Right Indicators of Urban Railway System: Combination of BSC and DEA Model

Somayyeh Danesh Asgari[1], Abdorrahman Haeri[2], Mostafa Jafari[3]

Received: 14.03.2017 *Accepted: 08.07.2017*

Abstract

With the expansion of cities and ever-increasing traffic dilemma closely connected to people's lives, public transportation has become one the essential needs of communities. Subway because of its benefits is an important part of our lives: alleviating urban transit pressure, high safety and reliability, mass transit capacity, low energy consumption, and low price. Therefore, its performance improvement led to increasing citizenry satisfaction seems essential. The most important point in evaluation and performance improvement is the proper selection of measures. The main purpose of this paper is to introduce a new approach for selection of right indicators. For this purpose, with respect to the cause and effect relationships in balanced scorecard, its measures are applied as input and output variables of three-stage data envelopment analysis model. At first, some indicators are supposed for each BSC's aspects and the efficiency of all stages in this basic model is computed. Then, individual inputs are considered in each stage and the efficiency of that stage is computed again in order to compare with the efficiency score of the same stage in the basic model. With interpreting of efficiency variations in each stage, appropriate measures are determined. An experimental example which contains 10 stations of Tehran subway is provided to illustrate the implementation of this model. The results indicate that efficiency of train, concurrent consideration of average density per each passenger and waiting at the station, and simultaneous consideration of average density per each passenger and the delay per trip are appropriate measures. The proposed approach in this study helps to managers and decision makers in transportation industry to recognize right indices for performance improvement.

Keywords: Urban Railway, Appropriate Measures, Balanced Scorecard (BSC), Three- Stage Data Envelopment Analysis (DEA), Transportation

---

Corresponding author Email: ahaeri@ iust.ac.ir

[1] Msc student, School of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran

[2]Assisstant Professor, School of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran

[3]Associate Professor, School of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran

1. Introduction

Transportation is one of the most important infrastructure of modern societies, which has been created to meet social and economic needs. With the expansion of cities and increasing growth of traffic volume, public transportation has become one of the vital needs of communities. On the other hand, increasing crude oil prices and air pollution have led to that the subway seems the best choice for public vehicle in urban trips. Except for the its role in alleviating urban transit pressure and unraveling urban transit dilemmas, high safety and reliability, passengers' convenience, mass transit capacity, reducing energy consumption, fast speed, and low price are other advantages of subway. In view of the importance and necessity of existence of subway in urban life, it is essential to its performance be measured. Performance measurement is a substantial issue for each system to detect strength and weakness points. In performance evaluation, we encounter different indicators and need an approach to determine which indicators are appropriate for this aim. On the other hand, in data gathering, two issues are important: first, the higher the indicator number, the greater the cost and complexity of problem would be expected, both for performance evaluation and for managers' decision making. The main purpose of this study is introducing a new approach in order to select the right indexes by integrating the balanced scorecard (BSC) and three- stage data envelopment analysis (DEA) model. So the main question of this study is raised as follows:

How can we determine the most appropriate measures regarding to efficiency and all operational perspectives of urban railway system?

The balanced scorecard model introduced by Kaplan and Norton in 1992 focuses on long- term goals instead of short- term ones. Data envelopment analysis is a non-parametric method for measuring the decision-making unit's (DMU) efficiency. It has been employed to evaluate the performance of different fields, such as, health care [Ghotbuee, Hemati and Fateminezhad, 2012], transportation [Nesterova et al., 2016], financial institutions [Titko, Stankevičienė and Lāce, 2014], hotel industry [Manasakis, Apostolakis and Datseris, 2013], education [García Valderrama, Revuelta Bordoy and Rodríguez Cornejo 2013], etc.Review of related literature suggests that performance measurement of urban railway system from different aspects, financial and non- financial, has received little attention. On the other hand, different methods for selection of right indices were surveyed which none of them have utilized DEA model to achieving this aim. Therefore, the aim of this study is to propose the first three- stage DEA-based approach to identify the most appropriate indicators for different strategic goals of urban railway systems. Our proposed approach has some advantages distinguishing this method from others such as non-interference of people's judgements in the final results, and selection of the appropriate measures in various perspectives- learning and growth, internal process, customer, and financial perspective, the most important aspects which must be considered in the organizations' performance measurement.

This paper is organized as follows: The literature review in Section 2 provides an overview of studies that measure the performance of urban railway system. The research method is given in section 3. An experimental example is presented in section 4 and its results is discussed in section 5. Finally, conclusion remarks are given in section 6 to summarize the contribution of the paper.

2. Literature Review

Performance measurement of subway has been studied in many researches. [Aydin, 2017] assessed the service quality of railway systems through investigation of passenger satisfaction. He identified the indicators need to be improved using satisfaction analysis, fuzzy trapezoidal numbers and TOPSIS. [Wanke, Barros and Figueiredo, 2016] utilized a stochastic DEA model and Beta regression to analyze the performance of different transportation modes, such as subway, in 285 cities of world during 2009- 2012. A vulnerability model was introduced by [Hong et al., 2017] to capture various levels of complementary strength between the two urban public transportation systems, bus and subway, with consideration of passengers' intermodal transfer distance preference. The results showed that the complementary relations meaningfully decrease the complementary urban public transportation systems' vulnerability. [Shen, Xiao and Wang, 2016] utilized the concept of the American customer satisfaction index and partial least squares (PLS) in order to provide an evaluation indicator system and measure the passenger satisfaction in subway system of China. [Maina, Forda and Robinson, 2016] applied an algorithm to minimize delay times and employs both safety performance functions and experimental Bayes before-after methodology to assess the impact on safety. Having use traffic data from 70 indicators signalized intersections located in Virginia Beach, they found that minimizing intersection delay time can lead to a safety improvement nearly 26.45 percent. [Aydin, Celic and Gumus, 2015] introduced a customer satisfaction framework for assessing the performance of railway system in Istanbul. They utilized fuzzy analytic hierarchy process (AHP), statistical analysis and trapezoidal fuzzy sets. Their proposed approach helps to identification of operational deficiencies and solving the complex decision making problems included vague information or uncertain data. [Mousa Ali, 2015] abridged some high-speed rail projects performed around the world such as Spain, Japan, France. They determined the most important factors and variable involved in these projects. The conclusion of their research can be useful in developing countries. [Mallikarjun, Lewis and Sexton, 2014] suggested a DEA- based model in order to evaluate the performance of railway system. They determined the resources of inefficiency using comparing the efficiency score of each decision making unit (DMU) to a suitable efficient benchmark system. [Qin, Zhang and Zhao, 2014] proposed a multi- stage network DEA model to measure the influence of organizational patterns on the efficiency of urban railway system's performance in China. The results indicated that the average overall efficiency score is relatively low which was mainly due to the financial and construction inefficiencies. [Stenström, Parida and Galar, 2014] introduced measures applied to assess performance of railway system. They compared these measures with indexes of European Standards. [Rassafi, Jafari and Javashir, 2014] evaluated sustainable urban transportation by system dynamic approach. They examined the validity of their model using real data related to city of Mashhad during the years 1994 to 2009. Their evaluation included indicator related to social, economic, and environmental aspects. [Boam, 2014] evaluated the technical efficiency of Canadian public transit services from 1990 to 1998 by bootstrap data envelopment analysis method. He discovered that peaking decreases transit efficiency and this result supported earlier studies applied data envelopment analysis. [Famurewa et al, 2014] introduced railway performance variables to link maintenance function to capacity situation using fuzzy inference system. They consider aspects of reliability, comfort, punctuality and safety. They evaluated the overall performance of five lines of Swedish Transport Administration by data between 2010 and 2012. [Powell, Gonzalen-Gil and Palacin, 2014] investigated important variables of usage of energy in the Tyneand Wear Metro system (UK). They detected that heat is the most important factor of energy consumption. [Stenström et al., 2013] developed a new approach that provided information on the performance of railway system, factors and linkage between them. They tested this model for Iron Ore Line Sweden. [Alizadeh et al., 2013] introduced factors and strategies to improve quality of Sanandaj city life by transit-oriented development. They evaluated the current conditions of the city to extracted these strategies. [Molhotra, Malhotra and Lermack, 2009] introduced some financial variables to benchmark the financial performance of seven North

