Asian Journal of Business Management 4(4): 376-385, 2012
ISSN: 2041-8752

Submitted: June 28, 2012 Accepted: August 08, 2012 Published: September 25, 2012

# Data Analysis and Decision Trees for Analysis and B2C Controls

Mohammad Nazaripour
University of Kurdistan, Sanandaj, Iran

**Abstract:** This study uses a two-step procedure for the evaluation of B2C controls, first, using a Data Envelopment Analysis (DEA) model, second, decision trees. The results of the DEA model indicate that retail firms and information service providers implement B2C controls more effectively than financial firms do. Controls for system continuity are implemented more effectively than entry controls. In financial firms, controls for system continuity, communication controls and entry controls, in a dropping order, are effectively followed in B2C approaches. Every company can determine its relative level of reduction in each part of controls in order to make the control system effective. The firms that effectively implement B2C controls are determined using a decision tree model. The decision tree model issued to suggest the level of controls and argued rules for controls guidance. This state the possibility of using decision trees for controls evaluation in B2C approaches.

**Keywords:** B2C approaches, B2C controls, Data Envelopment Analysis (DEA), decision trees

## INTRODUCTION

As organizations rely on IS (Information Systems) for strategic superiority and functioning's, management needs to pay more attention to IS protection issues due to a parallel increase in the impact of IS protection misuses. Jansen (2002) state that Controls of e-business approaches may ensure the protection, unity and controllability of the configured software, data and support organization. Pareek (2007) suggest that as controls are not only overpriced to put in place and work, but also increase audit efforts and slow down the implementation of business procedures, therefore it is necessary to automate controls in an optimum manner from a cost and regulatory position. How internal auditors or protection administrators make decisions of controls evaluation is in large part of judgment and experience. Thus, it is necessary to formulate a systematic approach for controls evaluation for protection management that relies on subjective judgments of internal auditors or protection administrators. Gordon and Loeb (2002) have shown an economic model to determine the optimal level of investment in information protection. Cavusoglu *et al.* (2004) stated a model depend on game theory for strategic investment decisions in protection controls; in the IT protection problem, the firm and hacker are players and the firm's get results from protection investment depends on the size of slashing it is subjected to Lee *et al.* (2005) suggest that industry attributes such as system in danger to protection risks and availability of value added networks can affect the economy of EDI (Electronic Data Interchange) controls. Pareek (2006) used an optimization algorithm on a linear scheduling model to identify controls that need to be tested to address the risks. Cerullo and Cerullo (2005) argued that the law of decreasing returns posits that, beyond a certain point, the effectiveness of protection supplied by additional controls will decrease and no longer improve the quality of the information systems. Guidance can be supplied by analyzing the data collected from tests used to quantify controls. In look out of the state of implementation and given the high cost and resources needed to develop controls embedded in the system, it is necessary to analyze the economy of controls in B2C (Business-to-Consumer) approaches (hereafter B2C controls). Depending on who performs the analysis, however, a wide range of protection quantifies may be implemented, resulting in either too few or too many B2C controls. This study proposes to investigate the evaluation of B2C controls, i.e., economy analysis and guidance of controls, using a Data Envelopment Analysis (DEA) model and decision trees. Previous studies have combined the use of DEA and decision trees in analyzing organizational units (Samoilenko and Osei-Bryson, 2008; Seol *et al*., 2007; Sohn and Moon, 2004). Using decision trees may support IS managers and assure them of the kinds of control quantifies that are necessary under given system circumstances. The firms that effectively implement B2C controls are determined using a decision tree model. The decision tree model is further used to suggest the level of controls and to suggest rules for controls guidance's. This suggests the possibility of using decision trees for controls evaluation in B2C approaches.

## LITERATURE REVIEW

Ott *et al*. (2008), suggests that audit management staff members face a continuous need to cut time in completing controls evaluation and testing. Automated data mining tools are capable of studding a large amount of information and searching for designs that may not be identified easily by manual means. The objective of controls evaluation is to assure customers, stakeholders and government agencies that controls are in place and effective. Few protection quantifies make the firm's IT environment unprotected to a wide range of potentially damaging risks and many protection quantifies lead to an increase in costs, to slowdowns and to delays in procedure approaches (Cerullo and Cerullo, 2005). The evaluation of controls is to ensure adequacy of controls for a balance between costs sustained for implementing controls and the resulting benefits derived. Bakshi (2004) suggests that controls evaluation provides a method for management to determine the current status of their information protections schedules; this evaluation involves controls improvement, if necessary through the timely detection and correction of weak controls. Evolution offers a means of identifying problems of controls and guidance's for improvement. Further, evaluation is done to reduce or remove costly and unsuccessful controls while creating valuable alternatives. Therefore the Business Procedure Model is depending on the identification of risks related with each business procedure. Sound knowledge of business procedures, information control objectives and company environments are crucial factors in the success of introduction of the controls evaluation model. Estimating an IT protection investment has been a sticking point and a reasonable methodology is needed to analyze protection investments (Cavusoglu *et al*., 2004). This study used DEA to analyze the economy of B2C controls. Charnes *et al*. (1995) argued that DEA is a mathematical scheduling formulation depends on technique that provides an effective boundary to suggest an estimate of the relative economy of each Decision Making Unit (DMU) in a problem set. DEA is developed around the concept of evaluating the economies of a decision alternative depend on its performance of creating outputs in means of input consumption. The economy of each DMU, relative to its peers, is defined as the ratio of that member's weighted sum of outputs to its weighted sum of inputs. The parametric approach, such as regression equation and discriminant analysis, however, needs assumptions about the functional form and the distribution of error terms. Those DMUs not on the boundary are scaled against a convex combination of the DMUs on the boundary facet closest to them. DEA is used in a wide range of conditions, such as software projects (Mahmood *et al*., 1996), information technology investments (Shao and Lin, 2002), technology commercialization projects (Sohn and Moon, 2004), EDI controls (Lee *et al*., 2005), Internet companies (Serrano-Cinca *et al*., 2005), service delivery procedures(Seol *et al*., 2007),and supplier evaluation and selection (Çelebi and Bayraktar, 2008).

