97-B-1

A Search Model with Subjective Judgments : Auditing of Incorrect Tax Declarations*

Seiji Iwasaki[1] and Kaoru Tone[2]

July 1, 1997

1) Yokohama Customs

2) Graduate School of Policy Science, Saitama University

*) Forthcoming in *Omega*. The views expressed in this paper are those of the authors and not those of the organizations they belong to. Comments on this paper should be sent to the following address:

Abstract

This paper proposes a new search model to deal with public sector decisions in which intangible factors need to be considered along with tangible ones. Such problems are often found in a wide range of administrative investigations and criminal inspections in Japan. In an effort to manage such problems, this article first constructs an Analytic Hierarchy Process (AHP) model to deal with the intangible factors. Then, we reinforce the AHP model by incorporating tangible factors which are not included in the model. The probit and/or logit models are applied to test the statistical significance of this combination, based on a past data set. This model can be considered as a two-stage procedure in the sense that the AHP results (the first stage) are utilized to construct a statistical model in the second stage, which aims at obtaining a better probability of detection. Although this study can foresee the probability of detection for individual objects based on this probability model, there is another important practical issue, i.e. the scheduling of associated officials. Therefore, this study proposes a scheduling model using mathematical programming methodology. Finally, the proposed model was applied to the investigation work of the Post-Clearance Audit Department of Japanese Customs. It was found that the probability of finding incorrect declarations would be improved from the current 60% to 75–80%. While our problem is not related to criminal activities directly, this study predicts that our approach might be applicable for other governmental investigations and inspections in the scope of this empirical study.

Keywords

AHP, logit/probit model, knapsack problem, public sector, auditing, law enforcement

1. Introduction

Many managerial problems have been solved by quantitative methods, including product mix planning by linear programming, crew scheduling by mixed-integer programming, and facility location problems by network optimization techniques.

However, some decisions include experts' subjective judgments arrived at via their experience and these cannot be solved by straightforward applications of quantitative methodologies. For example, a regional tax office deals with tax payments of more than several thousands companies, on average[1]([17]), using the limited man-hours available. Because of this, they cannot investigate all the companies in a region within a fixed time period, and therefore they may select only limited number on the basis of a variety of information, some of which is not quantitative([10]). Moreover, the implementation of administrative reform is being discussed in Japan as an urgent problem, and a reduction in the number of public officials is one of the important areas. Because of this, a more rationalized and efficient method for selection should be established. This research was undertaken to develop a new method to solve such decisional problems by taking both tangible and intangible factors into consideration. The method proposed in this study confirmed its applicability to other practical problems of maximizing the probability of finding incorrect statements and/or maximizing the amount of additional (revealed) taxes after investigations.

This paper presents the handling of the problem by means of a combination of management science and statistical techniques. To make clear the point of the research, the methodological outline will be explained by taking

[1]We will call tax-payers in general 'companies' in this paper.

the case of tax offices' investigations.

First, we construct an Analytic Hierarchy Process (AHP) model which deals with both qualitative (intangible) and quantitative factors, in order to rank the priority of companies. Then, we develop a stochastic model which incorporates an AHP score with other related factors, in order to increase probability of finding incorrectness. Actually, this process can be executed by estimating the coefficients of the employed regression factors, using a past data set under a maximum likelihood rule. After the validity of this stochastic model was confirmed, it can be used for the purpose of forecasting the probability of finding an incorrect statement for each company. Finally, the tax official scheduling problem will be solved via a 0-1 integer programming model, along with the objective of maximizing the total probability or maximizing the total additional taxes, so that we can obtain a feasible and optimal official schedule during a given time period.

The rest of the paper is organized as follows. Section 2 summarizes the methodological issues. Section 3 deals with the construction of a stochastic model which incorporates an AHP score with other related factors of importance. Then, in Section 4, the results thus obtained are used to construct an optimization model for tax official scheduling within their limited capacity. This process ensures the feasibility of the proposed method. Finally, in Section 5, this process is applied to a Post-Clearance Audit Department of Japanese Customs. A data set for the year of 1993 was employed to estimate the corresponding stochastic model. After confirming the validity of such a model, we then computed the probability for each importer (company) which was utilized in the optimization model for scheduling of the officials. It was found in this study that the probability of finding incorrect declarations is theoretically improved from the present 60% to 75-80%0. Thus, this method provides us with an optimal strategy for the identification of incorrect tax declarations. Although we applied this model to a customs problem, there are many other areas of interest to which it can be successfully applied.

2. Methodological Outline

As was explained in the preceding section, one of the main objectives of this research is to rank companies under investigation based on the likelihood of incorrect declarations. Our judgment will be performed on the basis of several criteria. Some criteria are subjective and intangible. For this purpose, we use the Analytic Hierarchy Process (AHP) first developed by Saaty([14]). Before going into the details of the AHP in our case, we will explain the reason why we chose it from among other multiple-criteria decision making methodologies. As was mentioned in Belton[2], the field of multiple-criteria decision making (MCDM) has expanded rapidly over the last two decades and continues to do so. Among the MCDM methodologies, the Multi-Attribute Utility (or Value) Theory (MAUT[11]) and the AHP are the approaches best suited for this kind of problem and the most widely used in practice. (See Belton[2].) Also, the review by Zanakis *et al.*[21] showed that these two have been most frequently employed as the models for measurement within the service and government sectors. Harker[9] compared the AHP with other decision methodologies, especially with the Delphi technique and the MAUT and we agree with his views. The reasons for evaluating the AHP over the MAUT are as follows: (1) We can easily incorporate the hierarchy structure of decision making problems into the AHP model. (2) We do not need any assumption on a von Neumann-Morgentern type utility function estimation. (3) We can evaluate the inconsistency of the decision maker. (4) The AHP is fitted for group decision situations. And lastly, (5) the AHP can easily be understood and handled by practitioners.

