^{\widehat{output}}, y_0^{\widehat{undesired\ output}})$ represents the projection points of an efficiency frontier.

$$\widehat{x_0} = x_0 - s^{input^*}$$

$$y_0^{\widehat{output}} = y_0^{output} + s^{output^*} \quad (3.3)$$

$$y_0^{\widehat{undesirbed\ output}} = y_0^{undesirbed\ output} - s^{undesired\ output^*}$$

The SBM model can produce the government efficiency indicator of this study.

### Analysis of decision tree

This study uses the R rpart package; rpart is CART classification and regression tree. The binary tree is repeatedly established from root through iteration till the isomorphism type of the tree nodes reach a certain standard or the operation termination conditions are triggered. The decision tree uses the efficiency indicators of SBM as the target variables, in order to obtain important pointers and achieve model accuracy.

## 4 Analysis results of government operation capability

### Analysis results of DEA

The study uses data of 2013-2015 for the DEA-SBM model. The input variables are employed population, annual expenditures, and electric power sales; the output variables include the welfare of the disadvantaged, sales of profit-seeking businesses, average disposable income per household; while undesired outputs include the concentration of PM, volume of garbage cleaned and transported, and penalties for industrial wastewater and criminal cases.

As shown in Table 4-1, there are 9 counties/cities with DEA-SBM indicators of 1 (the highest efficiency) in 2013; the other 13 counties/cities can improve their input and output; the efficiency indicator of Taichung City is the lowest.

**Table 4-2 DEA-SBM scores in 2013**

| Counties/cities | Score | Rank | 1/Score | Counties /cities | Score | Rank | 1/Score |
|---|---|---|---|---|---|---|---|
| Taipei City | 1 | 1 | 1 | Changhua County | 0.9819397 | 10 | 1.0183924 |
| Yunlin County | 1 | 1 | 1 | Pingtung County | 0.9802676 | 11 | 1.0201296 |
| Chiayi County | 1 | 1 | 1 | Keelung City | 0.9745176 | 12 | 1.0261487 |
| Taitung County | 1 | 1 | 1 | New Taipei City | 0.9627735 | 13 | 1.0386659 |
| Hualien County | 1 | 1 | 1 | Taoyuan City | 0.9575404 | 14 | 1.0443423 |
| Hsinchu City | 1 | 1 | 1 | Yilan County | 0.9537752 | 15 | 1.0484651 |
| Chiayi City | 1 | 1 | 1 | Nantou County | 0.9311214 | 16 | 1.0739738 |
| Kinmen County | 1 | 1 | 1 | Hsinchu County | 0.8770389 | 17 | 1.1402003 |
| Lienchiang County | 1 | 1 | 1 | Tainan City | 0.8515812 | 18 | 1.1742861 |
| | | | | Kaohsiung City | 0.8185517 | 19 | 1.22167 |
| | | | | Penghu County | 0.7920523 | 20 | 1.2625429 |
| | | | | Miaoli County | 0.7795309 | 21 | 1.2828229 |
| | | | | Taichung City | 0.7170804 | 22 | 1.3945437 |

Taichung City has the lowest efficiency, meaning it is deficient in the welfare of the disadvantaged, sales of profit-seeking businesses, and average disposable income per household; thus, the inputs, such as employed population, and the outputs, such as annual expenditures and the sale of electric power, are not adequate; while the undesired outputs, such as penalty for wastewater, must be reduced. (For other counties and cities, see Appendix).

**Table 4-3 Scores of input and output of Taichung City in 2013**

| SBM | 1/DEA_Score | | | |
|---|---|---|---|---|
| I/O (Input/output) | Original data | Relative value | Difference | Difference (%) |
| Taichung City | 1.394543748 | | | |
| Employed population | 1,276 | 1,276 | 0 | 0.00% |
| Annual expenditure | 109,538.02 | 109,538.02 | 0 | 0.00% |
| Sales of electric power | 17,505 | 17,505 | 0 | 0.00% |
| Welfare of the disadvantaged | 119,908 | 167,216.9517 | 47,308.95169 | 39.45% |
| Sales of profit-seeking business | 3,687,509,962 | 5,142,393,962 | 1,454,884,000 | 39.45% |
| Average disposable income per household | 928,938 | 7,589,818.795 | 6,660,880.795 | 717.04% |
| Concentration of PM | 78.52 | 546.4220774 | 467.9020774 | 595.90% |
| Volume of garbage cleaned and transported | 378,209 | 457,434.2076 | 79,225.20762 | 20.95% |
| Penalty of industrial wastewater | 23,030 | 16,200.98869 | -6,829.011307 | -29.65% |