American Class I freight railroad using data envelopment analysis and could identify the areas in which performance is poor and need to modify. [Åhrén and Parida, 2009] analyzed the measures involved in the performance of railway by developing a new approach evaluated the overall railway infrastructure effectiveness. They tested the verification of their model for a section of the Swedish railways. In another study, they investigated benchmarking and maintenance performance indexes for railway infrastructure. They found that benchmarking can be employed for maintenance performance indicators in order to continuous improvement [Åhrén and Parida, 2009]. [Nathanail, 2008] introduced a framework for railway operators to attend into superintendence and checking the services quality supplied to their passengers. In his proposed framework, there was six criteria contained 22 indicators. [Yu, 2008] applied classic DEA and network DEA to examine the efficiency and effectiveness of 40 railways during 2002. They presented that there is direct relationship between transportation service characteristic and performance evaluation and network DEA can identify the efficiency scores. [Lan and Lin, 2006] measured the performance of railway by stochastic input distance function with an inefficiency effect and a stochastic consumption distance function with inefficient effectiveness approaches. Also, they applied the same approaches in order to measure the performance of 39 worldwide railway systems over 8 years. They calculated the railway transport technical efficiency and service effectiveness for 44 global railways by proposing four-stage DEA [Lan and Lin, 2005]. [Lawrence and Erwin, 2003] applied data envelopment analysis model to examine the technical efficiency and service effectiveness for 76 railways in the world from 1999 to 2001. They demonstrated that railway efficiency and effectiveness depend on regions. [Costa and Markellos, 1997] introduced a new approach based on multi-layer perception neutral networks (MLPs) to assess the performance of public railway services. They surveyed this model for London Underground in the period of 1970 to 1994. On the other hand, for selection of key performance indexes, researchers have employed diverse methods. In some studies, judgment of experts is the basis of selection. Delphi method and interviews are two ways that can be implemented to achieve expert's opinion. [Varmazyar, Dehghan and Afkhami, 2016] assessed the performance of research and technology organizations through combination of MCDM and BSC model. They gathered the performance measurement indicators through BSC literature and selected final indices, 17 relevant indicators from 31 indicators, regarding to experts' viewpoint. [Shapouri and Keramati, 2015] collected BSC's criteria from literature and experts' viewpoints to draw a customer relations management strategy map by integrating BSC and MCDM models for internet service provider firm. [Shafiee, Lotfi and Saleh, 2014] employed the BSC and Data envelopment analysis (DEA) models to measure the performance of food supply chain. The efficiency measures of supply chain management were determined through literature review and expert ideas. [Tsai and Cheng, 2012] examined the key performance indicators for E-commerce and internet marketing of elderly products; they employed Delphi method along with questioner to develop 29 factors.

Some studies have utilized other methods in order to select the key indicators. [Murali and Pagazhendhi, 2016] ranked the after sales service strategies of firms involved in household appliances business through integrating BSC, AHP, and quality function deployment (QFD). [Yao and Lin, 2016] used BSC to evaluate E-government project and applied AHP to compute the weights of all indicators. [Yaghoobi and Haddadi, 2016] studied the performance of 5 functional units at information communication technology (ICT). In order to select the indicators, they implemented AHP tool. [Pan and Nguyen, 2015] recognized the key performance measurement criteria to achieve customer satisfaction via Decision Making and Trial Evaluation Laboratory (DEMATEL) method in manufacturing firms. [Chen and Tzeng, 2014] presented a performance measurement model based on hybrid dynamic MCDM- BSC model.

They determined the influential weights of BSC's aspects and indicators through ANP with DEMATEL (DEMATEL- based ANP). [Leung, Lam and Cao, 2006] presented a model through integrating Analytic Network Process (ANP) technique, AHP technique, and BSC. Using this model, they defined the relationships between perspectives of BSC and the weight of each perspective. [Wu, Tzeng and Chen, 2009] assessed the performance of bank by integrating fuzzy MCDM approach and BSC model. 23 evaluation factors were achieved from related literature of bank performance and their relative weights were calculated through fuzzy AHP.

3.Research Method

3.1 Balanced Scorecard

In 1992, the balanced scorecard was first introduced by Kaplan and Norton [Kaplan and Norton, 1992]. This model enables organizations to assess their performance from four different aspects: one financial and three non-financial perspectives. Also, it allows to managers to make causal relationships among their strategic measures based on perspectives of balanced scorecard instead of performance indicators in four separate perspectives. Strategic map was introduced to managers [Kaplan and Norton 2004] to comprehend how performance in each perspective pursue a ierarchical structure whereby improvements in learning and growth lead to better internal process which, inturn, grow the value propositions culminate in customers satisfaction, leading to financial performance ultimately. Figure 1, shows these relationships between BSC's perspectives. Kaplan and Norton were believed that the strategic relations between aspects would enable managers to examine the strategy using "if-then" propositions, so that continual improvement in each of the financial dimensions would be controlled to assess if it translated finally into financial performance. For example, an investment in information technology via an internet-based retailer (an accomplishment in learning and growth dimension) should increase sales income (a financial factor) otherwise managers receive a feedback and understand that it is necessary to the strategy map be depict again [Bento, Bento, and white, 2013].

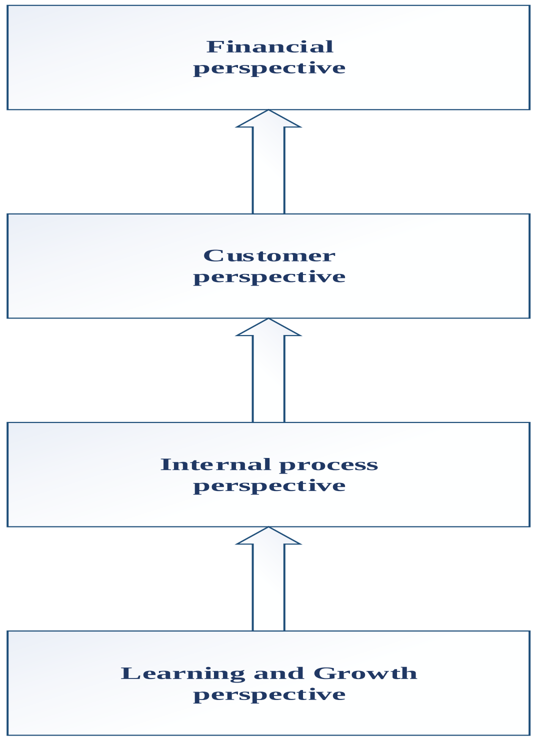

Figure 1. The causal relationships between perspectives of BSC

[Liang and Hou, 2007] recognized the cause and effect linkage between customer and financial dimensions in hotel industry but they could not observe any linkage between learning and growth and financial indexes. [Ittner and Larcker, 1998] investigated the causal linkages in the telecommunications industries. [Banker, Potter and Srinivasan, 2000] examined effects of customer satisfaction on financial performance in hospitality. They stated that 75 percent of financial services firms disregard the cause and effect relations between the four BSC perspectives. [Lucianetti, 2010] approved the negligence of those linkages among the adherents of BSC in Italy and deduced that organizations cannot be impart the all benefits of BSC's employment If they do not utilize the strategy map property. Regarding to the causal linkages in BSC model. The classic DEA model cannot be a proper alternative for our study. So, we use a cascade three-stage structure of network DEA model. Many researchers have carried out based on network-DEA that readers can refer to [Kao, 2009].