However, the AHP has not been commonly accepted as an established methodology and has been under active research and development. It has been criticized for its theoretical or axiomatic foundation. (See Belton and Gear[3], for example.) Therefore, its users need to be careful in applying it to their problems. In our case, as can be seen in Sections 3 and 5, we do test the statistical hypothesis on the relationship between the model scores and the past data, where the model scores are obtained by combining the AHP weights with other tangible factors. Thus, we can be, to some extent, free from the criticisms occasionally showered on the AHP results.

Figure 1 describes a typical hierarchy structure for the ranking process of selecting companies, which was constructed based on the publication in reference to the investigations by Japanese tax offices ([10]).

Figure 1

The purpose of this AHP application is the selection of companies, as indicated at the top of the hierarchy. The second level (criteria) consists of such factors as Information, Interval of Investigation, Type of Business, Contents of Declaration, and so forth. Finally, the third level comprises a list of companies. It is important to note that this figure is shown only for explanation of the method and does not reflect the hierarchy structure of the case study of Section 5. The contents of the criteria should be selected adaptively through expert opinion and experience of the problem. These may vary according to time and place.

One of difficulties in applying the AHP method to our problem is the 'large' number of alternatives (companies), as depicted at the bottom level of the hierarchy. Because of this, the number of pairwise comparisons becomes too large to handle. For such cases, there are several expedients, some of which can be summarized as follows:

1. Similar companies are classified into a reasonable number, (for example, less than 9 groups), and they are compared pairwise with each other with respect to the criteria in the upper level. This article refers to this as 'grouping'.

2. For each criterion, we evaluate companies at the bottom by an absolute measure(, e.g. 1 to 10 points), and obtain the score of the company as a weighted sum of the absolute measures, where the weight is associated with that of the criterion in the upper level. In this case, we need no pairwise comparisons between companies. This article calls this process, which was first developed by Saaty[15], 'absolute measurement'.

3. Harker[8] describes a set of techniques to reduce the number of pairwise comparisons that the decision maker must make during the analysis of

a large hierarchy. We can apply his 'incomplete' pairwise comparison technique. Tone[18] discussed a similar reduction technique within the framework of the geometric mean method.

4. Weiss and Rao[19] proposed the use of incomplete experimental designs, e.g., the method of balanced incomplete block designs (BIBD), for simplifying the data-collection tasks. This method reduces the number of pairwise comparisons, while still ensuring that every pair of attributes is replicated the same number of times in the design.

We can choose one of these expedients according to the situations of the companies. There are several variants in the absolute measure with regard to the range of points, e.g. a 5 point measure, a 7 point measure. The results and conclusions may differ according to the range. The choice depends largely on practitioners' experiences in the decision issue. In some cases, they may think that a 5 point system is enough for the problem. Another recourse is to choose the points measure that gives the highest fitness in the stochastic model described in Section 3.

The next step is to check the fitness of the AHP score to the actual findings of incorrect statements. This process can be carried out using a past data set. If both incorrect (false) statements were found in companies with high AHP scores and the relationship between incorrect statements and AHP scores was proved to be statistically significant, then we could utilize the AHP results for selection of future investigations. At this stage, this study will employ not only the AHP scores but also other considerable related factors, (e.g. annual turnover, annual profit, annual profit/employee, capital, the number of employees and so forth), all of which are quantitative, not included in the AHP criteria, and are judged to be significant for investigation by experts in the field. In the stochastic model, the explanatory variables consist of the AHP score and other numerical factors that are independent of the AHP score. The dependent (objective) variable is binary, i.e. 1 was assigned if incorrect statements were found and 0 in the opposite case. In order to cope with this binary variable, we employ probit and logit models and determine the coefficient of each explanatory variable by a maximum likelihood principle, which will be explained in the next section. The coefficients (including the constant term) will be examined in the form of statistical significance by a t-value and a χ^2-test that are carried out to check the fitness of the two models (probit and logit). Usually, we employ the model that gives the largest χ^2-value. If the stochastic model (probit or logit) thus obtained above is proved to be significant, we can use it for selecting future objects for investigation. Briefly, we should investigate companies with a high score according to the selected model. However, in implementing such a model, we should pay attention to two perspectives; (a) the man-hour capacity of tax officials and (b) the work load for investigation, which may differ from company to company. Thus, this study applies a knapsack type optimization method for solving this implementation issue.

3. The Stochastic Model

We constructed the following stochastic model for representing the probability of finding incorrect declarations in terms of explanatory variables. Let $x_1, x_2, \ldots, x_p$ be the explanatory variables incorporating the AHP score (x_1) and other considerable tangible factors ($x_2, \ldots, x_p$), e.g. annual turnover, annual profit, annual profit/ employee, capital and so forth, and z be the dependent variable associated with the probability of finding incorrect declarations. This study assumes the following model:

$$z = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p + \varepsilon, \tag{1}$$

where ε is a random error term.