According to Table 4-3, there are 10 counties with a DEA-SBM indicator of 1 (highest efficiency) in 2014; the other 12 counties can improve their input and output; the efficiency indicator of Kaohsiung City is the lowest.

**Table 4-3 DEA-SBM scores of 2014**

| Counties/cities | Score | Rank | 1/Score | Counties /cities | Score | Rank | 1/Score |
|---|---|---|---|---|---|---|---|
| Taipei City | 1 | 1 | 1 | Pingtung County | 0.9839405 | 11 | 1.0163216 |
| Yunlin County | 1 | 1 | 1 | Tainan City | 0.9817848 | 12 | 1.0185531 |
| Chiayi County | 1 | 1 | 1 | Yilan County | 0.9807762 | 13 | 1.0196006 |
| Taitung County | 1 | 1 | 1 | Changhua County | 0.9616386 | 14 | 1.0398917 |
| Hualien County | 1 | 1 | 1 | Nantou County | 0.9330617 | 15 | 1.0717405 |
| Keelung City | 1 | 1 | 1 | New Taipei City | 0.8848463 | 16 | 1.1301397 |
| Hsinchu City | 1 | 1 | 1 | Taoyuan City | 0.8835447 | 17 | 1.1318046 |
| Chiayi City | 1 | 1 | 1 | Penghu County | 0.8279055 | 18 | 1.2078674 |
| Kinmen County | 1 | 1 | 1 | Miaoli County | 0.8076365 | 19 | 1.2381807 |
| Lienchiang County | 1 | 1 | 1 | Hsinchu County | 0.7724447 | 20 | 1.2945911 |
| | | | | Taichung City | 0.7535297 | 21 | 1.3270877 |
| | | | | Kaohsiung City | 0.7480282 | 22 | 1.3368479 |

As shown in Table 4-4, there are 14 counties/cities with a DEA-SBM indicator of 1 (highest efficiency) in 2015, meaning government efficiency is improving year by year; while the other 8 counties/cities can improve their input and output; the efficiency indicators of Taoyuan City are the lowest.

**Table 4-4 DEA-SBM scores of 2015**

| Counties/cities | Score | Rank | 1/Score | Counties /cities | Score | Rank | 1/Score |
|---|---|---|---|---|---|---|---|
| Taipei City | 1 | 1 | 1 | Pingtung County | 0.9780446 | 15 | 1.0224482 |
| Tainan City | 1 | 1 | 1 | New Taipei City | 0.9038394 | 16 | 1.1063913 |
| Yilan County | 1 | 1 | 1 | Penghu County | 0.8563788 | 17 | 1.1677076 |
| Changhua County | 1 | 1 | 1 | Hsinchu County | 0.8526078 | 18 | 1.1728722 |
| Nantou County | 1 | 1 | 1 | Miaoli County | 0.846423 | 19 | 1.1814424 |
| Yunlin County | 1 | 1 | 1 | Kaohsiung City | 0.7511832 | 20 | 1.3312331 |
| Chiayi County | 1 | 1 | 1 | Taichung City | 0.7303366 | 21 | 1.3692317 |
| Taitung County | 1 | 1 | 1 | Taoyuan City | 0.6859151 | 22 | 1.4579063 |
| Hualien County | 1 | 1 | 1 | | | | |
| Keelung City | 1 | 1 | 1 | | | | |
| Hsinchu City | 1 | 1 | 1 | | | | |
| Chiayi City | 1 | 1 | 1 | | | | |
| Kinmen | 1 | 1 | 1 | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| County | | | | | | | |
| Lienchiang County | 1 | 1 | 1 | | | | |

**Analysis results of decision tree**

This study uses the data between 2013 and 2015 for decision tree analysis, and analysis results show that penalties of wastewater, garbage volume, planned urban area, and population have high repeatability, which may affect government efficiency.