3.1 Three- Stage DEA Model

Traditional DEA model considered the process as block-boxes and use a single model to transform primary inputs to outputs [Färe and Grosskopf, 2000]. In order to achieve useful information for performance improvement, the analysis must keep off from a black-box and analyze the efficiency of decision-making sub-units. [Färe and Grosskopf, 2000] opened the black-box and recognized the resources of inefficiency in organizations by proposing the network DEA. In order to performance evaluation, two-stage DEA model was proposed by Lewis and Sexton in 2003 [Sexton and Lewis, 2003]. In 2004, they introduced a model used for DMUs include of sub-DMUs. Some of those sub-DMUs consume factors produced by previous sub-DMUs and some of them produce outputs consumed by other sub-DMUs [Sexton and Lewis, 2004]. The BSC model can suggest a proper framework to organize some interconnected DEA model because it is based on cause and effect relationships [Kaplan and Norton, 1996]. A cascade system of h processes. X_{ij} and Y_{rj} and $Z_{pj}^{(t)}$ are respectively interpreted as input, output and the *p-th* intermediate product, $p = 1,\dots, q,$ of process t, $t = 1,\dots, h\text{-}1$, for DMU_j. intermediate products are outputs of process t and inputs of process $t+1$. Also, the intermediate products of the last process h are the outputs of the system. It is assumed that the number of products is the same for all process (only for simplification). This model has been shown in Figure 2

Following equations compute the efficiency of DMU_j:

$$E_k = \max \sum_{r=1}^{s} u_r Y_{rk}$$

$$s.t.$$

$$\sum_{i=1}^{m} v_i X_{ik} = 1$$

$$\sum_{r=1}^{s} u_r Y_{rj} - \sum_{i=1}^{m} v_i X_{ij} \le 0 \quad j=1,\dots,n$$

$$\sum_{p=1}^{q} w_p^{(1)} Z_{pj}^{(1)} - \sum_{i=1}^{m} v_i X_{ij} \le 0 \quad j=1,\dots,n$$

$$\sum_{p=1}^{q} w_p^{(t)} Z_{pj}^{(t)} - \sum_{p=1}^{q} w_p^{(t-1)} Z_{pj}^{(t-1)} \le 0 \quad j=1,\dots,n\ ,\quad t=2,\dots,h-1$$

$$\sum_{r=1}^{s} u_r Y_{rj} - \sum_{p=1}^{q} w_p^{(h-1)} Z_{pj}^{(h-1)} \le 0 \quad j=1,\dots,n$$

$$u_r, v_i, w_p^{(t)} \ge \varepsilon$$
$$r=1,\dots,s\ ,\quad i=1,\dots,m$$
$$p=1,\dots,q\ ,\quad t=1,\dots,h-1$$
(1)

$W_p^{(t)}$ is associated with the *p*-th intermediate product of process *t*.

For DMU_k , the efficiency of each process is computed as follow if u_r^*, v_i^*,and $w_p^{(t)*}$ be considered as optimal multipliers:

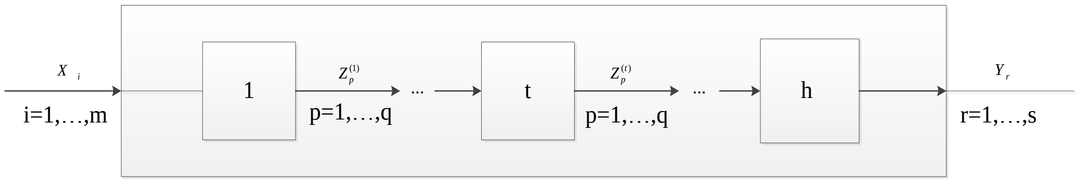

Figure 2. Cascade system source

$$E_k^{(1)} = \sum_{p=1}^{q} w_p^{(1)*} Z_{pk}^{(1)} \Big/ \sum_{i=1}^{m} v_i^* X_{ik} \tag{2}$$

$$E_k^{(t)} = \sum_{p=1}^{q} w_p^{(t)*} Z_{pk}^{(t)} \Big/ \sum_{p=1}^{q} w_p^{(t-1)*} Z_{pk}^{(t-1)} \tag{3}$$

t=2,...,h-1

$$E_k^{(h)} = \sum_{r=1}^{s} u_r^* Y_{rk} \Big/ \sum_{p=1}^{q} w_p^{(h-1)*} Z_{pk}^{(h-1)} \tag{4}$$

$E_k^{(t)}$, $t = 1, \dots, h$ is equal to $\sum_{r=1}^{s} u_r^* Y_{rk} \Big/ \sum_{i=1}^{m} v_i^* X_{ik}$ that is the efficiency of system. A DMU is called efficient if all of its processes be efficient [Kao, 2009].

3.2 Proposed Approach

The main purpose of this study is the identification of right indicators by integration of BSC and three- stage DEA model. Our proposed approach includes the following steps:

Step1. Making the Basic Model

In step 1, for each aspect of BSC, some variables are supposed and regarding to causal relationships in the BSC, these variables are supposed as inputs and outputs of three-stage DEA model. Thus, inputs of stage 1 are the variables of learning and growth perspective. And its outputs are selected from internal process dimension. As mentioned earlier, in our cascade three- stage DEA model, the inputs of each stage are identical with outputs of previous stage. Therefore, the outputs of stage 1 and inputs of stage 2 are alike. Output variables of stage 2 are the very factors of customer perspective. Ultimately, the input and output factors of stage 3 are respectively extracted from customer and financial perspectives. This three-stage DEA model is supposed as basic model. Figure 3 shows that input/ output variables of each stage are selected from which BSC's perspective Step 2. Calculation of the Basic Model's Efficiency Step3. Considering inputs individually and computing the efficiency In the next step, we consider only one of the input variables of stage 1, in the basic model, and compute its efficiency for all DMUs. Then, other input variables of stage 1 are considered and the efficiency of this stage is calculated again. This procedure repeats for stage 2 and stage 3. Step 4. Comparing Between Efficiency Scores Obtained to Basic Model's Efficiency

In this step, we compare the obtained efficiency scores of each stage in previous step to that of related stage in the basic model.

Step 5. Selection of Appropriate Measures Having compare the obtained efficiencies to primary efficiency scores in the basic model, we neglected changes of efficiency less than 0.1 while the efficiency variations equal to 0.1 or more are assumed significant. The analysis of observed results helps to specify that which input (s) is (are) appropriate measure (s) in each stage. Note that if there are more than two input variables in each stage, it is necessary to that multiplex combinations of input factors be considered. For example, if there are three inputs in one stage, it is essential to consider both individual inputs and The simultaneous use of input factors. Figure 4 displays the proposed approach by this research.

4 Experimental Results

According to the proposed approach in section 4, we investigate an empirical example in this section. We utilize the information of ten stations of urban railway in Tehran. These stations which are belonging to line 5 are respectively named Sadeghieh, Erame'-sabz, Azadi stadium, Chitgar, Iran khodro, Vardavard, Atmosphere, Karaj, Mohhamad shahr, Gol shahr. This line connects Tehran to Karaj which is one of the metropolises of Iran. Tehran- Karaj express- train (subway-line 5) was used in 1999. Because of its important role in passengers' trip, between two metropolises, we study that. The numeral data used in this section are extracted from [Sang trash, 2012] .

