Omega 41 (2013) 131–142

Contents lists available at SciVerse ScienceDirect

Omega

journal homepage: www.elsevier.com/locate/omega

Using Data Envelopment Analysis (DEA) for monitoring efficiency-based performance of productivity-driven organizations: Design and implementation of a decision support system

Sergey Samoilenko [a,*], Kweku-Muata Osei-Bryson [b]

[a] Department of CIS/CS, Virginia Union University, Richmond, VA23220, USA
[b] Department of Information Systems, Virginia Commonwealth University, Richmond, VA 23284, USA

ARTICLE INFO

Available online 6 February 2012

Keywords:
Productivity
Dynamic environments
Efficiency
DEA
DSS
Data mining
Clustering
Decision Tree

ABSTRACT

The competitive nature of the business environment requires the productivity-driven organization to be aware of its relative level of effectiveness and efficiency vis-à-vis its competitors. This suggests the need, first, for an effective mechanism that allows for discovering appropriate productivity models for improving overall organizational performance, and, second for a feedback-type mechanism that allows for evaluating multiple productivity models in order to select the most suitable one. In this paper our focus is on organizations that consider the states of their internal (e.g., possibly exemplified by *resource-based view*) and external (e.g., possibly exemplified by *positioning*) organizational environment in the formulation of their strategies. We propose and test a DEA-centric Decision Support System (DSS) that aims to assess and manage the relative performance of such organizations.

1. Introduction

Modern organizational entities typically operate in dynamic, competitive environments. Within this context, the critical issues of organizational survival and advancement often lead to calls for improvements in the levels of effectiveness and efficiency [64]. However, due to the relativity of the concepts of efficiency and effectiveness, productivity-driven organizations must take into consideration the performance of their competitors. For the dynamic nature of the business environment will cause the levels of performance of competing organizations to change over time, and if the efficiency of the competitors has improved, then a productivity-driven organization must respond with its own improvements in efficiency.

Although some improvements in productivity do not require any drastic structural transformations but simply call for a gradual type of improvements in the level of performance (e.g., TQM, BPI, etc.), significant changes in the levels of effectiveness and efficiency often require structural reorganizations (e.g., ERP, BPR, etc.) that could result in periods of unstable behavior, which, if not managed, could escalate and become chaotic [52]. Resultantly, in a dynamic business environment any static model that is used to describe the relationship between inputs and outputs will have limited usefulness and feasibility in periods of instability. This suggest the need, first, for an effective mechanism that allows for discovering appropriate productivity models for improving overall organizational performance [24] and, second for a feedback-type mechanism that allows for evaluating multiple productivity models in order to select the most suitable one.

The overall goal of this investigation is to propose and test a Decision Support System (DSS) that aims to assess and manage the relative performance of organizations. We focus on organizations that consider the states of their internal (e.g., possibly exemplified by *resource-based view*) and external (e.g., possibly exemplified by *positioning*) organizational environment in the formulation of their strategies, such that the achievement of an organizational goal is dependent on the level of performance that is commonly measured in terms of the levels of the efficiency of utilization of inputs, effectiveness of the production of outputs, and efficiency of conversion of inputs into outputs. This suggests that an important component technique of our DSS is Data Envelopment Analysis (DEA), which is widely used by researchers and practitioners for the purposes of measuring productivity and relative performance [74, 7, 17, 15, 73, 26, 63]. However, other techniques are also required for providing answers to several questions that are relevant to the organization's search for the productivity model that is most suitable with respect to survival and advancement. In this investigation we focus on the following questions related to system requirements:

We present our investigation as follows. Part One outlines the functionality and composition of the proposed system. Part Two

* Corresponding author. Tel.: +1 804 342 7381.
E-mail addresses: svsamoilenko@vuu.edu (S. Samoilenko), KMOsei@vcu.edu (K.-M. Osei-Bryson).

0305-0483/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.omega.2011.02.010

offers an overview of the structural elements of the proposed DSS. Part Three outlines the design of DSS. Part Four offers an illustrative example of the DSS in action. A brief conclusion follows.

2. The functionality and composition of the DSS

The dynamic nature of the business environment suggests the presence of a concept that is central to a productivity-driven organization, namely, that of the *superior stable configuration*. Given the goal of achieving a high level of efficiency of conversion of inputs into outputs, a superior stable configuration in the context of a productivity-driven organization may imply *a model of conversion of inputs into output (input–output model) characterized by a high level of efficiency*. Consequently, we put forward the following propositions:

Proposition 1. *Stability of the performance of a productivity-driven organization is dependent on the presence of the stable input–output model.*

Proposition 2. *Accomplishment of the organizational goal of a productivity-driven organization is dependent on the creation and implementation of a stable input–output model characterized by the high level of efficiency.*

Proposition 3. *In order to monitor performance of a productivity-driven organization, DSS must be able to create and identify superior stable configurations, represented by the input–output models characterized by the high level of efficiency.*

We suggest that the design of the proposed DSS must include two sets of functionalities: *externally-oriented*, and *internally-oriented*. The *externally-oriented* functionality of this DSS is directed towards evaluating the external competitive environment of a productivity-driven organization, as well as identifying the differences between the current state of the organization and the states of its competitors. The *internally-oriented* functionality, on the other hand, is directed towards the optimization of the level of productivity of the organization, as well as towards an identification of the factors impacting the efficiency of the input–output process. We suggest that such a DSS could be implemented using a combination of parametric and non-parametric data analytic and data mining techniques including Data Envelopment Analysis (DEA), Cluster Analysis (CA), Decision Tree (DT), Neural Networks (NN), and Multivariate Regression (MR). The suggested functionality of this DSS is presented in Table 2.