Table 4-8 Branches of the decision tree

| 2013 | 2014 | 2015 | 2013~2015 |
|---|---|---|---|
| Penalty of industrial wastewater | Garbage volume | Garbage volume | Garbage volume |
| Land area | Planned city area | Planned city area | Planned city area |
| Planned city area | Penalty of industrial wastewater | Penalty of industrial wastewater | Criminal case |
| Population | Population | Population | Power sales |
| Average disposable income per household | | | Sales of profit-seeking business |
| | | | Concentration of PM |

In 2013, the main branches include penalties for wastewater, land area, planned urban area, population, and average disposable income per household.

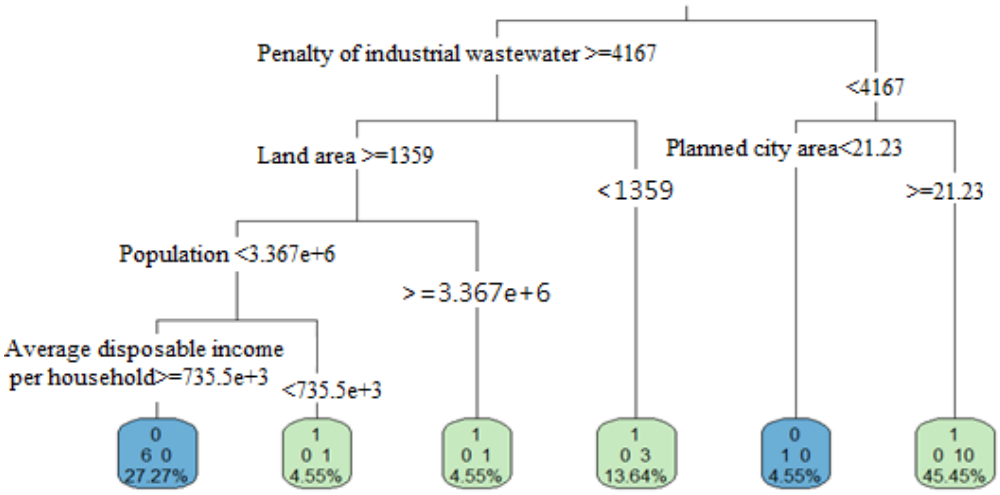


**Figure 4-5 Decision tree analysis chart of 2013**

The main branches in 2014 include garbage volume, planned urban area, penalties of industrial wastewater, and population.

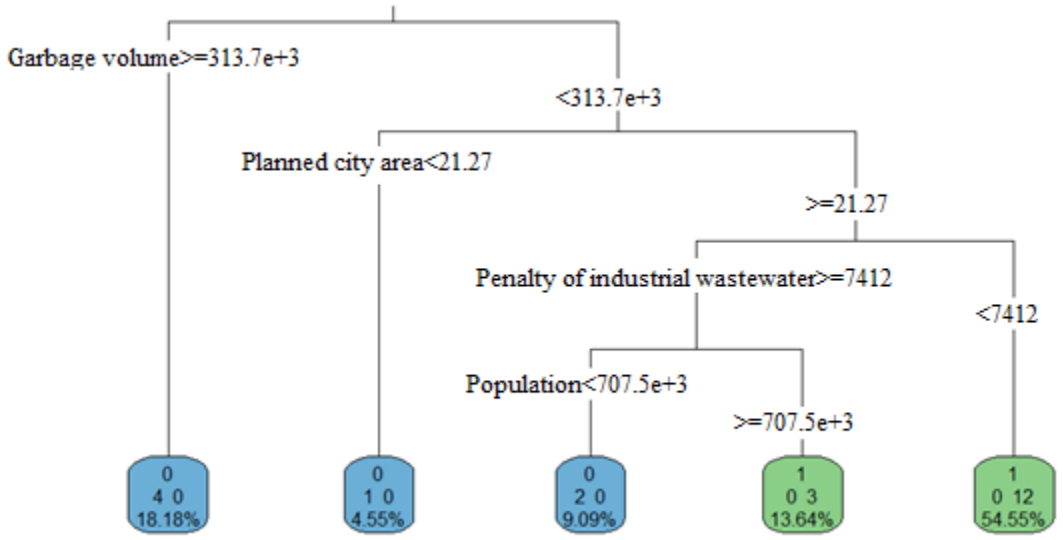


**Figure 4-6 Decision tree analysis chart of 2014**

The main branches in 2015 include garbage volume, planned urban area, penalties of industrial wastewater, and population.