In order to estimate the coefficients β_j $(j = 0, \ldots, p)$, this study uses past data concerning n investigations:

$$(x_{1i}, x_{2i}, \cdots, x_{pi}, y_i), \quad (i = 1, \ldots, n) \tag{2}$$

where y_i is binary, i.e. $y_i = 1$ when incorrect statements found and $y_i = 0$ for otherwise.

As an expected correspondence between z and y, we assume that if $z > 0$, then $y = 1$ and if $z \leq 0$, then $y = 0$.

Let F be the cumulative distribution of the random variable z. Then, the probability of $y_i = 1$ is given by

$$P(y_i = 1|z_i) = F(\beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi}), \tag{3}$$

where $z_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi}$.

The relationship between z and y needs some attention. Suppose that we assume $z = y$ and apply the stochastic model below to the data (2).

$$y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi} + \varepsilon. \tag{4}$$

Then, we can find the coefficient $(\beta_0^*, \beta_1^*, \ldots, \beta_p^*)$ by the least squares principle. This is a computationally simple method. However, in this case, the model value calculated by

$$Y_i = \beta_0^* + \beta_1^* x_{1i} + \cdots + \beta_p^* x_{pi}, \tag{5}$$

does not necessarily satisfy the relationship

$$0 \leq Y_i \leq 1. \tag{6}$$

This condition is definitely necessary to interpret Y_i as probability. Even if we further impose the condition (6) in estimating the coefficient, the parameters in the model (4) inevitably have a limited interpretation and range of validity. Furthermore, this type of straight application of the least squares to binary data has several statistical drawbacks with respect to the fundamental assumptions underlying the least squares, since the dependent variable y_i takes only the values 0 and 1. (See Cox[5] pp.16-18, in detail.) Therefore, we need a type of model in which the constraint (6) is automatically satisfied. In many respects, the simplest way of representing the dependence of a probability on explanatory variables so that the constraint (6) is inevitably satisfied, is to postulate a dependence for $i = 1, \cdots, n$,

$$P(y_i = 1) = \frac{\exp(z_i)}{1 + \exp(z_i)}, \tag{7}$$

where $z_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi}$. This model is called the linear logistic model or the *logit* model.

Another candidate is the standard normal distribution which is expressed as

$$P(y_i = 1) = \int_{-\infty}^{z_i} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx. \tag{8}$$

this model is called the *probit* model.

(The following statistical areas are discussed in detail in [1], [5], [13] and [16], in the reference list.)

3.1. Logit and Probit Models

As mentioned above, this study employs two models; logit and probit, for expressing the cumulative distribution F, since they are representative of the latent trait for dealing with binary random variables.

1. **Logit Model**

   In this case, F is the cumulative distribution of a logistic distribution and is expressed as:

   $$\Lambda(z) = \frac{e^z}{1+e^z}.$$

   Thus, the probability of $y_i = 1$ becomes

   $$\Lambda(z_i) = \Lambda(\beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi}). \tag{9}$$

2. **Probit Model**

   F is the cumulative distribution of the standard normal distribution which is expressed as:

   $$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx. \tag{10}$$

   The probability of $y_i = 1$ is expressed by

   $$\Phi(z_i) = \Phi(\beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi}).$$

3.2. Estimation of Coefficients

We estimated the coefficient $(\beta_0, \ldots, \beta_p)$ of these models using past data sets $(x_{1i}, \ldots, x_{pi}, y_i)$ $(i = 1, \ldots, n)$. For this purpose, the following likelihood function is fully utilized:

$$L(\beta_0, \ldots, \beta_p) = \prod_{y_i=0} F(z_i) \times \prod_{y_i=1} [1 - F(z_i)], \tag{11}$$

where $z_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi}$. The logarithmic likelihood function, $\log L$, is maximized in $(\beta_0, \ldots, \beta_p)$. This task requires solving $(p+1)$ simultaneous nonlinear equations in the $(p+1)$ unknown parameters.

There are numerous procedures for finding numerically the maximum of a relatively complicated function like (11). Usually, the Newton-Raphson iterative solution (Fletcher [7]) of the maximum likelihood equation is effective, especially if the number of parameters is not so large.

3.3. The Standard Error of Estimated Coefficients

Let the column vector x_i be $(1, x_{1i}, \ldots, x_{pi})^T$, $\hat{\beta}$ be the maximum likelihood estimation of $\beta = (\beta_0, \cdots, \beta_p)^T$ and $f(z)$ be the probability density function of F. Then, if the number of data approaches infinity, the distribution of $\sqrt{n}(\hat{\beta} - \beta)$ comes to display the multivariate normal distribution $N(\mathbf{0}, B^{-1})$, where the $(p+1) \times (p+1)$ matrix B is defined by

$$B = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \frac{f^2(x_i^T \beta)}{F(x_i^T \beta)(1 - F(x_i^T \beta))} x_i x_i^T. \tag{12}$$

In the case of a large enough n, this term can be approximated by

$$\hat{B} = \frac{1}{n} \sum_{i=1}^{n} \frac{f^2(\hat{z}_i)}{F(\hat{z}_i)(1 - F(\hat{z}_i))} x_i x_i^T, \tag{13}$$

where $\hat{z}_i$ is the model value for $(x_{1i}, \ldots, x_{pi})$.