In order to ensure successful audits, organizations may maintain a minimized list of controls that is manageable and easy to understand (Bakman, 2007). For instance, network protection, and virus protection and support procedures are the two most important controls for a small business (Busta and Strong, 2006). Ott *et al*. (2008) argued that auditors can use data mining techniques such as statistical modeling to uncover designs that can help organizations identify procedure improvements, recognize cheating and improve risk management. As auditing depending on data mining may need additional resources to ensure management data analysis for a continuous assurance procedure, auditors may plan sufficiently and have a reasonable position before embarking on a data mining exercise. Decision trees are a rapid and effective method of classifying data set entries and can offer good decision support potentials. A decision tree is a tree in which each non-leaf node signify a test on an attribute of cases, each branch correlates to a consequence of the test and each leaf node signify a class forecast. The quality of a decision tree depends on both its classifying accuracy and its size. Classifying using decision trees categorizes a set of cases in a database into different classes according to a classifying model. Two kinds of data sample are used for the classifying task. A training sample is first analyzed and a classifying model is constructed depend on the characteristics available in the data of the training sample. Such a classifying model is then used to categorize a test sample. For example, we can use the classifying model learned from the existing customers' data to predict what services a new customer would like. A case in the training sample set includes of multiple attributes (independent and dependent factors) and a known class label related with them. The independent factors are represented as an attribute-value vector, $x = (\chi_1, \chi_2, \ldots, \chi_j)$. Purpose that the cases can fall into j classes, that is, $C = (c_1, c_2, \ldots, c_j)$. Then, a training sample can be indicated by $M = \{(x_m, y_m)\}$ where $x_m \in X$ (all possible attribute space) and $y_m \in C$ (all possible cases), $m = 1, \ldots, M$ (the size of the model set). On the other hand, since all the cases in a test sample have no known class levels, a test sample is indicated by $S = \{(x_\varepsilon, y_\varepsilon)\}$ where $x_s \in X$ and $y_s \in \Theta$, $s = 1, \ldots, S$ (the size of the test sample). A decision tree can be persuade that will make it possible to assign a class to the dependent factor of a new case in the test sample depend on the values of independent factors. Approaches of a decision tree depend classifying include target marketing, agitate forecast, medical diagnosis and so on. For instance, Bernstein and Provost (2001) used decision trees in the development of a knowledge discovery assistant, in order to categorize different methods used to solve a specific

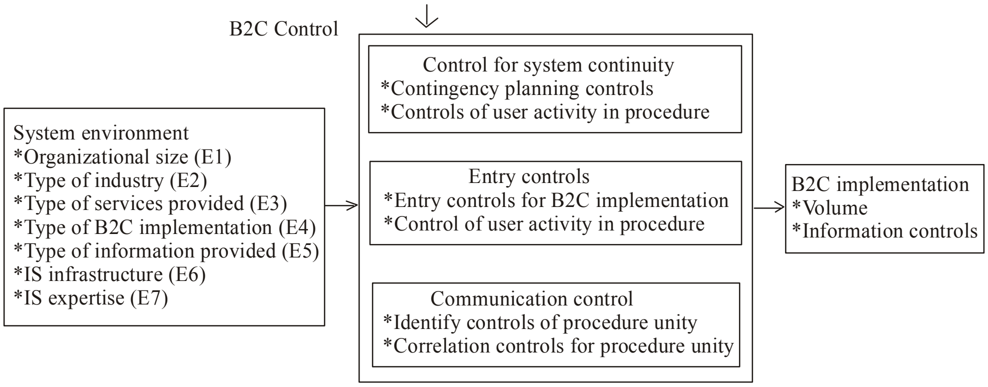


Fig. 1: System environments, B2C controls, and B2C implementation

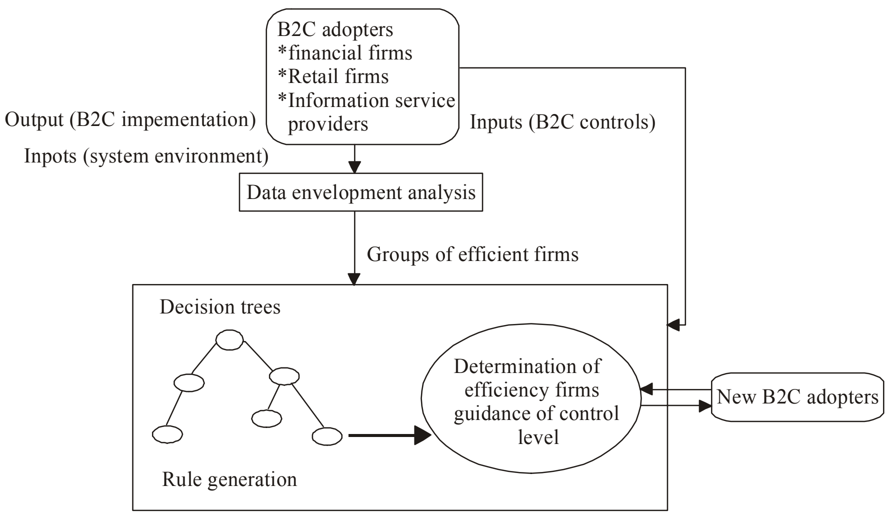


Fig. 2: The assessment process of B2C controls using DEA decision trees

problem. Endou and Zhao (2002) examined a decision tree implementation method that relied on an evolution of the training data set used. The training data set was developed to give the best coverage of the knowledge. Zmazek *et al*. (2003) used decision trees to predict radon gas concentration from other environmental factors, leading to a possible future earthquake forecast system.