Figure 3. Selection of input/ output variables from measures of BSC's perspectives

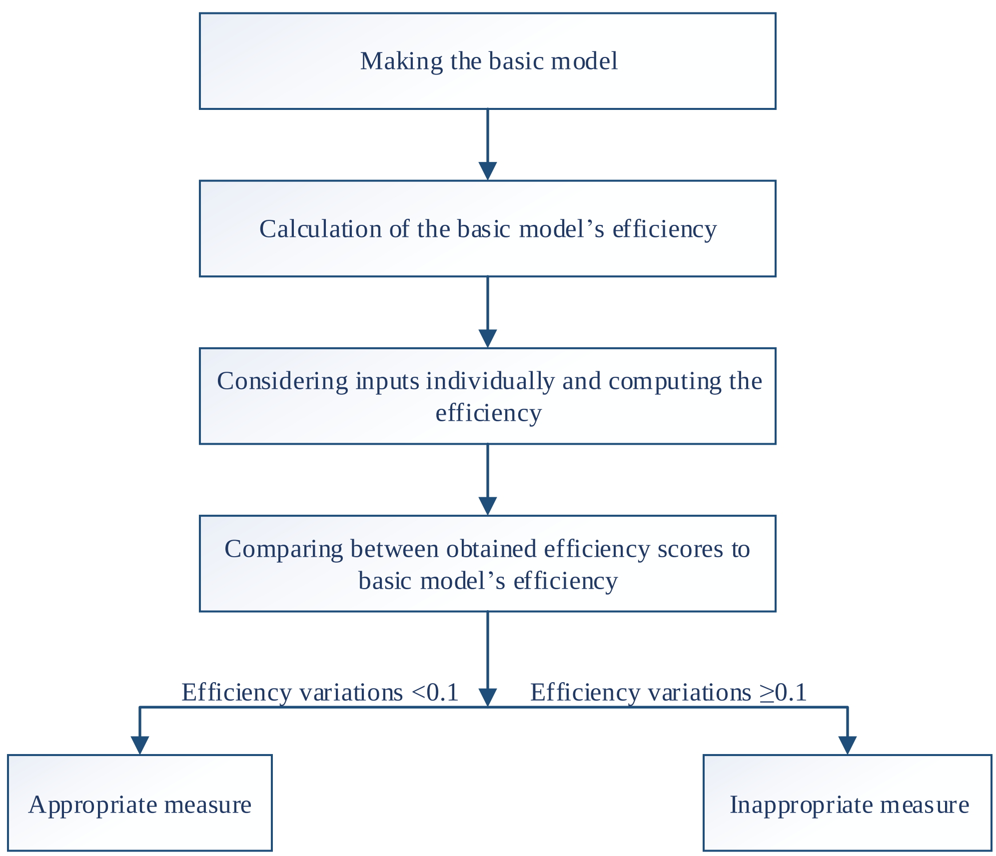

Figure 4. The proposed approach

Table 1. Measures of BSC model

perspective	Indicator	Description
L:Learning and growth	(L1) Staff's knowledge level	The number of educated employees divided by total number of employees
	(p1) Efficiency of train	The number of trips divided by the total passengers' capacity
P: Internal process	(P2) Number of delayed trip	The number of delayed trip divided by total trips
	(C1) Average density per each passenger	The number of passengers per area of train. Increasing this factor lead to this increasing customer satisfaction
C: Customer	(C2) Waiting at the station	Interval time between arrival of two consecutive trains at the station
	(C3) The delay per trip	Total number of delays to total trips
	(F1) Labor costs	Labor costs per train's operation
F: Financial	(F2) Labor costs per each trip	Labor costs divided by total number of trips

Note that these measures have been selected to elucidate the details of implementation of our proposed approach and you can use this approach for any set of indicators and DMUs. With respect to the cause and effect relationships in BSC, mentioned variables are applied as inputs and outputs of three-stage DEA model. In stage 1, staff's knowledge level (L1) is input selected from learning and growth perspective. Also, Efficiency of train (P1) and number of delayed trip (P2) are utilized as the outputs of stage 1. These factors are chosen from internal process perspective. Note that the outputs of stage 1 are assumed as inputs of stage 2 and The output parameters for stage 2 are average density per passengers (C1), waiting at station (C2), and delay per trip (C3). These variables belong to customer perspective. Again, these parameters are applied as inputs of stage 3 and outputs of this stage are labor costs (F1) and labor costs per each trip (F2) adopted financial perspective. Table 2 summarized the numeral values of DAE input/output indicators. After determination input/output variables of DEA model, we should make the basic model illustrated in Figure 5.

Table 2. DEA input/output variables for different stages

DMU	**Input of stage1**	**Inputs of stage 2 (outputs of stage 1)**		**inputs of Stage 3 (outputs of stage 2)**			**outputs of stage 3**	
	L1	**P1**	**P2**	**C1**	**C2**	**C3**	**F1**	**F2**
1	0.1268	46.1826	0.0140	46.1826	0.0866	0.1499	0.9565	124.5035
2	0.2500	7.6581	0.0412	37.0062	0.1144	0.3816	0.9480	795.8625
3	0.2857	0.6405	0.0412	3.0952	0.1144	0.3816	0.8353	8384.125
4	0.2857	7.6536	0.0412	36.9844	0.1144	0.3816	0.8353	701.6526
5	0.2857	4.5347	0.0412	21.9129	0.1144	0.3816	0.8353	1184.243
6	0.3000	8.2519	0.0412	39.8755	0.1144	0.3816	0.7988	622.2993
7	0.2857	2.5049	0.0412	12.1045	0.1144	0.3816	0.8353	2143.851
8	0.1935	41.268	0.0412	199.4195	0.1144	0.3816	1.1826	184.2302
9	0.2632	3.5334	0.0412	17.0744	0.1144	0.3816	0.7317	1331.372
10	0.1667	38.4036	0.0412	185.5777	0.1144	0.3816	2.0508	343.3139

Figure 5. Basic three-stage DEA model

Table 3. Efficiency scores of DEA model

	Stage1	**Stage2**	**Stage3**
DMU1	1.000	0.821	0.943
DMU2	0.667	0.960	0.533
DMU3	0.583	1.000	1.000
DMU4	0.583	0.960	0.470
DMU5	0.583	0.960	0.515
DMU6	0.556	0.960	0.442
DMU7	0.583	0.960	0.559
DMU8	0.885	0.960	0.405
DMU9	0.633	0.960	0.466
DMU10	1.000	0.960	0.712

5 Table 3 shows the efficiency of this model calculated by equations (2), (3), and (4). At the following, we survey the influence of input factors. That is, we assume the outputs of each stage are fixed and input factors are changed and the efficiency of that stage is calculated. This efficiency score is compared to the efficiency of the same stage in the basic model (Table 3). Efficiency variations equal to 0.1 or more are supposed significant and variations less than 0.1 are negligible. The efficiency of DMUs has been obtained by GAMS software.

4.1Efficiency Variation of Stage 2 by Individual Inputs

In this example, we examine the efficiency changes only in stage 2 and 3 because in stage 1, there is only one input. To investigate efficiency variations in stage 2, efficiency of train (P1) is employed as individual input of this stage and the efficiency is computed. The model of this supposition has been shown in Figure 6a. Then, we compute the efficiency in condition that number of delayed trip (P2) is individual input of stage 2 (Figure 6b). Table 4a and 4b draw a comparison between obtained efficiency scores and efficiency of stage 2 in the basic model.

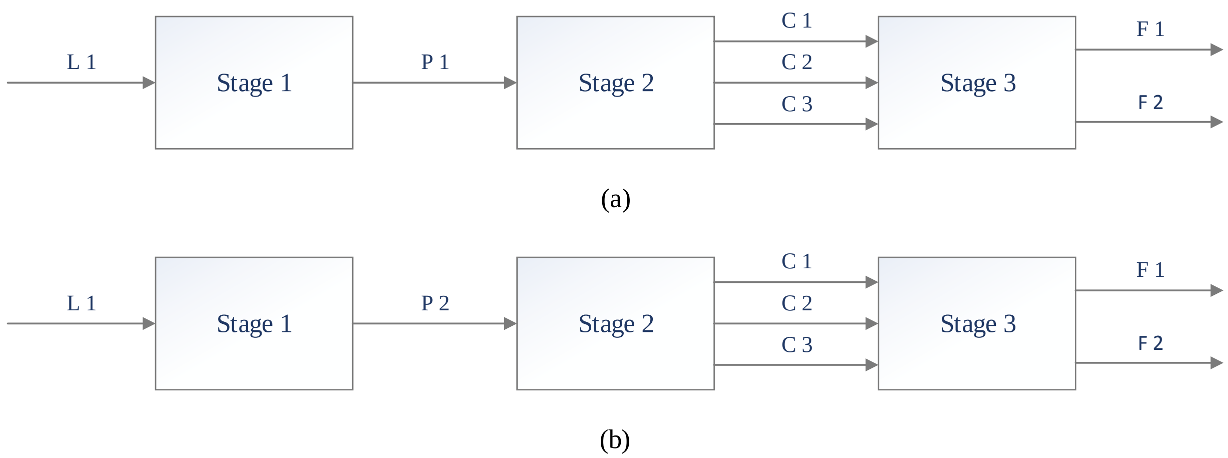

Figure 6. (a) Three-stage DEA model based on P1 input. (b) Three-stage DEA model based on P2 input