From the results of the above definitions, the square root of the diagonal elements of $\hat{B}^{-1}$ gives the standard deviation for each estimated coefficient. Moreover, this research tests the significance of the coefficients by their t-values.

3.4. Hypothesis Testing of the Model

This study can extend our previous results to establish a fitness test related to the stochastic model in the following manner. The null hypothesis, that the factors $x_1, \ldots, x_p$ are irrelevant to the results of the investigation, is expressed as:

$$H_0 \quad : \quad \beta_1 = 0,\ \beta_2 = 0, \cdots, \beta_p = 0.$$

The alternative hypothesis becomes:

$$H_1 \quad : \quad \beta_1 \neq 0,\ \beta_2 \neq 0, \cdots, \beta_p \neq 0.$$

This study can apply a likelihood ratio χ^2-test for this purpose in the following manner:

1. Under the null hypothesis $H_0 : \beta_1 = 0,\ \beta_2 = 0, \cdots, \beta_p = 0$, we estimate the constant β_0 from the likelihood function (11). Let the logarithmic likelihood be $\log L_0$.

2. Under the alternative hypothesis of employing all explanatory variables, compute the logarithmic likelihood $\log L_1$.

3. Let

$$T = -2(\log L_0 - \log L_1).$$

   If the null hypothesis is valid, then T displays approximately the χ^2 distribution with the degree of freedom p. Thus, we can check the hypothesis by means of the test statistic T.

4. Man-hour Scheduling

In the preceding stages of this research, we established the stochastic model, logit or probit, for forecasting the probability of finding incorrectness, i.e.

$$\Lambda(\beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi}) \quad \text{(logit)},$$

or

$$\Phi(\beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi}) \quad \text{(probit)},$$

depending upon the model chosen for our investigation.

Let p_j be the probability of the j-th candidate company ($j = 1, \ldots, l$) which is calculated by the above formula.

At this stage, we need to consider several factors for implementing our search procedure. One is the work load for investigation, which may differ from company to company, and the others are the limit of available man-hours for investigation and the upper limit of number of companies to be investigated for a given time period.

Let d_j ($j = 1, \ldots, l$) be the work load (man-hour) for the j-th company, D be the total available man-hours and N be the upper limit of investigations. Officials can determine the work load d_j by considering such factors as the scale and the variety of information of the company j and so on. Then, we have the following knapsack type problem:

$$\max \quad w = \sum_{j=1}^{l} p_j z_j \tag{14}$$

$$\text{subject to} \quad \sum_{j=1}^{l} d_j z_j \leq D \tag{15}$$

$$\sum_{j=1}^{l} z_j \leq N, \quad z_j \in \{0,\ 1\}. \tag{16}$$

The objective is to find the optimal assignment of officials so that it maximizes the sum of the probability of finding incorrectness under the total man-hours and the upper limit constraints.

The above model has several variants, among which we will select a simple but important one with multi-resource constraints. The objective is to maximize the expected total additional tax.

$$\max \quad w \quad = \quad \sum_{j=1}^{l} p_j a_j z_j \tag{17}$$

$$\text{subject to} \quad \sum_{j=1}^{l} d_{jk} z_j \quad \leq \quad D_k \quad (k = 1, \ldots, K) \tag{18}$$

$$\sum_{j=1}^{l} z_j \quad \leq \quad N, \quad z_j \in \{0,\ 1\}\,, \tag{19}$$

where a_j is the expected additional tax paid by the j-th company, d_{jk} is the work load of resource k for investigating company j, and $D_k \quad (k = 1, \ldots, K)$ is the limit of the available amount of resource k. We can solve these models by using software for linear programming problems with 0-1 integer variables.

5. A Case Study

In this section, the applicability of the proposed method is verified by a case study. The subject chosen is a selection problem of importers to be investigated by a Post-Clearance Audit Department of Japanese Customs.

Before going into a detailed analysis, we will describe briefly the Customs investigation system in Japan. Generally, the Customs employs the post-clearance audit system based on the self-assessment made by importers. Then Customs officials examine the documents in order to confirm whether the declared value was right or not after clearance. If incorrect declarations are found as a result of investigation, officials usually recommend importers to correct their declarations. However, in cases where officials judge the false declaration was made intentionally, importers would be punished via

Customs Law and so on.

Practically, because of limited man-hours, they cannot investigate all importers, and the Customs selects companies based on their experience and a variety of information. It is reported that the rate of finding incorrect declarations is approximately 60% and this percentage has remained same for the past several years ([6]). It is therefore considered to be an urgent subject for the Customs to develop a more efficient selection methodology for this purpose.

The following case study utilizes the Customs data of the year 1993 to build up the proposed stochastic model and the scheduling model of officials. Then, we applied the models thus obtained for the data of 1994. The companies under investigation are classified into three categories denoted as A, B and C, depending on their type of industry.