## PROPOSED METHODOLOGY

The relationships among environments, B2C controls and implementation are shown in Fig. 1. The proposed framework of controls evaluation is including of two steps, as displayed in Fig. 2. The first step is to utilize DEA with multiple inputs of B2C controls and multiple outputs of B2C implementation. DEA classifies DMUs, i.e., B2C adopters, into classification of effective and unsuccessful depend on the resulting economy scores. The second step uses decision trees where factors affecting B2C controls are include of organization related factors, system related factors and infrastructure related factors. Rules for the determination of the level of controls in effective adopters are determined. The objective of B2C controls is to ensure that an organization realizes its goals through B2C approaches. If users are not sure about the superiority's of the system, B2C approaches cannot be followed and implemented; before B2C approaches are implemented, management may demand assurance that sufficient controls are in place. Thus, for B2C implementation, B2C controls may be designed effectively indicating that B2C may be considered as input of DEA model. B2C controls have three parts, i.e., controls for system continuity, entry controls and communication controls (Table 1):

- Controls for system continuity
  - Contingency planning controls (C1_1)
  - Supports of data and schedules (C1_2)
- Entry controls
  - Entry controls for B2C approaches (C2_1)
  - Controls of User activity in procedure (C2_2)
- Communication controls

Table 1: B2C controls

| Controls class | Controls | Objectives | Description |
|---|---|---|---|
| Controls for system continuity | Contingency planning controls (C1_1) Supports of data and schedule (C1_2) | Availability | Procedures for the recovery of information systems department's services follow the support of critical resources. |
| Entry controls | Entry controls for B2C implementation (C2_1) | Unity, classified | Procedures designed to ensure that access to data and programs is controls of user activity in procedure (C2_2) controlled. |
| Communication controls | Identify controls for procedureunity (C3_1) Correction controls for procedure unity (C3_2) | Unity, classified | Procedures used by company to ensure security in inbound and outbound transactions. Procedures designed to ensure that error are identified and corrected during input of data and the process of data is authorized and appropriate during communication. |

- o Informal controls for procedure unity (C3_1)
- o Correction controls for procedure unity (C3_2)

A test containing a considerable quantify of controls was used to determine the size of controls. The explanation of the first version of quantify of B2C controls were reviewed through an interview with six B2C request practitioners (one from each company); two IS professors made a final review. The answers were placed on a five-point Likert-type scale (1= disagree, 5 = agree). Average value was used for multi-item quantify. The data used in validating the research model was gathered as part of a larger examination concerning B2C controls. Implementation of B2C approaches has two aspects: volume and information contents. The implementation success of B2C approaches is defined by the size of implementation of B2C approaches, as represented by volume and information contents. Volume in B2C approaches correlate to the percentage of an organization's transactions with consumers that are handled through B2C approaches. Volume represents the proportion to which a firm's information exchange and procedure are handled through B2C approaches or need a parallel system to be performed. Information contents indicate the size to which information related to products and services is supplied effectively to customers. This concept highlights how the organization's adoption of B2C approaches improved the presentation of information. This includes the size to which search and browse functions are implemented and how product and service information is supplied to facilitate information exchange and a greater volume of information output. The guidance of B2C controls depend on the relationship between system environments and B2C controls. Lee and Han (2000) proposed factors affecting internal and external controls of inter-organizational systems in the conditions of EDI. Using the theories of innovation adoption, some researchers have argued various factors that have an effect on protection adoption such as firm size, industry type and top management support (Lee and Kozar, 2005). Industry type and organizational use of IT were regarded as the two factors that influence protection adoptions (Yeh, Chang, 2007) Previous studies suggest that the business risk level is related to the current firm concern about protection risk (Kotulic and Clark, 2004) and propose relationships among organizational factors, IS protection quantifies and IS protection effectiveness (Kankanhalli *et al*., 2003). The relationship between system environments and controls can be used to design controls (Lee and Kim, 2009). Achieving a balance among the organization, people, procedure and technology is crucial for effective information protection (Anderson, 2008). Factors affecting B2C controls are include of organization related factors, system related factors and infrastructure related factors:

- Organization related factors
- o Organizational size (E1) (number of employees)
- o Type of industry (E2) (1 = financial firms, 2 = retail firms, 3 = information service providers)
- System related factors
- o Type of services supplied (E3)
- o Type of B2C approaches (E4) (1 = shopping malls 2 = intermediaries, 3 = information service)
- Infrastructure related factors
- o IS infrastructure (E5) (Likert-type scale)
- o IS expertise (E6) (Likert-type scale)
- o Needs for IS protection (E7) (Likert-type scale)

Respondents chose one of the following states in order to determine type of services supplied (E3):

- B2C approaches that provide company information including some information about products or services.
- B2C approaches that provide company information including some information (e.g., price details) about products or services. However, only conventional purchasing is possible.
- B2C approaches that provide company information including some information (e.g., price details) about products or services and on-line purchasing facilities. However, billing occurs conventionally.
- B2C approaches that provide information on specific business purpose.
- Others.