Table 4. (a) Efficiency of stage 2, input: (P1) or (P1, P2)

	Stage2 (P1,P2)	Stage2 (P1)
DMU1	0.821	0.199
DMU2	0.960	0.960
DMU3	1.000	1.000
DMU4	0.960	0.960
DMU5	0.960	0.960
DMU6	0.960	0.960
DMU7	0.960	0.960
DMU8	0.960	0.960
DMU9	0.960	0.960
DMU10	0.960	0.960

Table 4. (b) Efficiency of stage 2, input: (P2) or (P1, P2)

	Stage2 (P1,P2)	Stage 2(P2)
DMU1	0.821	0.821
DMU2	0.960	0.390
DMU3	1.000	0.390
DMU4	0.960	0.390
DMU5	0.960	0.390
DMU6	0.960	0.390
DMU7	0.960	0.390
DMU8	0.960	0.390
DMU9	0.960	0.390
DMU10	0.960	0.390

As shown in Table 4a, when only efficiency of train (P1) is supposed as input, efficiency changes are not tangible but DMU 1. A good explanation of significant efficiency variations in this DMU is that input of DMU 1 (Table 2) is too larger than input of other DMUs. Table 4b shows changes of efficiency are meaningful when number of delayed trip (P2) is used as individual input of stage 2. Since the input of DMU 1 is less than input of other DMUs –in origin station, trains travel with a precise schedule and in intermediate stations, trains may be delayed because of changes in the number of passengers- the efficiency score in this DMU is large.

Note that we can identify DMUs (subway station) needing to performance improvement. We can even determine that these DMUs should be focused on which indicators in order to improve the performance. For example, the efficiency of DMU1 is equal to 0.199 (Table 4a) when the efficiency of train (P1) is considered as input. This efficiency score shows that DMU1 must improve its performance in terms of this measure. Similar cases in all stages are investigable likewise.

5.1 Efficiency Variations of Stage 3 by Individual Inputs

In order to investigate the efficiency variations of stage 3, we can consider six different states for selection of input variable:

a) Average density per each passenger (C1)
b) Waiting at the station (C2)
c) The delay per trip (C3)
d) Average density per each passenger and waiting at the station (C1, C2)
e) Average density per each passenger and the delay per trip (C1, C3)
f) Waiting at the station and the delay per trip (C2, C3)

These states have been exhibited in Figure 7a to 7f.

Table 5a to 5f show the calculated efficiency scores of DMUs in all above- mentioned states and compare them with efficiency score of stage 3 in the basic model separately.

(a)

(b)

(c)

(d)

(e)

(f)

Figure 7. Three-stage DEA model based on the (a) C1 input (b) C2 input (c) C3 input (d) C1, C2 inputs (e) C1, C3 inputs (f) C2, C3 inputs

Table 5. (a) Efficiency of stage 3, input: (C1) or (C1, C2, C3)

	Stage3 (C1,C2,C3)	**Stage 3 (C1)**
DMU1	0.943	0.077
DMU2	0.533	0.095
DMU3	1.000	1.000
DMU4	0.470	0.084
DMU5	0.515	0.141
DMU6	0.442	0.074
DMU7	0.559	0.256
DMU8	0.405	0.022
DMU9	0.466	0.159
DMU10	0.712	0.041

Table 5. (b) Efficiency of stage 3, input: (C2) or (C1, C2, C3)

	Stage3 (C1,C2,C3)	Stage 3 (C2)
DMU1	0.943	0.359
DMU2	0.533	0.278
DMU3	1.000	0.449
DMU4	0.470	0.245
DMU5	0.515	0.258
DMU6	0.442	0.233
DMU7	0.559	0.284
DMU8	0.405	0.336
DMU9	0.466	0.234
DMU10	0.712	0.583

Table 5. (c) Efficiency of stage 3, input: (C3) or (C1, C2, C3)

	Stage3 (C1,C2,C3)	Stage 3 (C3)
DMU1	0.943	0.805
DMU2	0.533	0.320
DMU3	1.000	0.445
DMU4	0.470	0.282

DMU5	0.515	0.292
DMU6	0.442	0.268
DMU7	0.559	0.312
DMU8	0.405	0.391
DMU9	0.466	0.262
DMU10	0.712	0.678

Table 5. (d) Efficiency of stage 3, input: (C1, C2) or (C1, C2, C3)

	Stage3 (C1,C2,C3)	**Stage 3 (C1, C2)**
DMU1	0.943	0.527
DMU2	0.533	0.485
DMU3	1.000	1.000
DMU4	0.470	0.427
DMU5	0.515	0.498
DMU6	0.442	0.398
DMU7	0.559	0.559
DMU8	0.405	0.350
DMU9	0.466	0.461
DMU10	0.712	0.609

Table 6. (e) Efficiency of stage 3, input: (C1, C3) or (C1, C2, C3)

	Stage3 (C1,C2,C3)	Stage 3 (C1, C3)
DMU1	0.943	0.943
DMU2	0.533	0.533
DMU3	1.000	1.000
DMU4	0.470	0.469
DMU5	0.515	0.512
DMU6	0.442	0.442
DMU7	0.559	0.545
DMU8	0.405	0.405
DMU9	0.466	0.463
DMU10	0.712	0.712

Table 5. (f) Efficiency of stage 3, input: (C2, C3) or (C1, C2, C3)

	Stage3 (C1,C2,C3)	Stage 3 (C2, C3)
DMU1	0.943	0.805
DMU2	0.533	0.323
DMU3	1.000	0.450
DMU4	0.470	0.285

DMU5	0.515	0.295
DMU6	0.442	0.271
DMU7	0.559	0.316
DMU8	0.405	0.395
DMU9	0.466	0.265
DMU10	0.712	0.685

In state (a), when average density per each passenger (C1) is used as input of stage 3, inspection of Table 5a shows that the efficiency variations are meaningful. In DMU 3, since input value is much less than other input, the efficiency is large. Upon survey of Table 5b, it is clear that efficiency changes are significant if waiting at the station (C2) be individual input of stage 3. Also, when the delay per trip (C3) is chosen as input, Table 5c exhibits meaningful variations in efficiency scores. In state (d), in which average density per each passenger (C1) and waiting at the station (C2) are simultaneously considered as inputs, Table 5d shows intangible efficiency variations. Upon comparison efficiency scores in Table 5e, it is clear that if average density per each passenger (C1) and the delay per trip (C3) are considered as inputs, efficiency can approximately remain constant. Finally, in state (f) in which waiting at the station (C2) and the delay per trip (C3) are input variables, Table 5f exhibits significant variations in efficiency scores.

5. Discussion

In this study, we integrated BSC and DEA model in order to select appropriate measures. We utilized measures of BSC's perspectives as inputs and outputs of three-stage DEA model (basic model) and computed the efficiency of all stages as prime efficiency scores (Table 3). The causal relationships between measures of various perspectives have been shown in Figure 8. Then, we considered each input factor individually and computed the efficiency of stages again. Having compare the obtained efficiencies to primary efficiency scores in the basic model, we neglected changes of efficiency less than 0.1. In order to providing an empirical evaluation of our proposed approach, we tested this approach for urban railway of Tehran. Since there is only one input in stage 1, we examined only stage 2 and 3. In stage 2, when efficiency of train (P1) is considered as input, efficiency variations are not tangible. Vice versa, if number of delayed trip (P2) be assumed as input, the efficiency changes are meaningful. These results indicate that efficiency of train (P1) is more appropriate measure than number of delayed trip (P2).