This section is divided into three parts. In Section 5.1, we select a small data set in Category A from the 1993 data and follow the proposed method step by step so that the numerical treatments can be better understood. Then, in Section 5.2, the whole data set is analyzed. It was found that the probability of finding incorrect declarations can be improved from the present 60% to 75-80%. Lastly, in Section 5.3 some observations will be presented on why this improvement was made possible by our proposed method.

5.1. A Small Data Set Case

Although this study utilizes the whole data set kept by Yokohama Customs, it is impossible to describe the numerical treatments in detail. Therefore, we trace the calculation processes in the case of a limited number of data. For this purpose, the data of twenty three companies (T1 to T23) in Category A were chosen. The 23 companies were further divided into 4 groups according to their imported materials. The number of samples are six each for Group 1 to Group 3 and five for Group 4. (See Table 5.)

5.1.1 Subjective Judgment by AHP

The hierarchy structure in the AHP model was established by experts' opinions (Figure 2). The structure is conceptually similar to Figure 1. A part of the pairwise comparisons and AHP scores for Group 1 (T1 to T6) are shown in Tables 1 to 4. (Those for other groups are not shown.). It should be noted that the AHP scores of Group 4 are multiplied by 5/6, since this group contains 5 companies while others contain 6. Thus, the summation of scores in this group is 5/6. (See Table 5.) We call this operation *normalization.* The AHP scores of 23 companies thus obtained are exhibited in *AHP* column in Table 5.

Figure 2

Table 1 to Table 4

5.1.2 Construction of Stochastic Model

In addition to the AHP scores, the proposed method deals with other tangible factors, such as annual turnover, profit, capital, and number of employees in the company. We examined the correlation analyses between the probability of finding incorrect declarations and each tangible factor and found that an index which represents the *scale of foreign trade* is significantly correlated.

After these preliminary surveys with the intangible (*AHP*) factor and the tangible factor (*scale of foreign trade*), we developed the logit and the probit models, including two explanatory variables, expressed by AHP and x. Accordingly, the mathematical expression becomes

$$z = \beta_0 + \beta_1 AHP + \beta_2 x. \tag{20}$$

The coefficients β_0, β_1 and β_2 are estimated by utilizing 23 sets of data shown in Table 5, where the column y_i takes the value 1, if incorrect declarations are found, and 0 otherwise.

Table 5

By using the maximum likelihood estimation, the following model equations were obtained. Standard deviation, t-value and χ^2-value for each coefficients are shown in Table 6.

Logit Model

$$z = -4.460 + 21.81AHP + 8.468 \times 10^{-3}x + \epsilon. \tag{21}$$

Probit Model

$$z = -2.543 + 11.89AHP + 4.973 \times 10^{-3}x + \epsilon. \tag{22}$$

Table 6

Although the logit model has a larger χ^2-value as shown in Table 6, we chose the probit model, since the latter has slightly larger t-value for all three coefficients.[2]

We now consider a knapsack type 0-1 integer programming problem corresponding to the expressions (14) to (16) in Section 4. The coefficient p_j in the objective function is estimated through the Logit model (21) as described in Section 3. Since the demonstration of this small sample is purely for explanatory purposes, the work load d_j is chosen randomly from the interval [4, 12], and we set D (the total man-hours available) = 40 and N (the upper limit of investigations) = 9. See Table 7.

By using the software XPRESS -MP ([20]), we obtained the optimal solution $\{z_i^*\}$ as exhibited in Table 7, where $z_i^* = 1$ means 'to go' and $z_i^* = 0$ 'not to go'. Eventually, this optimal solution chose companies with a high p_i value, and eight out of nine companies investigated were found to have declared incorrectly. Since the purpose of this part of the case study is to demonstrate the proposed processes, we do not intend to compare the optimal solution with the actual investigation results.

Table 7

[2]It is reported that there is, in most cases, no significant difference between the probit and the logit models in binary data analysis ([16]).

5.2. Results Obtained Using the Whole Data

Yokohama Customs investigated several hundreds of importers in 1993. Estimation of the coefficients, β_0, β_1 and β_2, was carried out using the data set[3] for 1993, of which a portion is exhibited in Table 8. Table 9 presents the results along with the t-value for each estimated coefficient and the χ^2-value for each model. All the coefficients and both models were found to be statistically significant.

Table 8

Furthermore, Table 10 shows the correlation coefficient between the model and the actual value for each category. More exactly, these coefficients were calculated as follows. Suppose that the i-th importer in category A has the model value z_i, and the actual probability of finding incorrectness in the class to which the importer belongs is $\bar{z}_i$. The class was determined by a level of the AHP score and one tangible factor. The correlation coefficient was calculated as the one between $\{(z_i, \bar{z}_i)\}(i \in \text{Category } A)$. Again, all coefficients were found to be statistically significant. Here, we chose the one with both a larger χ^2 value and coefficients of correlation in this case.

[3]Due to the confidential nature of the subject matter, we cannot disclose actual data employed.

Table 9

Table 10

Then the numerical computations of a knapsack type 0-1 integer program were performed with a data set for the year 1994 in three categories of importers, under several varieties of constraints regarding the available man-hours and the upper limit of investigations. We obtained approximately 80% as the average of these forecasted probabilities of selected importers, i.e., those of companies with $z_i^* = 1$. Consequently, if we implement our customs investigation based on the optimal solution $\{z_i^*\}$, the probability of identifying incorrect declarations may be increased from the present 60% to 80%.