Table 2: The industry distribution of companies adopting B2C implementation

| | Financial firms | Retail firms | Information service providers | Total |
|---|---|---|---|---|
| No. of companies | 24 | 40 | 20 | 84 |
| Percent | 25.9 | 51.2 | 22.9 | 100 |

**Data description:** The main data collection method was respondents from B2C approaches adopters. The data was collected as part of a larger study on B2C controls (Lee and Kim, 2009). A test is used as a guide for evaluating the status of B2C controls, system environments and implementation of B2C approaches. This will develop understanding of where protection needs improvement (Bakshi, 2004). The sample data were filled with past cases that were collected from interviews and discussions with IS personnel. The sample data include of 84 companies that had successfully implemented B2C approaches; data were drawn from a population of more than 300 companies in Iran that have followed B2C approaches. The population covers a wide range of businesses that have begun business-to-consumer e-commerce. Quantifies for unstable were adapted from related literature. They were quantified on five-point Likert-type scales. One or two members of the B2C approach staff or management took part in the interview. They were believed to have sufficient knowledge about their implementation. The data was collected as part of a larger examination concerning B2C controls. One goal of this study was to analyze the differences in control economy among firms in different industries. This study uses three groups of firms, i.e., financial firms, retail firms and information service providers. The number of adopters of B2C approaches in each group is shown in Table 2. The total number of firms in the sample is 24 financial firms, 40 retail firms and 20 information service providers. DMU is an individual B2C approaches adopter.

## RESULTS AND DISCUSSION

The unstable output and Reliability and validity test of controls was performed for the collected data. The Cronbach's alphas are shown in Table 3. All scales exceed 0.5, which shows moderate to high reliability. The content validity of the items was created through the adoption of constructs that have been validated by other researchers and through a pretest with six IS professional. This study adapts to quantify used by previous studies and pretests them with practitioners and experts to increase the content validity of the instrument. Thanassoulis and Emrouznejad (1996) was used DEA software to code the DEA model in the PC version. In order to analyze the economy of firms adopting B2C controls, this study used the scale the DEA model as a means for radial improvement and continuous returns. This model provides a radial improvement in both inputs and outputs. DEA enabled a series of analyses of the differences in the economy in various combinations of inputs and outputs and the degree of reduction needed in specific modes of controls. DEA identified effective and unsuccessful B2Capproaches adopters when all six unstable of the B2C controls and two unstable of the implementation of B2C approach and performance were used. Separate economy analyses were used for financial firms, retail firms and information service providers. Table 4 shows the average economy of information service providers and retail firms are higher than that of financial firms.

Table 5 shows the number of effective firms and the average economy when specific classes of input unstable were used. Input unstable was divided into

Table 3: Measurement properties of factors and controls

| Controls | Items | Mean | Individual item reliability |
|---|---|---|---|
| Contingency planning controls (C1_1) | C1_1_1 | 4.7 | 0.78 |
| | C1_1_2 | 3.9 | |
| Supports of data and schedule (C1_2) | C1_2_1 | 4.6 | 0.68 |
| | C1_2_2 | 5.2 | |
| Entry controls for B2C implementations (C2_1) | C2_1_1 | 4.5 | 0.85 |
| Controls of user activity in procedure (C2_2) | C2_2_1 | 3.8 | 0.75 |
| Identify controls for processing unity(C3_1) | C3_1_1 | 4.2 | 0.86 |
| | C3_1_2 | 4.2 | |
| Correction controls for procedure unity (C3_2) | C3_1_3 | 3.7 | 0.73 |
| | C3_2_1 | 4.2 | |
| Information contents (IMP3) | C3_2_2 | 3.9 | 0.80 |
| | E13_1 | 4.3 | |
| | E13_2 | 4.1 | |
| | E13_3 | 4.4 | |
| | E13_4 | 4.5 | |

Table 4: DEA efficiency results

| Type of firms | Financial firms | Retail firms | Information service providers |
|---|---|---|---|
| Number of efficient firms (%) | 2 (9.4) | 16 (26.2) | 9 (30.8) |
| Number of inefficient firms (%) | 21 (76.6) | 30 (54.7) | 13 (55.1) |
| Average economy | 60.7 | 71.6 | 75.9 |

Table 5: Average efficiency in specific class of input

| Type of firms | Financial firms | Retail firms | Information service providers |
|---|---|---|---|
| Controls for system continuity | 54.5 (3) | 67.8 (6) | 64.8 (4) |
| Entry controls | 47.8 (0) | 50.4 (9) | 55.9 (4) |
| Communication controls | 48.6 (1) | 53.4 (11) | 59.7 (6) |

Input class columns. Number in parenthesis indicates the number of efficient firms; Peer scaling method is contribution to targets

Table 6: Average efficiency of firms

| Input class | Input variables | Output variable | Financial firms Mean | Retail firms Mean | Information service providers Mean |
|---|---|---|---|---|---|
| Controls for system continuity | Contingency planning controls | Volume | 23.6 (1) | 30.4 (2) | 31.4 (1) |
| | | Information contents | 43.7 (0) | 47.8 (1) | 50.8 (2) |
| | Supports of data and schedule | Volume | 16.7 (0) | 26.8 (3) | 24.9 |
| | | Information contents | 32.0 (0) | 52.3 (3) | 44.8 (0) |
| Entry controls | Entry controls for B2C implementation | Volume | 12.9 (0) | 28.9 (3) | 18.9 (0) |
| | | Information contents | 12.9 (0) | 23.8 (3) | 20.4 (0) |
| | Controls of user activity in procedure | Volume | 17.5 (0) | 27.4 (2) | 28.4 (1) |
| | | Information contents | 23.9 (0) | 28.3 (2) | 40.1 (1) |
| Communication controls | Detection controls for procedure unity | Volume | 12.5 (0) | 27.8 (1) | 31.0 (3) |
| | | Information contents | 41.2 (1) | 23.5 (1) | 40.0 (1) |
| | Correction controls for procedure unity | Volume | 13.2 (0) | 24.8 (1) | 29.3 (2) |
| | | Information contents | 17.8 (0) | 18.9 (2) | 22.8 (1) |