In stage 3, when each indicator is individually considered as input variable, changes of efficiency are not intangible. This means that none of these indicators individually cannot be appropriate measures. Another scenario for selection of appropriate measures in stage 3 is that to consider indicators to each other. The results indicate that when average density per each passenger (C1) and waiting at the station (C2) are inputs of stage 3, efficiency changes are not significant. Also, when average density per each passenger (C1) is considered with the delay per trip (C3) as input, it causes that efficiency changes be intangible. Simultaneous consideration of waiting at the station (C2) and the delay per trip (C3) as input lead to variations in efficiency be significant. In other words, Simultaneous consideration average density per each passenger (C1) and waiting at the station (C2) or concurrent consideration of average density per each passenger (C1) and the delay per trip(C3) can be considered as appropriate measures while concurrent considering waiting at the station (C2) and the delay per trip (C3) cannot culminate in an appropriate measure. This results mean that the urban stations utilized the selected measures efficiently regarding to efficiency scores. Also, the causal linkages among these indicators are more strongly. The relationships between appropriate measures according to our proposed approach have been showed in Figure 9a and 9b.

Also, they can adapt their decisions to urban transit system's strategic objectives for implementation of improvement projects.

6. Conclusion

Subway is one of the important needs of each modern community. In order to identify its weaknesses and performance improvement culminated in citizenry satisfaction, it is necessary to define appropriate measures for performance evaluation. In other words, one of the most important points in performance evaluation of each system such as urban rail transit is selection of appropriate indicators. This paper presented a novel approach for choosing right indicators through integrating of balanced scorecard and three-stage data envelopment analysis. Employing the DEA model for this purpose leads to select the right indicators utilized efficiently by DMUs (subway stations). Also, considering the BSC helps to select measures having the strongest causal linkages. It allows urban transit's policy makers to focus on long-term goals instead of short- term ones.

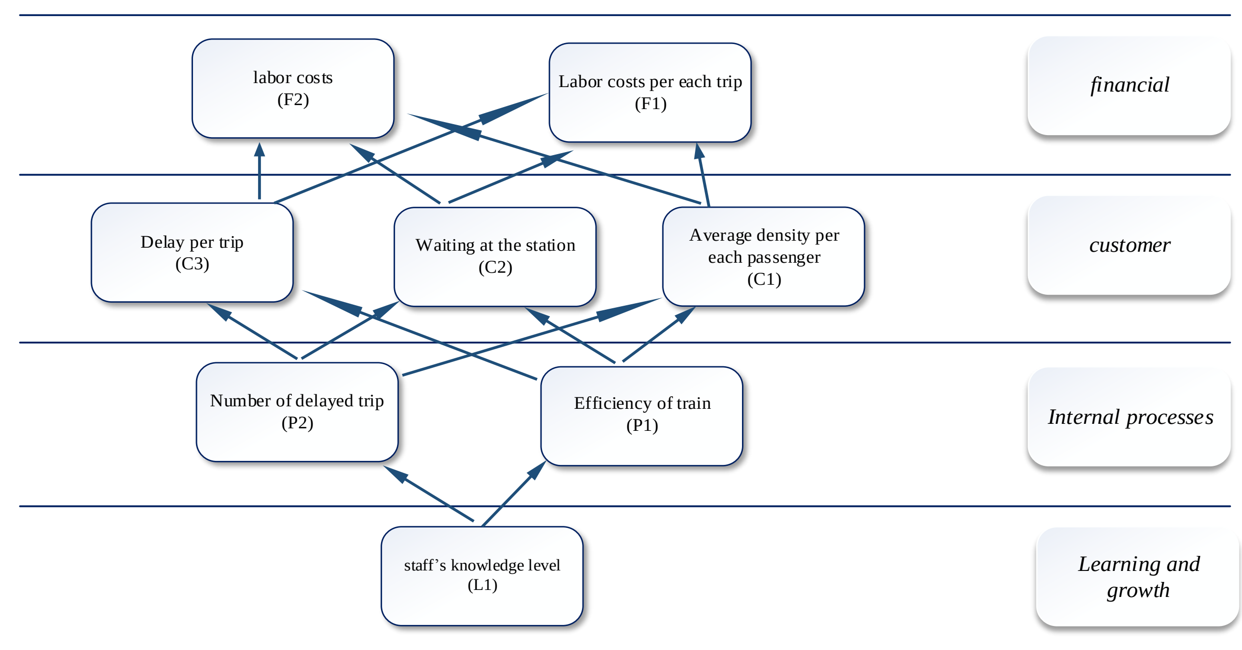

Figure 8. The cause and effect relationships between measures

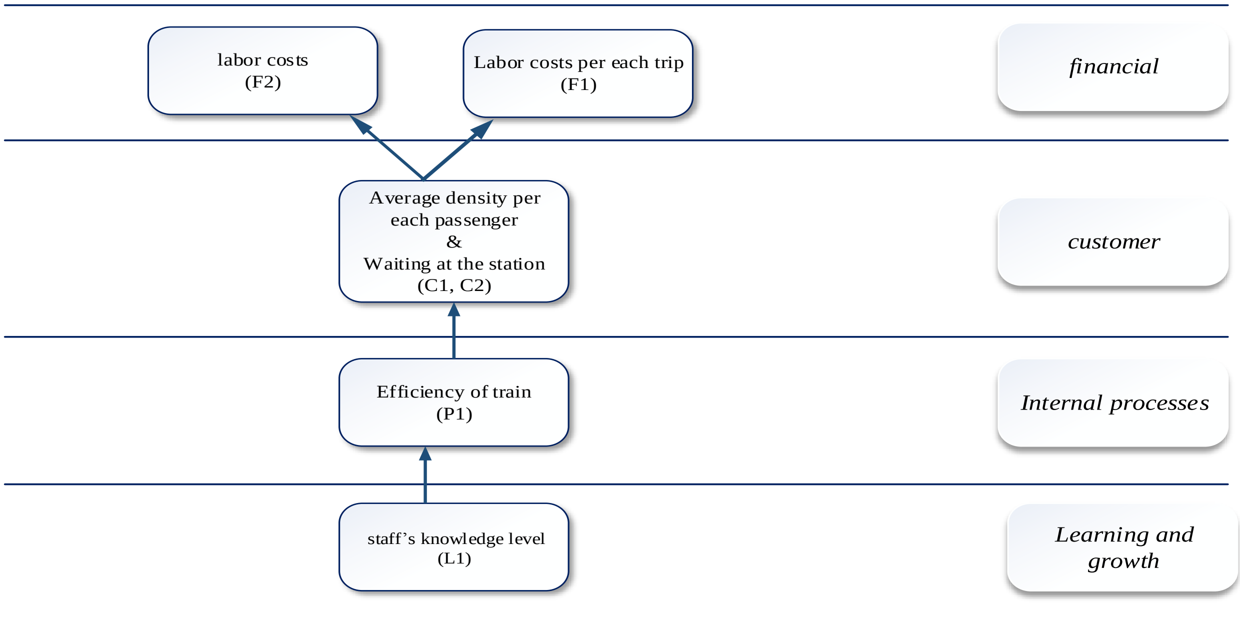

Figure 9. (a) Relationships between selected appropriate measures according to our proposed approach