5.3. Some Observations on the Case Study

This study applied our method to the data set of Yokohama Customs in the year of 1993 and 1994. More concretely, first, the 1993 year data were utilized to establish the stochastic models and then the models were applied to forecast the probability of finding incorrect declarations of companies for the year 1994. Based on the probability thus estimated, the $0-1$ integer linear program was solved optimally to determine which companies should be investigated within the available man-hours and the upper limit constraints. The results indicate that the probability of finding incorrect declarations could be raised from the present 60% to 80%, at least theoretically. This is a remarkable difference and improvement.

We believe this improvement was caused by the combination of experts knowledge and management science methodologies. From long experience, the expert officials know which factors are important in determining where to investigate. However, these factors would be mostly intangible and difficult to prioritize. The AHP succeeded in measuring the priorities of these intangible factors and in scoring the incorrectness of target companies. Furthermore, the stochastic model, which combines the AHP score with other tangible factors, had its creditability verified by statistical tests. The improvement is the result of deliberate integration of expert knowledge and management science methodologies.

6. Conclusion

This article proposed a new search model with subjective judgments and then applied it to the selection work of the Post-Clearance Audit Department of Japanese Customs. It found that the probability of finding incorrect declarations could be improved from the current 60% to 80%. This improvement is realized by the combination of wide expert knowledge and mathematical methodology.

However, it should be noted that the data set used in our experiment was obtained from a conventional selection process (not by the proposed one) and hence, if the Japanese Customs could replace its selection process by the one proposed in this research, we expect that the rate of finding incorrectness might be further improved. It is hoped that this model will be applied to other similar kinds of administrative investigations and criminal inspections.

In this research, AHP scores were used together with logit and probit models as suitable for forecasting the probability of finding incorrect customs declarations. Then, we solve a resource allocation problem for finding the optimal assignment of tax officials, using the results of the stochastic model as data.

Zanakis *et al.* [21] surveyed 306 articles on program evaluation and fund allocation methods within the service and government sectors, which appeared in 93 journals. They found that few publications dealt with evaluation and allocation as an integrated process. Actually, in only two, (Bodily[4] and Khorramshahgol and Steiner[12]), were such approaches within criminal and law enforcement applications. Our approaches are different from theirs in following ways. The former deals with the spatial design problem of the police mobile units, incorporating the multi-objectives of administrations, citizens and service personnel. Using multi-attribute utility theory, alternative locations are evaluated according to the preferences for efficiency and equality of service of the three interest groups. Iterative improvements and heuristics are utilized to find a satisfactory solution. Bodily's paper is different from ours in the problem area examined and methodologies employed. It concerns a stationary allocation problem, while ours deals with a search model. The latter paper by Khorramshahgol *et al.* deals with the problem of allocating funds to competing projects. Investment decisions normally involve several, often conflicting and non-commensurable, goals. So, they applied the goal programming approach to this problem, where the weights to the goals were determined by the Delphi method. Thus, this work differs from ours in methodology. Instead of the Delphi method, we utilized the AHP results coupled with the stochastic model and tested its statistical significance. (Saaty[14] described comparisons between the AHP and the Delphi method.)

Thus, we believe and hope that this article will present a new methodology for law enforcement and criminal applications.

Finally, as an important future research task, this study will extend our interest to an investigation regarding which stochastic models will be applicable to other decisional issues in both the public and private sectors.

Acknowledgment

The authors wish to thank two anonymous referees for their valuable suggestions and the staff of Yokohama Customs for supporting the study.

References

[1] Amemiya, T., *Advanced Econometrics,* Harvard University Press, 1985.

[2] Belton, V., "A Comparison of the Analytic Hierarchy Process and a Simple Multi-Attribute Value Function," *European Journal of Operational Research,* Vo.26, pp.7-21, 1986.

[3] Belton, V. and T. Gear, "On a Shortcoming of Saaty's Method of Analytic Hierarchies," *Omega*, Vol. 11, pp.228-230, 1984.

[4] Bodily, S.E., "Police Sector Design incorporating Preferences of Interest Groups for Equality and Efficiency," *Management Science*, Vol. 24, pp.1301-1313, 1978.

[5] Cox, D.R., *Analysis of Binary Data,* Methuen, 1970.

[6] Customs and Tariff Bureau, Ministry of Finance, Japanese Government, *The Third Annual Report of Customs and Tariff Bureau,* Nippon Kanzei Kyokai, 1994.

[7] Fletcher, R., *Practical Methods of Optimization,* Vol. 1, John Wiley & Sons, 1980.

[8] Harker, P.T., "Incomplete Pairwise Comparisons in the Analytic Hierarchy Process," *Mathematical Modelling,* Vol. 9, pp.837-848, 1987.

[9] Harker, P.T., "The Art and Science of Decision Making: The Analytic Hierarchy Process," in *The Analytic Hierarchy Process,* B.L. Golden, E.A. Wasil and P.T. Harker (Eds.), Springer-Verlag, pp.3-36, 1989.

[10] Hotti, T., *The Scene of the Tax Investigations by Tax Offices,* Gyousei, 1995.