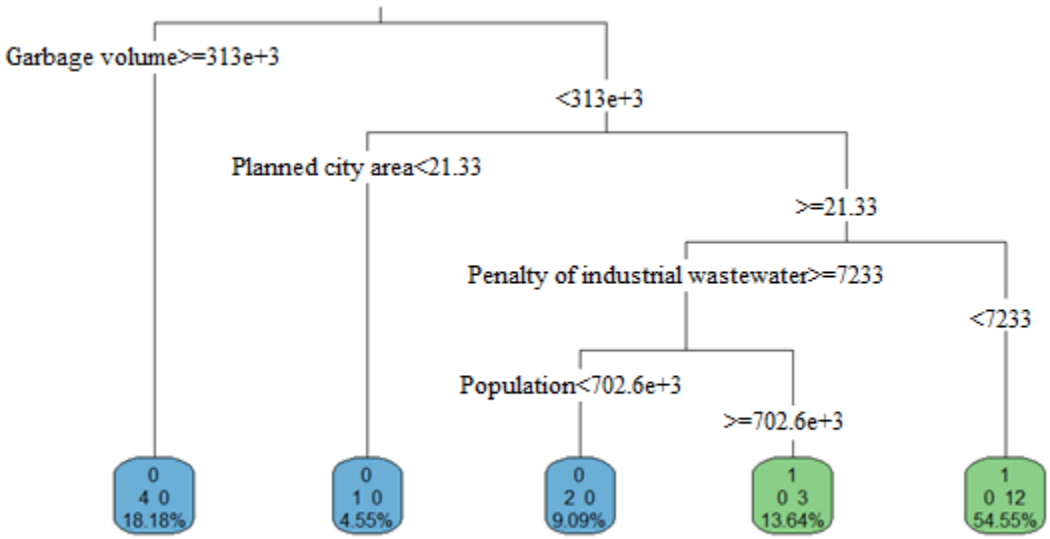


**Figure 4-7 Decision tree analysis chart of 2015**

Garbage volume is an important branch of the decision tree, and belongs to the top-most branch between 2014 and 2015. The garbage volume was less than 313,700 (metric tons) between 2014 and 2015, and this means that government efficiency has positive growth. As the government advocates environmental protection and recycling, the garbage volume has decreased year by year since 2001, as shown Figure 4-3.