Number in parenthesis indicates the number of efficient firms

Table 7: Test of efficiency difference between firms

| Input class | Financial firms-retail firms | Financial firms-information service providers | Retail firms-information service providers |
|---|---|---|---|
| Controls for system continuity | -8.5 (-2.35, 0.012**) | -8.7 (-2.10, 0.028**) | 0.2 (0.06, 0.881) |
| Entry controls | -6.8 (-1.78, 0.072*) | -9.4 (-2.13, 0.034**) | -2.6 (-0.52, 0.586) |
| Communication controls | -5.9 (-1.22, 0.123) | -10.1 (-2.30, 0.020**) | -4 (-0.90, 0.350) |

Values in the parenthesis indicate t-value and significance (*: $p<0.1$; **: $p<0.05$; ***: $p<0.01$)

three classes of controls, controls for system continuity,entry controls and communication controls. The output variable was the implementation of B2C approaches (volume, information contents). The average efficiencies were produced when unstable of input and output were used (Table 6). As the statistical significance of the differences among approaches needs to be shown, the average efficiencies were produced and the paired WilkoxonTest was performed to see whether average economy was different (Table 7 and 8). The retail firms and information service providers develop B2C controls more effectively than financial firms do. Controls for system continuity are developed more effectively than entry controls and communication controls. The results indicate that the level of controls in retail firms and information service providers is considered as more suitable than that in financial firms. The size of controls in financial firms is more uncontrolled than demanded by system needs. Due to the nature of financial transactions involving cash flow needs control for financial firms are being stronger; the level of controls in financial firms may be higher compared to other industries under the condition of the same size of B2C implementation. Users in the sample, on the other hand, due to their belief in controls, are less worried about the size of controls in financial firms than those in other industries. Therefore it would be unsuccessful to implement expensive controls in the system if the sensitivity and safety of the system were observed to be high. B2C controls may be suitable implemented for the protection and unity of the system using limited resources and expertise. Thus the economy is dependent on the nature of users' conception of controls and users' belief in controls; the greater the belief in controls, the lower the economy of B2C controls. In two industries, controls for system continuity are implemented less effectively than those for entry. Communication controls are developed less effectively than entry controls in financial firms while controls for system continuity are implemented less effectively than communication controls in financial and retail firms. Thus, in financial firms, controls for system continuity, communication controls and entry controls, in dropping order, are effectively created in B2C approaches. This indicates the relative importance of the two classes of controls in financial firms. The implementation of controls for system continuity is not effective due to the low observed importance placed on these controls. Controls for system continuity, such as contingency planning and supports of data and schedules, are not crucial in financial firms compared to other firms. Table 9 shows the slack of controls for the 24 financial firms (the number of effective financial firms is two). The uneconomically of each

Table 8: Test of efficiency difference between controls

| Input class | Controls for system continuity (entry controls) | Controls for system continuity (communication controls) | Entry controls (communication controls) |
|---|---|---|---|
| Financial firms | 10.3 (5.00, 0.000***) | 6.9 (2.79, 0.005***) | -3.2 (-2.50, 0.016**) |
| Retail firms | 12.8 (4.24, 0.000***) | 10.6 (2.93, 0.004***) | -2.7 (-1.41, 0.132) |
| Information service providers | 9.8 (2.34, 0.021**) | 5.2 (1.11, 0.255) | -5 (-1.42, 0.150) |

Values in the parenthesis indicate t-value and significance (*: p< 0.1; **: p<0.05; ***: p<0.01

Table 9: Slack analysis of controls for financial firms

| Firm | Efficiency | Contingency planning controls | Supports of data and schedule | Entry controls for B2C implementations | Controls of user activity in procedure | Identify controls for procedure unity | Correction controls for procedure unity |
|---|---|---|---|---|---|---|---|
| 71 | 37.87 | 3.8 | 4.1 | 4.89 | 3.96 | 3.99 | 4.95 |
| 68 | 38.76 | 4.0 | 4.2 | 4.97 | 4.20 | 4.00 | 2.30 |
| 66 | 46.67 | 2.4 | 4.2 | 4.97 | 3.98 | 3.30 | 4.98 |
| 80 | 52.20 | 2.2 | 2.3 | 3.30 | 3.10 | 3.10 | 3.10 |
| 75 | 56.66 | 2.3 | 2.2 | 3.10 | 2.00 | 2.10 | 3.20 |
| 62 | 57.03 | 2.0 | 2.1 | 3.30 | 4.10 | 3.20 | 3.20 |
| 74 | 58.98 | 1.4 | 3.2 | 4.20 | 3.30 | 2.20 | 4.30 |
| 87 | 59.67 | 2.1 | 2.1 | 2.20 | 2.20 | 2.10 | 2.10 |
| 73 | 59.87 | 1.4 | 2.1 | 2.10 | 1.30 | 2.30 | 3.30 |
| 88 | 60.34 | 2.2 | 2.1 | 3.50 | 2.40 | 2.10 | 2.00 |
| 77 | 60.96 | 2.2 | 2.3 | 3.30 | 1.30 | 2.60 | 2.20 |
| 76 | 61.04 | 2.1 | 2.1 | 1.50 | 1.20 | 2.10 | 3.20 |
| 89 | 61.45 | 1.3 | 2.1 | 2.10 | 2.50 | 2.10 | 2.30 |
| 78 | 63.69 | 1.1 | 2.2 | 2.00 | 2.20 | 1.60 | 3.30 |
| 70 | 65.65 | 2.1 | 2.0 | 3.10 | 2.30 | 1.50 | 1.20 |
| 82 | 67.99 | 1.2 | 1.3 | 1.30 | 1.30 | 1.60 | 1.30 |
| 83 | 68.87 | 1.2 | 1.3 | 2.40 | 3.20 | 2.00 | 1.20 |
| 44 | 70.94 | 1.4 | 1.1 | 2.10 | 1.30 | 1.60 | 1.40 |
| 81 | 70.79 | 1.0 | 1.3 | 2.20 | 2.30 | 1.50 | 2.00 |
| 67 | 75.40 | 1.1 | 1.2 | 1.40 | 0.70 | 3.50 | 4.20 |
| 86 | 76.92 | 1.5 | 0.9 | 0.90 | 1.50 | 0.80 | 0.20 |
| 69 | 80.55 | 2.1 | 0.9 | 2.30 | 2.50 | 3.00 | 0.30 |
| 91 | 90.73 | 0.4 | 0.3 | 1.30 | 1.50 | 1.80 | 1.10 |
| 79 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |

Input: B2C controls; output: B2C implementation; Firms (DMUs) are ordered in order of efficiency

configuration of controls is supplied in Table 9. The amount of decrease in inputs shows potential improvements of the unsuccessful DMUs without worsening the other inputs or outputs. Every company can identify the level of relative amount of reduction in each part of controls in order to make the control system effective. As an example, the needed reduction of controls is greater than therefore firm 80 in controls, C2_1, C2_2, C3_1, C3_2 (Table 1). For Firm 96, C2_1, C2_2, C3_1need to be reduced more than three. The usage level for these controls may be reduced to a greater size than that for the other controls. Firm 84 needs a reduction of less than 1.0 in every input variable except C1_2. On average, DEA analysis shows that the controls for system continuity need more reduction of the usage level than entry controls and communication controls do. B2C request adopters may increase the usage level of entry and communication controls less than controls for system continuity. The sources of uneconomical may be identified to provide a direction for controls evaluation. Internal auditors may be interested in identifying the specific mode or part of controls that adversely affect implementation of B2C approaches or performance. Economy of B2C approaches systems can be decreased by reducing the use of unsuccessful controls and selecting the resource-minimizing mix. A popular tree building algorithm is Quinlan (1992) ID3 (Iterative Dichotomizer 3). The tree building procedure starts by selecting an attribute to place at the root node and at each succeeding level the subsets generated by preceding levels are further partitioned until the procedure reaches a relatively homogenous terminal node or leaf node including of a majority of the examples in a single class. An extension of ID3 includes Quinlan's C4.5 and C5, which model both discrete and continuous unstable (1992). Additional change includes handling of missing values, reducing of the decision tree and rule derivation. This study used a C4.5 learning scheme implemented using Visual Basic in an Excel spreadsheet. There are two criteria for growing the tree and breaking anode: minimum node size, maximum purity (% of records in the node with majority class) and maximum depth. This study sets these criteria as 4 records, 100% and 20, respectively. This setting allows the sufficient training of the decision tree. The training of decision trees using different criteria does not affect the classifying performance of the decision tree. Further, the purpose

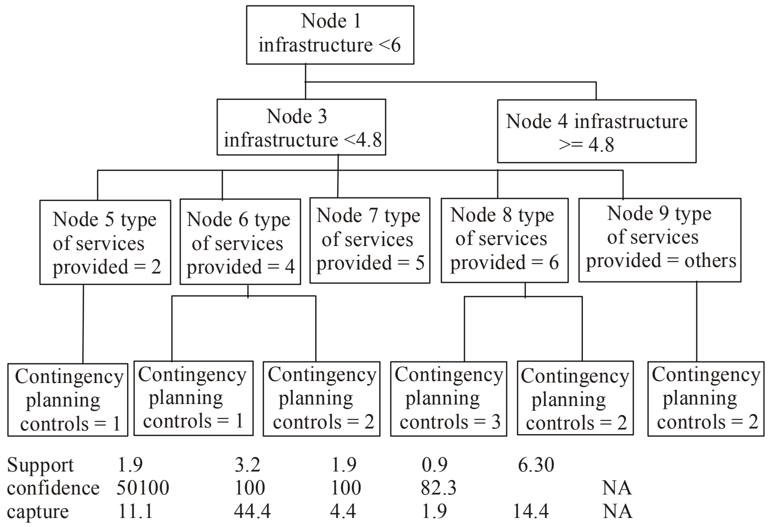


Fig. 3: An illustrative example of decision trees for recommendation of controls level, NA: Not applicable

Table 10: Rules generated from decision tree A and B. 1) Tree A is to determine efficient firms, 2) Tree B is to recommend the level of B2C controls (contingency planning controls)

| Trees | Rules | Contents | Support (%) | Confidence (%) | Catch (%) |
|---|---|---|---|---|---|
| Tree A | Rule 1 | If type of services provided = 2 then not efficient | 2.60 | 63.7 | 2.40 |
| | Rule 2 | If requirements for IS security ≥5.98 then not efficient | 49.4 | 81.3 | 56.97 |
| | Rule 3 | If type of information provided = 2 then not efficient | 12.87 | 91.9 | 15.7 |
| | Rule 4 | If type of information provided = 3 then not efficient | 50.4 | 72.1 | 48.3 |
| | Rule 5 | If type of services provided = 5 then not efficient | 6.70 | 53.1 | 4.87 |
| | Rule 6 | If type of services provided = 6 then not efficient | 29.5 | 90.2 | 35.2 |
| | Rule 7 | If organizational size b 8 then efficient | 6.70 | 69.4 | 14.5 |
| | Rule 8 | If organizational size ≥8 then not efficient | 91.3 | 74.6 | 93.4 |
| | Rule 9 | If organizational size b 800 then not efficient | 81.8 | 69.6 | 78.8 |
| Tree B | Rule 1 | If IS infrastructure ≥5 then contingency planning controls = 2 | 21.2 | 92.8 | 40.6 |
| | Rule 2 | If IS infrastructure <4.17 then contingency planning controls = 2 | 28.6 | 58.4 | 41.2 |
| | Rule 3 | If IS infrastructure <5 then contingency planning controls = 2 | 45.2 | 55.8 | 65.7 |
| | Rule 4 | If type of information provided = 2 then contingency planning controls = 3 | 10.1 | 55.3 | 12.0 |
| | Rule 5 | If type of information provided = 3 then contingency planning controls = 2 | 51.7 | 50.7 | 65.7 |
| | Rule 6 | If type of services provided = 5 then contingency planning controls = 3 | 46.1 | 50.8 | 49.9 |
| | Rule 7 | If IS infrastructure ≥5 then contingency planning controls = 3 | 51.8 | 71.9 | 73.9 |