[11] Keeney, R.L. and H. Raiffa, *Decision with Multiple Objectives: Preference and Value Tradeoffs,* John Wiley, New York, New York, 1976.

[12] Khorramshahgol R. and H.M. Steiner, "Resource Analysis in Project Evaluation: A Multicriteria Approach," *J. of Operl. Res. Soc.*, Vol. 39, pp.795-803, 1988.

[13] Pindyck, R.S. and D.L. Rubinfeld, *Econometric Models and Economic Forecasts,* Second ed., McGraw-Hill, 1981.

[14] Saaty, T.L., *The Analytic Hierarchy Process,* McGraw-Hill, 1980.

[15] Saaty, T.L., "Absolute and Relative Measurement with the AHP. The Most Livable Cities in the United States," *Socio-Eco. Plann. Sci.* Vol.20, No. 6, pp.327-331, 1986.

[16] Takeuchi, K. *et al.*, *Analysis of Scientific Data (Basic Statistics 3),* The Statistic Section, Department of Social Science, College of Arts and Science, University of Tokyo, University of Tokyo Press, 1992.

[17] Tax Agency, Japanese Government, *The 120th Annual Statistical Report of Tax Agency,* Okura Zaimu Kyokai, 1994.

[18] Tone, K., "Two Technical Notes on the AHP Based on Geometric Mean Method," in *Proceedings of the Fourth International Symposium on the Analytic Hierarchy Process,* Faculty of Business Administration, Simon Fraser University, Canada, pp.375-381, 1996.

[19] Weiss, E.N. and V.R. Rao, "AHP Design Issues for Large-Scale Systems," *Decision Sciences,* Vol. 18, No. 1, pp.43-61, 1987.

[20] XPRESS-MP, Release 7, Dash Associates.

[21] Zanakis, S.H., T. Mandakovic, S.K. Gupta, S. Sahay and S. Hong, "A Review of Program Evaluation and Fund Allocation Methods Within the Service and Government Sectors," *Socio-Eco. Plann. Sci.* Vol.29, No.1, pp.59-79, 1995.

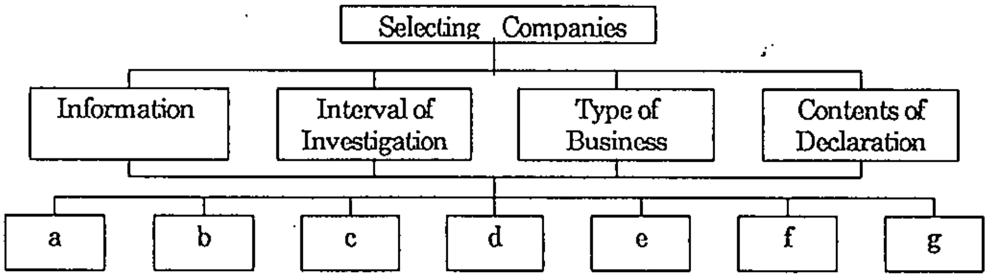

Fig. 1. Hierarchy Structure for Selecting Companies

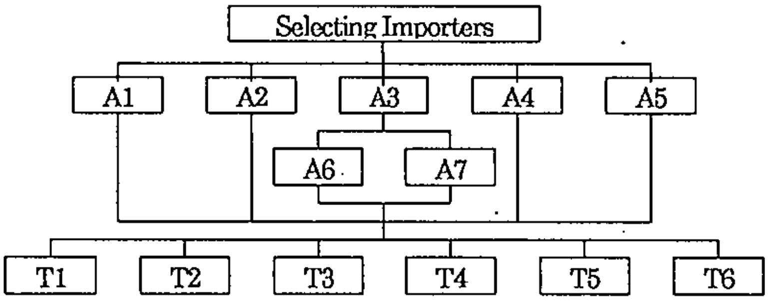

Fig.2. Hierarchy Structure for Selecting Importers

Table 1. Pairwise Comparisons of Selecting Importers

	A1	A2	A3	A4	A5	weight
A1	1	1/6	1/3	4	1/7	0.0697
A2	6	1	2	6	1/3	0.2412
A3	3	1/2	1	3	1/7	0.1182
A4	1/4	1/6	1/3	1	1/8	0.0374
A5	7	3	7	8	1	0.5336

C. I. =0. 095
C. R. =0. 085

Table2. Pairwise Comparisons of A3

A3	A6	A7	weight
A6	1	5	0.8333
A7	1/5	1	0.1667

C. I. =0. 000
C. R. =0. 000

Table 3. 1. Pairwise Comparisons of A1

A1	T1	T2	T3	T4	T5	T6	weight
T1	1	1/2	7	6	9	7	0.3264
T2	2	1	8	7	9	8	0.4431
T3	1/7	1/8	1	1/2	5	1	0.0580
T4	1/6	1/7	2	1	7	2	0.0923
T5	1/9	1/9	1/5	1/7	1	1/5	0.0223
T6	1/7	1/8	1	1/2	5	1	0.0580

C. I. =0. 095
C. R. =0. 077

Table 3. 2. Pairwise Comparisons of A2

A2	T1	T2	T3	T4	T5	T6	weight
T1	1	5	1/2	1	3	1/3	0.1442
T2	1/5	1	1/7	1/5	1/3	1/8	0.0313
T3	2	7	1	2	4	1/2	0.2419
T4	1	5	1/2	1	3	1/3	0.1442
T5	1/3	3	1/4	1/3	1	1/5	0.0636
T6	3	8	2	3	5	1	0.3748