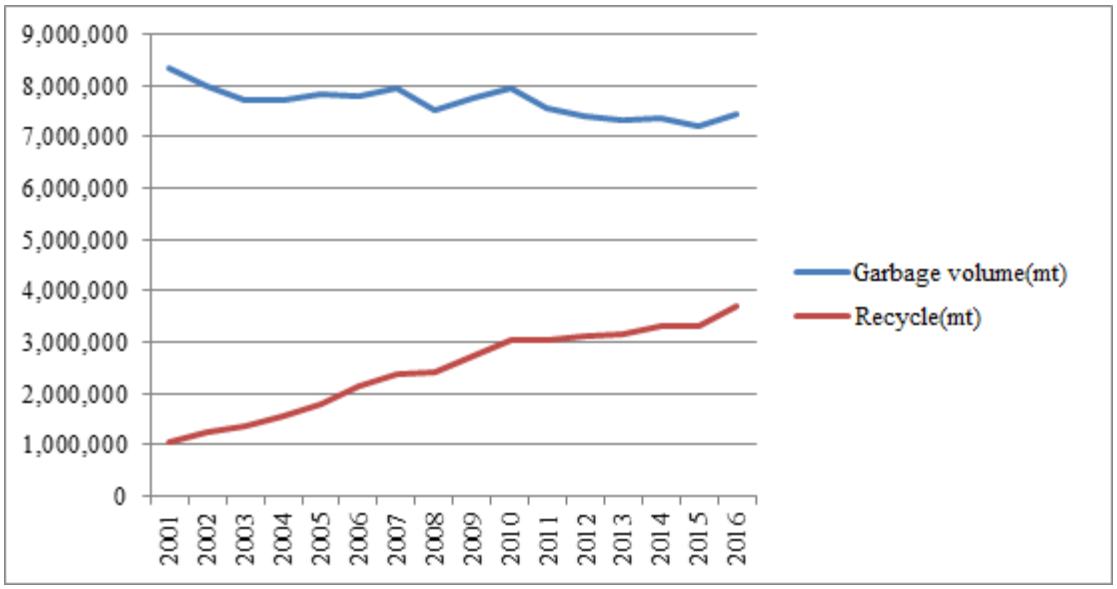


**Figure 4-3 Broken line chart for generated garbage volume and recovered resources between 2001 and 2016**

For example, a garbage volume of more than 313,700 (metric tons) is related to city size, population, and industrial development. In 2014 the garbage volumes of New Taipei City, Taoyuan City, Taichung City, and Kaohsiung City exceeded the standard. The average daily generated garbage volume per capita is 0.383 (kg) in 2014, meaning the annual garbage volume per capita is 0.383*365=139.795 (kg); Taoyuan City, Taichung City, and Kaohsiung City exceeded the average volume; while New Taipei City has a large population, and thus, did not exceed the average volume.

**Table 4-9 Predictive error of generated garbage volume**

| | Population | Predicted production | Garbage volume | Error |
|---|---|---|---|---|
| New Taipei City | 3,966,818.000 | 554,541.3223 | 388,898.000 | -165,643.322 |
| Taipei City | 2,702,315.000 | 377,770.1254 | 274,166.000 | -103,604.125 |
| Taoyuan City | 2,058,328.000 | 287,743.9628 | 353,137.000 | 65,393.037 |
| Taichung City | 2,719,835.000 | 380,219.3338 | 384,775.000 | 4,555.666 |
| Tainan City | 1,884,284.000 | 263,413.4818 | 256,680.000 | -6,733.482 |
| Kaohsiung City | 2,778,992.000 | 388,489.1866 | 408,745.000 | 20,255.813 |

The main branches of the decision tree between 2013 and 2015 are garbage volume, criminal cases, planned urban area, sales of electric power, sales of profit-seeking businesses, and PM. In this study, the data of the 22 counties/cities of the 3 years are divided into 70% validation sets and 30% test sets, and the accuracy rate of the model is 97.87%, which could almost predict the efficiency of government operation during the 3 years.