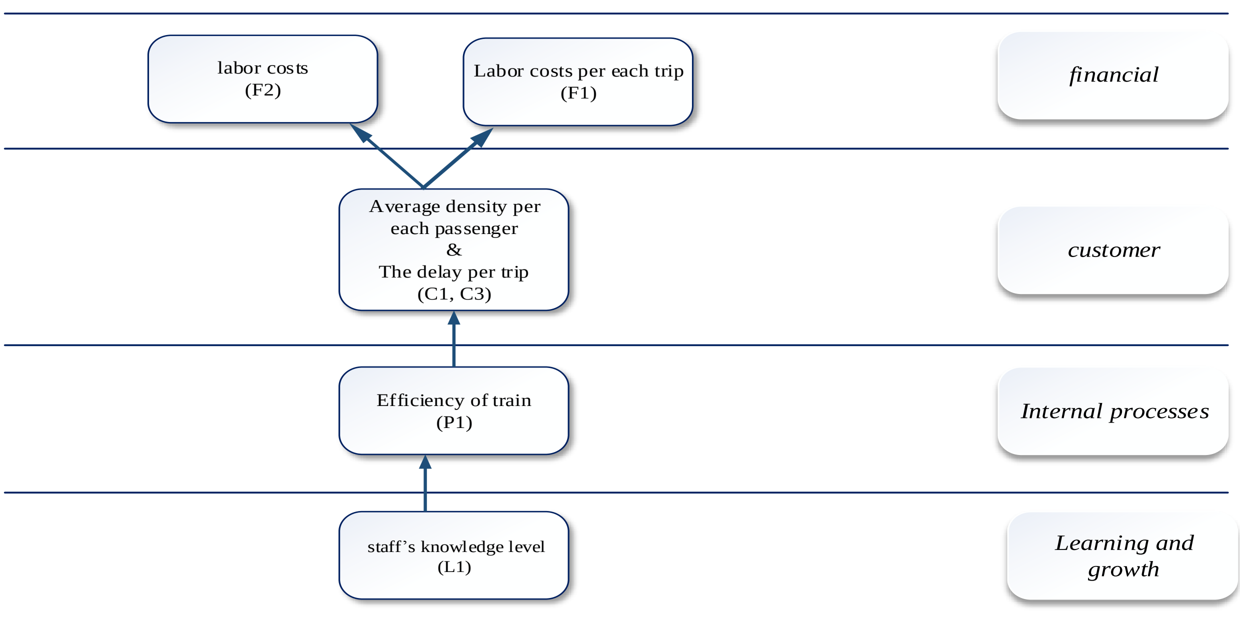

Figure 9. (b) Relationships between selected appropriate measures according to our proposed approach

At first, regarding to the BSC's cause and effect relationships, the indicators of each perspective were applied as input and output variables of three-stage DEA model. The efficiency of this structure supposed as basic model was calculated. At the next step, individual inputs were considered in each stage and the efficiency of that stage was computed again and made comparison to the efficiency score of the same stage in the basic model. The efficiency changes equal to 0.1 or more were deemed significant and variations less than 0.1 were ignored. With the analysis of observed results, we can determine the most appropriate indicators in each BSC's perspective. To learning how to utilize this approach, we implemented it for indicators and data of Tehran's urban railway. Data are related to 10 stations of line 5 which connect Tehran to Karaj. Results indicate that efficiency of trains (P1) is the right indicator of stage 2. Also, there are two appropriate measures in stage 3: 1) Concurrent consideration of average density per each passenger (C1) and waiting at the station (C2). 2) Simultaneous consideration of average density per each passenger (C1) and the delay per trip (C3).

The proposed approach in this study, is a worthwhile guideline for managers and decision makers in transportation industry. Since the appropriate indices can reflect the progress of urban railway system toward their goals, managers can be aware of unexpected problems. Also, policy makers can adapt their decision to strategic objectives of urban public transportation system for implementation of improvement projects. Comprehension how public transit and investments can be employed most to result in better performance is a vital factor for improvement of services and increasing citizenry satisfaction through urban railway system as well. The methodology outlined in this paper can also be applied to other industries and organizations. In various industries, these indicators are different thus, in order to improve operational performance through these indicators, the propriety of measures should be adjusted based on operational program of organization, customer requirement and environmental changes.

7. References

-Ahmed Mousa Ali, M. (2014) ''Lessons for Policy Makers in Non-High Speed Rail Countries: A Review'', International Journal of Transportation Engineering, Vol. 2, No. 4, pp. 323-338.

-Åhrén, T. and Parida, A. (2009) "Maintenance performance indicators (MPIs) for benchmarking the railway infrastructure: a case study'', Benchmarking: An International Journal, Vol. 16, No. 2, pp. 247-258.

-Åhrén, T., and Parida, A. (2009) "Overall railway infrastructure effectiveness (ORIE) A case study on the Swedish rail network'', journal of quality in maintenance engineering, Vol. 15, No. 1, pp. 17-30.

-Alizadeh, H., Shahmoradi, B., Abdi, M. H. and Rahimi, A. (2014) "Developing Transit-oriented Strategies for Sanandaj City Center, Iran", International Journal of Transportation Engineering, Vol. 1, No. 3, pp. 141-150.

-Aydin, N. (2017) "A fuzzy-based multi-dimensional and multi-period service quality evaluation outline for rail transit systems", Transport Policy, Vol. *55*, pp. 87-98.

-Aydin, N., Celik, E. and Gumus, A. T. (2015) "A hierarchical customer satisfaction framework for evaluating rail transit systems of Istanbul", Transportation Research Part A: Policy and Practice, Vol. *77*, pp. 61-81.

-Banker, R. D., Potter, G. and Srinivasan, D. (2000) "An empirical investigation of an incentive plan that includes nonfinancial performance measures", The accounting review, Vol. 75, No. 1, pp. 65-92.

-Bento, A., Bento, R. and White, L. F. (2013) "Validating cause-and-effect relationships in the balanced Fscorecard", Academy of Accounting and Financial Studies Journal, Vol. 17, No. 3, pp. 45.

-Boame, A. K. (2004) "The technical efficiency of Canadian urban transit systems", Transportation Research Part E: Logistics and

Transportation Review, Vol. 40, No. 5, pp. 401-416.

-Chen, F. H., and Tzeng, G. H. (2014) "Probing organization performance using a new hybrid dynamic MCDM method based on the balanced scorecard approach", Journal of Testing and Evaluation, Vol. 43, NO. 4, pp. 924-937.

-Costa, Á., and Markellos, R. N. (1997) "Evaluating public transport efficiency with neural network models", Transportation Research Part C: Emerging Technologies, Vol. 5, No. 5, pp. 301-312.

-Famurewa, S. M., Stenström, C., Asplund, M., Galar, D. and Kumar, U. (2014) "Composite indicator for railway infrastructure management", Journal of Modern Transportation, Vol. 22, No. 4, pp. 214-224.

-Färe, R. and Grosskopf, S. (2000) "Network DEA", Socio-Economic Planning Sciences, Vol. 34, No. 1, pp. 35-49.

-García Valderrama, T., Revuelta Bordoy, D. and Rodríguez Cornejo, V. (2013) "Balanced Scorecard and Efficiency: Desing and Empirical Validation of strategic Map in the University by Means of DEA".

-Ghotbuee, A., Hemati, M. and Fateminezhad, R. (2012) "An empirical study based on BSC-DEA to measure the relative efficiencies of different health care centers in province of Semnan, Iran", Management Science Letters, Vol. 2, No. 7, pp. 2643-2650.

-Hong, L., Yan, Y., Ouyang, M., Tian, H. and He, X. (2017) "Vulnerability effects of passengers' intermodal transfer distance preference and subway expansion on complementary urban public transportation systems" Reliability Engineering & System Safety, Vol. 158, pp. 58-72.

-Ittner, C. D. and Larcker, D. F. (1998) "Are nonfinancial measures leading indicators of financial performance? An analysis of customer satisfaction", Journal of accounting research, Vol. 36, pp. 1-35.

-Kao, C. (2009) "Efficiency measurement for parallel production systems", European Journal of Operational Research, Vol. 196, No. 3, pp. 1107-1112.