C. I. =0. 023
C. R. =0. 019

Table 4. Results of AHP

	AHP score
T1	0.2470
T2	0.1581
T3	0.2628
T4	0.0953
T5	0.0756
T6	0.1610

Table 5. Small Sample Data for Estimating Coefficients

i	company	group	AHP AHP_i	foreign trade * x_i	detection y_i
1	T1	1	0.247	562.58	1
2	T2		0.158	455.90	1
3	T3		0.263	92.49	1
4	T4		0.095	55.96	0
5	T5		0.076	6.91	0
6	T6		0.161	15.09	1
7	T7	2	0.128	407.9	1
8	T8		0.298	44.22	1
9	T9		0.161	24.08	0
10	T10		0.289	34.44	1
11	T11		0.060	51.49	0
12	T12		0.063	50.59	0
13	T13	3	0.225	475.63	1
14	T14		0.139	92.82	0
15	T15		0.184	52.00	1
16	T16		0.241	547.12	1
17	T17		0.136	556.19	1
18	T18		0.075	49.17	0
19	T19	4	0.217	25.78	1
20	T20		0.188	658.10	1
21	T21		0.285	128.35	0
22	T22		0.072	26.87	0
23	T23		0.072	22.41	0

* y_i = 1, if incorrect declarations found, = 0 if not found.

Table 6. Results of Statistical Tests

Model	β_0	β_1	β_2	
	t-value *level*	*t*-value *level*	*t*-value *level*	χ^2-value *level*
Logit	-2.354 5%	2.135 5%	1.645 20%	17.08 1%
Probit	-2.665 5%	2.381 5%	1.900 10%	17.01 1%

* *level* means significant level.

Table 7. Computed Results for Small Sample Problem

i	company	group	probability p_i	work load d_i	solution z_i^*	correspondence to y_i
1	T1	1	0.999	6	1	yes
2	T2		0.945	5	1	yes
3	T3		0.852	9	0	–
4	T4		0.128	6	0	–
5	T5		0.054	4	0	–
6	T6		0.290	7	0	–
7	T7	2	0.843	4	0	–
8	T8		0.887	9	1	yes
9	T9		0.305	10	0	–
10	T10		0.856	8	1	yes
11	T11		0.058	10	0	–
12	T12		0.062	12	0	–
13	T13	3	0.994	6	1	yes
14	T14		0.334	8	0	–
15	T15		0.461	8	0	–
16	T16		0.999	12	1	yes
17	T17		0.967	7	1	yes
18	T18		0.080	11	0	–
19	T19	4	0.565	5	0	–
20	T20		0.998	4	1	yes
21	T21		0.931	10	1	no
22	T22		0.060	12	0	–
23	T23		0.058	4	0	–

Table 8. Past Data Set for Estimating Coefficients

Category	Factor \ Importer	I_1	I_2	I_3	·····	I_{n-1}	I_n
A	*AHP*	0.095	0.177	0.139	·····	0.104	0.307
	x	55.96	446.8	2835	·····	41.19	105.7
	y	0	0	1	·····	0	1
B	*AHP*	0.209	0.198	0.270	·····	0.243	0.085
	x	45.23	10.30	125.1	·····	678.1	5.909
	y	0	1	1	·····	1	0
C	*AHP*	0.125	0.191	0.264	·····	0.074	0.072
	x	70.44	61.75	184.5	·····	20.82	9.537
	y	0	1	1	·····	1	0

* *n* is different for each category

Table 9. Results of Coefficients Estimation and Statistical Tests

Category	Model	β_0 (t-value) (level)	β_1 (t-value) (level)	β_2 (t-value) (level)	χ^2-value (level)
A	Probit	−1.450 (−3.805) (1 %)	6.051 (3.213) (1 %)	7.631×10^{-3} (2.850) (1 %)	25.486 (1 %)
	Logit	−2.559 (−3.787) (1 %)	10.71 (3.223) (1 %)	4.666×10^{-2} (2.898) (1 %)	26.411 (1 %)
B	Probit	−1.124 (−2.948) (1 %)	4.155 (2.107) (5 %)	8.177×10^{-3} (2.568) (5 %)	16.770 (1 %)
	Logit	−1.948 (−2.972) (1 %)	7.152 (2.150) (5 %)	1.428×10^{-2} (2.441) (5 %)	17.123 (1 %)
C	Probit	−1.230 (−2.825) (1 %)	4.641 (2.041) (5 %)	9.593×10^{-3} (2.042) (5 %)	12.680 (1 %)
	Logit	−1.995 (−2.972) (1 %)	7.449 (1.946) (10 %)	1.578×10^{-2} (1.894) (10 %)	12.480 (1 %)

* model: $z = \beta_0 + \beta_1 AHP + \beta_2 x + \varepsilon$

Table 10. Correlation Coefficients between Model and Data

Category	Model	correlation	level
A	Probit	0.8751	1%
	Logit	0.8628	5%
B	Probit	0.9060	1%
	Logit	0.9169	1%
C	Probit	0.9509	1%
	Logit	0.9465	1%