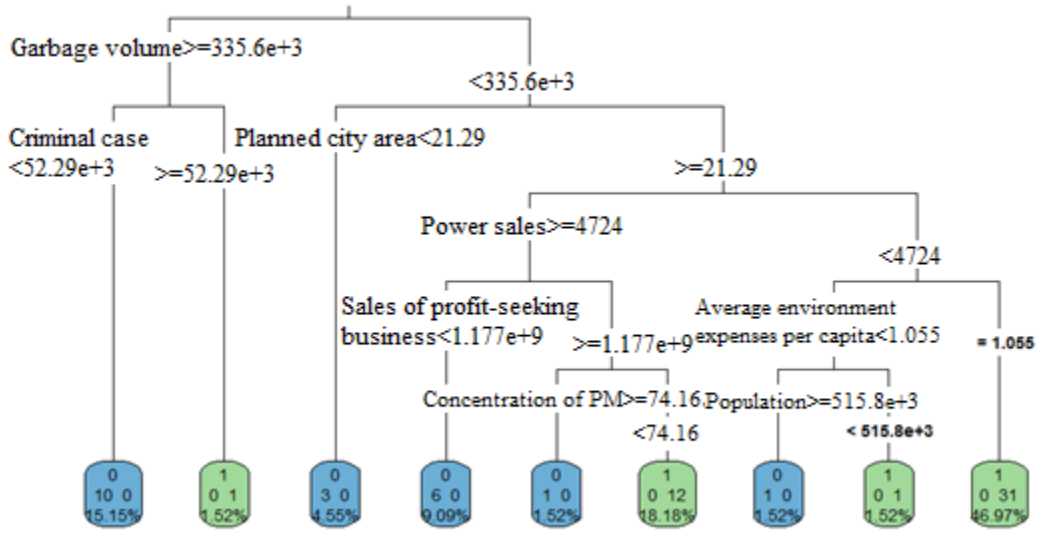


**Figure 4-4 Decision tree analysis chart of 2013-2015**

**Policies for improvement of government efficiency**

In order to improve the efficiency of government, many policies have been proposed, including waste disposal, industrial water recycling, wastewater treatment, birth allowance, and energy conservation. The outcomes of the implementation of these policies in the 6 cities are, as follows:

Waste disposal: after implementation of various polices, such as "Don't Throw That Trash on the Ground", "Mandatory Sorting and Recycling of Garbage", "Extended responsibility of manufacturers and source reduction", the restricted use of plastic bags and dishes, excessive packaging, and the waste volume per capita per day was reduced from 1.14kg in 1998 to 0.52kg in 2008, thus, the recycling rate of household wastes reached 42%, which can be compared to advanced countries.

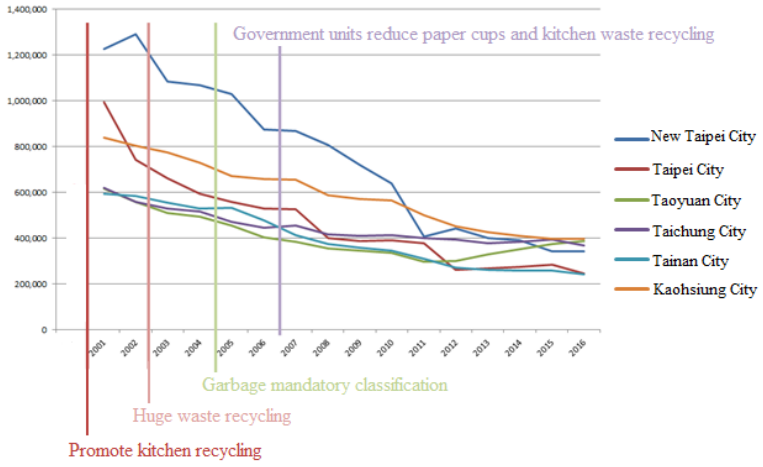


**Figure 4-5 Relation chart between the garbage policies of the 6 cities and time**

Reuse of water resources: the industrial wastewater pollution of Taoyuan City and Kaohsiung City exceeded the standard considerably, thus, the penalty is more, as shown in Figure 4-5. In order to effectively solve the problems of water pollution, the government should pay more attention to the pollution of waters, rivers, and seas, in order that relevant environment protecion policies shall be promoted.