While the five data analytic techniques that we use in the design of the proposed DSS have been utilized in IS research in a stand-alone fashion, they are also very frequently used in combination. For example, DEA is widely employed for the purpose of evaluating productivity and performance (e.g., [35, 59, 57, 6, 73, 2, 34, 24, 38]), but it has also been used to complement other data analytic techniques: cluster analysis (e.g., [61, 30, 36, 41]), neural network induction (e.g. [54, 10, 27, 42, 71]), decision tree induction (e.g., [55, 53, 71]), regression analysis (e.g., [19, 6, 47, 56]), and other methods([37, 26, 50]).

3. Overview of the structural components of the DSS

3.1. Data Envelopment Analysis (DEA)

Data Envelopment Analysis (DEA) is a nonparametric method for measuring the efficiency of a *decision-making unit* (DMU). Any group of entities that receives the same set of the inputs and produces the same set of outputs could be designated as a DMU; it could be a group of people, companies, hospitals, schools, industries, or countries. To determine the relative efficiency of each DMU in the group, DEA collapses inputs and outputs defined by the model into a ratio of a single meta-input and meta-output, and uses methods of linear programming to calculate the efficiency score for each DMU, where obtained score is reflective of the performance [60, 8, 41, 74]. This comparison results in a ranking of the DMUs in terms of their *relative efficiency*, where the highest-ranking DMUs are considered relatively efficient and assigned a perfect score of 1, while the rest of the DMUs in the sample are considered to be relatively inefficient. Resultantly, DEA 'envelops' the data set with the *efficiency frontier* consisting of the relatively efficient DMUs. The two commonly mentioned orientations of DEA models are the *Input-Oriented* and the *Output-Oriented* [12]. An *Input-Oriented* model is concerned with the minimization of the use of the inputs for achieving a given level of the output [14]. A relatively efficient DMU under *input-orientation* cannot reduce its levels of inputs any further to achieve a given level of output, while the relatively inefficient DMUs (with the scores of greater than "0" but less than "1") could. An *Output-Oriented* DEA model, conversely, is concerned with the maximization of the level of the outputs per given level of inputs. A relatively efficient DMU under *output-orientation* cannot increase its levels of outputs any further while relying on a given level of inputs, while the relatively inefficient DMUs (with the scores of greater than "1") could. Thus, while in both cases a relatively efficient DMU is assigned a score of "1", a relatively inefficient DMU will receive a score of greater than "1" under *output-orientation*, and a score in the [0, 1) interval under *input-orientation*.

DEA is a flexible method [16, 65, 1, 37] that can be applied under different underlying economic assumptions about the returns to scale [58] yield different DEA models [25]. An assumption of the *constant return-to-scale* (CRS) model reflects the situation where the changes in output are in the same proportion as the changes in inputs (e.g., changes of 50% in inputs correspond to the changes of 50% in outputs), while assumptions of the *variable returns-to-scale* (VRS) model reflects increasing (e.g., changes of 25% in inputs correspond to the changes of 50% in outputs), and *non-increasing returns-to-scale* (NIRS) model reflects decreasing (e.g., changes of 50% in inputs correspond to the changes of 25% in outputs) returns to scale. We direct the interested reader to the comprehensive presentations of the theoretical underpinnings of the DEA by Cook and Zhu [18] and Cooper et al. [20].

3.2. Cluster analysis (CA)

Clustering is a popular non-directed learning data mining technique for partitioning a dataset into a useful set of mutually exclusive clusters such that the similarity between the observations within each cluster (i.e., subset) is high, while the similarity between the observations from the different clusters is low (e.g., [45, 49, 44, 67, 22]). There are different reasons for doing clustering, and one of them is to find a set of natural groups (i.e., segmentation), and the corresponding description of each group. This is relevant if there is the belief that there are natural groupings in the data. Jain et al. [32] noted that there are three approaches for assessing cluster validity: (1) external assessment which involves comparing the generated segmentation (i.e., set of clusters) with an a priori structure, typically provided by some domain experts; (2) internal assessment which attempts to determine if the generated set of clusters is "intrinsically appropriate" for the data; and (3) relative assessment which involves comparing two segmentations (i.e., two sets of clusters) based on some performance measures and measure their relative performance. Our use of cluster analysis is based on the assumption