Tree A: Tree for determination of efficient firms, # of nodes = 30, # of leaf nodes = 20, and number of levels = 5, % of misclassified = 7.1% (training data), 32.2% (test data); Tree B: tree for guidance of controls level, # of nodes = 27, # of leaf nodes = 16, number of levels = 7, % of misclassified = 6.2% (training data), 30.0% (test data)

of this study is to use a decision tree and extract rules in controls guidance rather than to obtain a high classifying ratio (Fig. 3). Two responses are keep out from further analysis because they consist of minority group in the classifications of type of services supplied and it is difficult to train decision trees with these responses. Numbers of training data and test data are 84and 10, respectively. Test data are randomly selected from the sample. Table 10 shows the rules for tree A and B:

- Tree A is to determine effective firms
- Tree B is to suggest the level of B2C controls (contingency planning controls)

Tree A has 25 nodes, 15 leaf nodes, and five levels. In tree A, percent of misclassified is 7.1% for training data and 32.2% for test data. Tree B has 22 nodes, 11 leaf nodes and 7levels. Tree B has 6.2 and 30.0% of misclassifying for training data and test data, respectively. Three quantifies of rules are suggested in Table 10. Support, Confidence and Catch. Support is the percent of training data for which the Left Hand Side (LHS) of the Rule is true. If for an observation the LHS of the rule is true, the rule uses for that observation. This quantifies how widely appropriate the rule is. Confidence is the percent of data out of the training data for which the (LHS) of the Rule is true for which the Right Hand side is also true. This means the percent of data for which the rule is true out of the data in which the rule uses. This quantifies the accuracy of the rule. One minus confidence is equal to the percent of the miss classifying. Catch is the percent of cases correctly catcher by this rule out of the data that satisfies the Right Hand Side of Rule (RHS). This is more of a reflection of the structure of the problem. A rule with catcher close to 100% indicates that the rule has been able to catch that part of the predictor space very well. If there is a rule with catcher close to 100%, that means, in the predictor space, all observations with this class sit close to each other. For example, consider a rule: If IS expertise ≥5.8 and IS infrastructure <6 then Contingency planning controls = 3. In the conditions of the above rule, the quality metrics are explained as follows. There are 84 cases in the training data, out of which 38 cases satisfy the Right Hand Side of Rule (RHS) (i.e., contingency planning controls = 3). The above rule uses to four cases and out of these, four cases satisfy RHS. Then, Support is 3.2% (= 4/95) and Confidence is 100% (= 4/4). Catch is 6.3% (= 4/47). There are 84 cases in the training data, out of which 38 cases satisfy the Right Hand Side of Rule (RHS) (i.e., Contingency planning controls = 2). The above rule uses to six cases and out of these, five cases satisfy RHS. Then, Support is 6.3% (= 6/95) and Confidence is 82.3% (= 5/6). Catch is 14.4% (= 6/47). Percent of misclassifying 1 minus Confidence or 17.7% (100-82.3%).

## CONCLUSION

The proposed methods can result in a greater return on the auditor's time and expense because the company need not request business management to supply unnecessary supporting documentation. After the analysis is completed for the first time, future evaluations of controls will need much less effort as the completed test creates a baseline. DEA can validate the economy of controls for the implementation of B2C approaches. The scientifically and reasonable used approach may lessen the burden of retrofitting protection quantifies, saving resources in areas of auditing and protection management. DEA can help identify firms adopting controls effectively and can signal unsuccessful controls that need to be reduced to be effective. Protection contraventions of web-depend on B2C approaches, which approaches are often united with mission-critical financial reporting approaches in Enterprise Resource Planning (ERP), are increasing at an alarming rate. Therefore the controls are in place and effective is important and this can be given through control evaluation. Firms, however, cannot install every possible control; as such a strategy is not economically practical. Considering the economy of controls, the size of B2C controls can be adjusted in relation to the size of the implementation of B2C approaches. Then, system environments can be examined in order to suggest the suitable level of controls. This study conducts a controls evaluation in two parts: determination of economy of B2C controls using DEA and guidance of level of controls using decision trees. Considering the protection incidences, limited internal audit resources and legal needs, the proposed methods will create action plans to define what controls must be introduced, increased or removed. The action plan is demanded for controls that are rated as limited, deficient or surplus. Actual and objective quantifies of information protection performance such as data loss and number of fraud transactions are good output unstable. At this time, the study sample does not include those unstable. It will be better to include them in the sample in a future study. The study further examined the economy differences among industries and control classes, performing multi-industry and multi-class controls examination to determine the economy of B2C controls. The results of the examination of the DEA model indicate that retail firms and information service providers implement B2C controls more effectively than financial firms do. Controls for system continuity are implemented more effectively than entry controls. In financial firms, controls for system continuity, communication controls and entry controls, in a dropping order, are effectively created in B2C approaches. Decision trees support auditors by creating two kinds of rules, i.e., rules for determining firms that effectively implemented B2C controls and rules for arguing the level of B2C controls. This study does not evaluate the impact of controls on risk reduction and the relation between risks and controls of B2C approaches may be further examined in order to increase our understanding of B2C control strategy and to mitigate protection risks. Further, the study results may help practitioners such as auditors and protection administrators by providing a systematic approach for the determination of the mode and level of controls in the procedure of implementation of B2C approaches. Every company can determine the relative amount of reduction in each part of controls in order to make the control system effective. Auditors can arrange controls with system environments and implementation of B2C approaches. This allows the control strategy to achieve a balance with organization, system, infrastructure related factors and utilization of B2C approaches. This study of the economy of B2C controls has significant suggestions for researchers and practitioners. As studies that examine economy in controls and protection have been lacking, this study provides insight to researchers by suggesting an overall normative approach to the evaluation and design of controls in terms of economy analysis of controls and decision analysis framework to evaluate the fit between system environments and controls.