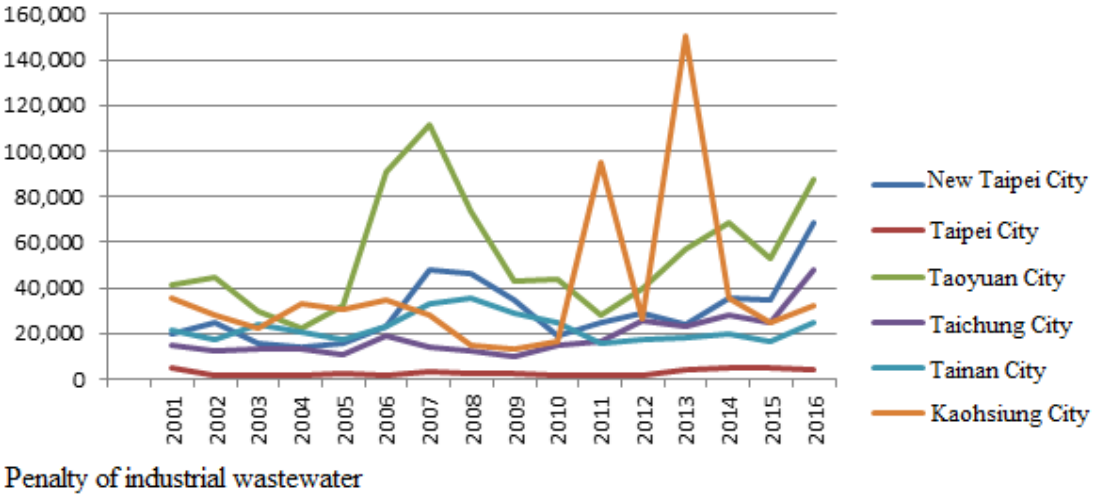


**Figure 4-6 Broken line chart for the industrial wastewater penalties of the 6 cities**

Birth allowance: In recent years, Taiwan's fertility rate continues to drop, the aging population increases, and the labor population declines, which has caused a reduction in national revenues. Due to such issues, people's tax burdens have become heavier and

heavier, the dependency ratio increases, and young people may face economic pressure. As a result, the Taiwan market shrinks, the economy cannot grow, and state finance records have a huge black hole. In order to increase the fertility rate, the government has proposed birth allowance measures, and hopes that such measures can increase the fertility rate, in order to solve the state financial crisis.

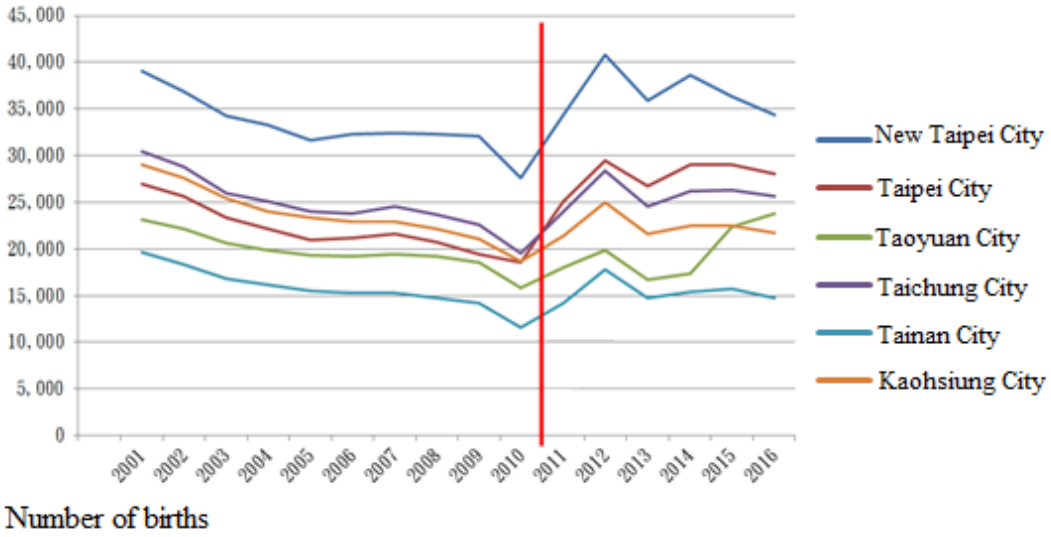


**Figure 4-7 Number of newborns of the 6 cities after implementation of the birth allowance policy**

## 5 Conclusion and Research Limitations

This study selects 15 social, economic, and environmental variables from the open government data between 2013 and 2015. DEA-SBM is used to analyze the input, undesired output, and output to obtain economic indicators. The decision tree is used for K-means, and the decision tree results of the three years can be obtained to discuss the relation and content of each variable, as well as their impact on the governmental efficiency of Taiwan. Finally, the garbage and wastewater treatment of the counties/cities may affect government efficiency. According to this study, the undesired outputs may have considerable impact on the input of the counties/cities. In the future, more efficient counties/cities can be taken as examples to reduce undesired outputs, while increasing desired outputs 老的产出), in order that governmental efficiency can be improved, which can affect the social, economic, and environmental developments of each county/city. As the government improves social, economic, and environmental efficiency, the efficiency of the governments of the less efficient counties/cities can be improved, and the sustainable development of environment protection can be promoted. According to

analysis, it can be inferred that the efficiency of the DEA-SBM model in new municipalities must be improved as counties and cities are merged, thus, Taoyuan, Taichung, Tainan, and Kaohsiung will include the land, as well as the old counties and cities. According to the DEA-SBM model for 2015, the government efficiency of Tainan City is the highest. This study analyzes whether local governments can protect the environment while developing their economy. As shown in the analysis results, if local governments attach importance to air quality and water and land pollution, they can develop sustainable operation capabilities, which can contribute to providing better living environments and development spaces in Taiwan. It is hoped that future government can pay more attention to our land, formulate policies to improve the cities, and develop sustainable operation capabilities.