-Kaplan, R. S. and Norton, D. P. (1992) "The Balanced Scorecard--Measures That Drive Performance", Harvard Business Review, Vol. 70, No. 1, pp. 71-79.

-Kaplan, R. S. and Norton, D. P. (1996) "The balanced scorecard: translating strategy into action. Harvard Business Press".

-Kaplan, R. S. and Norton, D. P. (2004) "The strategy map: guide to aligning intangible assets", Strategy & Leadership, Vol. 32, No. 5, pp. 10-17.

-Lan, L. W. and Lin, E. T. (2005) "Measuring railway performance with adjustment of environmental effects, data noise and slacks", Transportmetrica, Vol. 1, No. 2, pp. 161-189.

-Lan, L. W. and Lin, E. T. (2006) "Performance measurement for railway transport: stochastic distance functions with inefficiency and ineffectiveness effects", Journal of Transport Economics and Policy (JTEP), Vol. 40, No. 3, pp. 383-408.

-Lawrence, W. and Erwin, T. (2003) "Technical efficiency and service effectiveness for railways industry: DEA approaches", Journal of the Eastern Asia Society for Transportation Studies, 5.

-Leung, L. C. Lam, K. C. and Cao, D. (2006) "Implementing the balanced scorecard using the analytic hierarchy process & the analytic network process", Journal of the Operational Research Society, Vol. 57, No. 6, pp. 682-691.

-Lewis, H. F. and Sexton, T. R. (2004) "Network DEA: efficiency analysis of organizations with complex internal structure", Computers & Operations Research, Vol. 31, No. 9, pp. 1365-1410.

-Liang, C. J. and Hou, L. C. (2007) "A dynamic connection of balanced scorecard applied for the hotel", Journal of Services Research, Vol. 7, NO. 1, pp. 91.

-Lucianetti, L. (2010)" The impact of the strategy maps on balanced scorecard performance", International Journal of Business Performance Management, Vol. 12, No. 1, pp. 21-36.

-Maina, E.V., Forda, A. and Rabinson, R.M. (2016) "Impact of Optimally Minimizing Delay Times on Safety at Signalized Intersections in Urban Areas, Case Study: The City of Virginia Beach", International Journal of Transportation Engineering, Vol.3, No. 4, pp. 277-288.

-Malhotra, R., Malhotra, D. K. and Lermack, H. (2009) "Using data envelopment analysis to analyze the performance of North American class I freight railroads" In Financial Modeling Applications and Data Envelopment Applications, Emerald Group Publishing Limited, pp. 113-131.

-Mallikarjun, S., Lewis, H. F. and Sexton, T. R. (2014) "Operational performance of US public rail transit and implications for public policy", Socio-Economic Planning Sciences, Vol. 48(1), pp. 74-88.

-Manasakis, C., Apostolakis, A., and Datseris, G. (2013) "Using data envelopment analysis to measure hotel efficiency in Crete", International Journal of Contemporary Hospitality Management, Vol.25, No. 4, pp. 510-535.

-Murali, S., and Pugazhendhi, S. (2016) "An integrated model to identify and rank the after sales service strategies of firms engaged in household appliances business", International Journal of Services and Operations Management, Vol.*24,* NO. 1, PP. 99-124.

-Nathanail, E. (2008) "Measuring the quality of service for passengers on the Hellenic railways", Transportation Research Part A: Policy and Practice, Vol. 42, No. 1, pp. 48-66.

-Nesterova, N. S., Goncharuk, S. M., Anisimov, V. A., and Anisimov, A. V. (2016) "Strategy development management of Multimodal Transport Network", In *MATEC Web of Conferences*, Vol. 86, p. 05024, EDP Sciences.

-Pan, J. N., and Nguyen, H. T. N. (2015) "Achieving customer satisfaction through product–service systems", European Journal of Operational Research, Vol. 247, No. 1, pp. 179-190.

-Powell, J. P., González-Gil, A., and Palacin, R. (2014) "Experimental assessment of the energy consumption of urban rail vehicles during stabling hours: influence of ambient temperature", Applied Thermal Engineering, Vol. 66, No. 1, pp. 541-547.

-Qin, F., Zhang, X., and Zhou, Q. (2014) "Evaluating the impact of organizational patterns on the efficiency of urban rail transit systems in China", Journal of Transport Geography, Vol. 40, pp. 89-99.

-Rassafi, A.A., Ostad Jafari, M. and Javanshir, H. (2014) "An Appraisal of Sustainable Urban Transportation: Application of a System Dynamics Model", International Journal of Transportation Engineering, Vol. 2, No. 1, pp. 47-66.

-Sangtarash, L. (2012) "Performance evaluation (performance) subway stations, using a combination of balanced scorecard and data envelopment analysis", Master's thesis, Iran University of Science and Technology.

-Sexton, T. R. and Lewis, H. F. (2003) “Two-stage DEA: An application to major league baseball”, Journal of Productivity Analysis, Vol. 19, No. (2-3), pp. 227-249.

-Shafiee, M., Lotfi, F. H. and Saleh, H. (2014) “Supply chain performance evaluation with data envelopment analysis and balanced scorecard approach”, Applied Mathematical Modelling, Vol. 38, No. 21, pp. 5092-5112.

-Shapouri, F. and Keramati, A. (2015) “A framework for constructing customer relationship management strategy map based on multiple criteria decision-making approach”, International Journal of Electronic Customer Relationship Management, Vol. *9*, NO. 2-3, PP. 175-188.

-Shen, W., Xiao, W. and Wang, X. (2016) “Passenger satisfaction evaluation model for Urban rail transit: A structural equation modeling based on partial least squares”, Transport Policy, Vol. 46, pp. 20-31.

-Stenström, C., Parida, A. and Galar, D. (2014) “Performance indicators of railway infrastructure”, The international Journal of railway technology, Vol. 1, No. 3, pp. 1-18.

-Stenström, C., Parida, A. Galar, D., and Kumar, U. (2013) “Link and effect model for performance improvement of railway infrastructure”, Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, Vol. 227, No. 4, pp. 392-402.

-Titko, J., Stankevičienė, J. and Lāce, N. (2014) “Measuring bank efficiency: DEA application”, Technological and Economic Development of Economy, Vol. 20, No. 4, pp. 739-757.

-Tsai, Y. C. and Cheng, Y. T. (2012) “Analyzing key performance indicators (KPIs) for E-commerce and Internet marketing of elderly products: A review”, Archives of gerontology and geriatrics, Vol. 55, No. 1, pp. 126-132.

-Varmazyar, M., Dehghanbaghi, M. and Afkhami, M. (2016) “A novel hybrid MCDM model for performance evaluation of research and technology organizations based on BSC approach”, Evaluation and Program Planning, Vol. 58, pp. 125-140.

-Wanke, P., Barros, C. P. and Figueiredo, O. (2016) “Efficiency and productive slacks in urban transportation modes: A two-stage SDEA-Beta Regression approach”, Utilities Policy, Vol. 41, pp. 31-39.

-Wu, H. Y., Tzeng, G. H. and Chen, Y. H. (2009) “A fuzzy MCDM approach for evaluating banking performance based on Balanced Scorecard”, Expert Systems with Applications, Vol. 36, No. 6, pp. 10135-10147.

-Yaghoobi, T., Yaghoobi, T. Haddadi, F., and Haddadi, F. (2016) “Organizational performance measurement by a framework integrating BSC and AHP”, International Journal of Productivity and Performance Management, Vol. 65, NO. 7, pp. 959-976.

-Yao, J. and Liu, J. (2016) “E-Government Project Evaluation: A Balanced Scorecard Analysis”, Journal of Electronic Commerce in Organizations (JECO), Vol. 14, NO. 1, pp. 11-23.

-Yu, M. M. (2008) “Assessing the technical efficiency, service effectiveness, and technical effectiveness of the world’s railways through NDEA analysis” Transportation Research Part A: Policy and Practice, Vol. 42, No. 10, pp. 1283-1294.