## REFERENCES

Anderson, K., 2008. Business model for information security. Inform. Syst. Cont. J., 3: 51-52.

Bakman, A., 2007. If compliance is son critical, why are we still failing audits? Inform. Syst. Cont. J., 5: 37-40.

Bakshi, S., 2004. Control self-assessment for information and related technology. Inform. Syst. Cont. J., 1: 55-62.

Bernstein, A. and F. Provost, 2001. An intelligent assistant for the knowledge discovery process. New York University - Leonard Stern School of Business, Center for Digital Economy Research, CeDER Working Paper # IS-01-01, 2001.

Busta, B. and J. Strong, 2006. Expert consensus on the top IT controls for a small business. Inform. Syst. Cont. J., 6: 22-24.

Cavusoglu, H., B. Mishra and S. Raghunathan, 2004. A model for evaluating IT security investments. Commun. ACM 47(7): 87-92.

Çelebi, D. and D. Bayraktar, 2008. An integrated neural network and data envelopment analysis for supplier evaluation under incomplete information. Exp. Syst. Appl., 35(4): 1698-1710.

Cerullo, M.J. and V. Cerullo, 2005. Threat assessment and security measures justification for advanced IT networks. Inform. Syst. Cont. J., 1: 35-43.

Endou, T. and Q.F. Zhao, 2002. Generation of comprehensible decision trees through evolution of training data. Proceedings of IEEE Congress on Evolutionary Computation (CEC), pp: 1221-1225.

Gordon, L.A. and M.P. Loeb, 2002. The economics of information security investment. ACM Trans. Infor. Sys. Security, 5(4): 438-457.

Jansen, K., 2002. Siebel's e-business application and controls. Inform. Syst. Cont. J., 2: 43-50.

Kankanhalli, A., H.H. Teo, K.K. Tan and B.C.Y. Wei, 2003. An integrative study of information systems security effectiveness. Int. J. Inform. Manag., 23: 139-154.

Kotulic, A.G. and J.G. Clark, 2004. Why there aren't more information security research studies. Inform. Manag., 41: 597-607.

Lee, S. and I. Han, 2000. The impact of organizational contexts on EDI controls. Int. J. Account. Inform. Syst., 1: 153-177.

Lee, S. and K. Kim, 2009. Using case based reasoning for the design of controls for internet based information systems. Exp. Syst. Appl., 36(3): 5582-5591.

Lee, S., K. Lee and I. Kang, 2005. Efficiency analysis of controls in EDI applications. Inform. Manag., 42(3): 425-439.

Lee, Y. and K.A. Kozar, 2005. Investigating factors affecting the adoption of anti-spyware systems. Commun. ACM, 48(8): 72-78.

Mahmood, M.A., K.J. Pettingell and A.I. Shaskevich, 1996. Measuring productivity of software projects: A data envelopment analysis approach. Decision Sci., 27(1): 57-80.

Ott, J., A. MacLeod and K.M. Fan, 2008. Computer-assisted audit techniques: Value of data mining for corporate auditors. Inform. Syst. Cont. J., 3: 45-48.

Pareek, M., 2006. Optimizing controls to test as part of a risk-based audit strategy. Inform. Syst. Cont. J., 2: 39-41.

Pareek, M., 2007. Automating controls. Inform. Syst. Cont. J., 3: 45-47.

Quinlan, J.R., C4.5, 1992. Programs for Machine Learning. Morgan Kaufman, San Mateo, CA.

Samoilenko, S. and K.M. Osei-Bryson, 2008. Increasing the discriminatory power of DEA in the presence of the sample heterogeneity with cluster analysis and decision trees. Exp. Syst. Appl., 34(2): 1568-1581.

Seol, H., J. Choi, G. Park and Y. Park, 2007. A framework for benchmarking service processusing data envelopment analysis and decision tree. Exp. Syst. Appl., 32(2): 432-440.

Serrano-Cinca, C., Y. Fuertes-Callén and C. Mar-Molinero, 2005. Measuring DEA efficiency inInternet companies. Decision Supp. Syst., 38(4): 557-573.

Shao, B.B.M. and W.T. Lin, 2002. Technical efficiency analysis of information technology investments: A two-stage empirical investigation. Inform. Manag., 39(5): 391-401.

Sohn, S.Y. and T.H. Moon, 2004. Decision Tree based on data envelopment analysis for effective technology commercialization. Exp. Syst. Appl., 26(2): 279-284.

Thanassoulis, E. and A. Emrouznejad, 1996. Warwick Windows DEA Version 1.02: User's Guide, Warwick Business School. University of Warwick, Coventry, United Kingdom, First Print.

Zmazek, B., L. Todorovski, S. Dzˇeroski, A. Vaupoticˇ and I. Kobal, 2003. Application ofdecision trees to the analysis of soil radon data for earthquake prediction. Appl. Radiat. Isotop., 58: 697-706.