# References

Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. Government Information Quarterly, 32(4), 399-418.

Cao Mengtian; Luo Luyang(2017) The policy performance evaluation of new energy automobile industry based on DEA-SBM method. 2017 3rd International Conference on Information Management (ICIM).1-5.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. European journal of operational research, 2(6), 429-444.

Chehata., N., Guo, L., & Mallet, C., (2009), “Airborne LiDAR Feature Selection for Urban Classification using Random Forests,”IntArchPhRS, 38(3), 207-212.

Crimmins, S.M., Dobrowski, S.Z., & Mynsberge, A.R., (2013), “Evaluating Ensemble Forecasts of Plant Species Distributions under Climate Change,”Ecological Modelling, 266, 126-130.

Gurin,J(2014)Open Data Now,McGraw Hill Education.

Lin, Cheng-Hsien(2015). Dynamic Performance Perspective with Matthew Effect. Diss. 56-68

Martin, C. (2014). Barriers to the Open Government Data Agenda: Taking a Multi-Level Perspective. Policy & Internet, 6(3), 217-240.

Odom, M. D. and R. Sharda.(1990). A neural network model for bankruptcy prediction. IEEE INNS IJCNN 2: 33-38.

Pan, Y. H. (2015). The Evaluation of Car Icons and Head Up Display Icons Design Using a Decision Tree.

Peck, E., Perri 6, Gulliver, P., & Towell, D. (2004). Why do we keep on meeting like this? The board as ritual in health and social care. Health Services Management Research, 17(2), 100-109.

Quinlan, J. R. (1993.)C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers,

Ran Wang; Sam Kwong; Xi-Zhao Wang; Qingshan Jiang(2015) Segment Based Decision Tree Induction With Continuous Valued Attributes. IEEE Transactions on Cybernetics. 1262 – 1275.

S. R. Nandanwar; S. B. Warkad; K. M. S. Y. Konara; M. L. Kolhe; N. P. Patidar(2016) Load management for voltage security using probabilistic fuzzy decision tree method. 2016 IEEE International Conference on Information and Automation for Sustainability (ICIAfS). 1 – 5.

Satya Samyukta Kambhampati; Vishal Singh; M. Sabarimalai Manikandan; Barathram Ramkumar(2015) Unified framework for triaxial accelerometer-based fall event detection and classification using cumulants and hierarchical decision tree classifier. Healthcare Technology Letters. 101 – 107.

Sayogo, D. S., Pardo, T. A., & Cook, M. (2014, January). A framework for benchmarking open government data efforts. In System Sciences (HICSS), 2014 47th Hawaii International Conference on (pp. 1896-1905). IEEE.

Schrock, A. R. (2015). Civic hacking as data activism and advocacy: A history from publicity to open government data. New media & society, 18(4), 581-599.

Shadbolt, N., O'Hara, K., Berners-Lee, T., Gibbins, N., Glaser, H., & Hall, W. (2012). Linked open government data: Lessons from data. gov. uk. IEEE Intelligent Systems, 27(3), 3-24.

Sonal Singhal; Shrey Jain; Nithin B; Bhawna Rawat; H. O. Gupta(2016) Performance evaluation of Uttarakhand electric utility data using slack based measure and Two-stage DEA-SBM modelling. 2016 International Conference on Electrical Power and Energy Systems (ICEPES)27-34.

Yao, C. C. (2014). The Study of a Bayes Classifier Enhanced by Least-Mean-Square Learning Approach.

Yisen Wang; Shu-Tao Xia(2017) Unifying attribute splitting criteria of decision trees by Tsallis entropy. 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2507 – 2511.

Yu, H., & Robinson, D. G. (2012). The new ambiguity of'open governme