Congcong Yang and Alfred Taudes and Guozhi Dong

Efficiency Analysis of European Freight Villages-Three Peers for Benchmarking

Paper

Efficiency Analysis of European Freight Villages-Three Peers for Benchmarking

Congcong Yang, Alfred Taudes, Guozhi Dong

Abstract Measuring the performance of Freight Villages (FVs) has important implications for logistics companies and other related companies as well as governments. In this paper we apply Data Envelopment Analysis (DEA) to measure the performance of European FVs in a purely data-driven way incorporating the nature of FVs as complex operations that use multiple inputs and produce several outputs. We employ several DEA models and perform a complete sensitivity analysis of the appropriateness of the chosen input and output variables, and an assessment of the robustness of the efficiency score. It turns out that about half of the 20 FVs analyzed are inefficient, with utilization of the intermodal area and warehouse capacity and level of goods handed the being the most important areas of improvement. While we find no significant differences in efficiency between FVs of different sizes and in different countries, it turns out that the FVs Eurocentre Toulouse, Interporto Quadrante Europa and GVZ Nürnberg constitute more than 90% of the benchmark share.

Keywords Freight Village, benchmarking, performance measurement, Data Envelopment Analysis (DEA)

Congcong Yang
Institute for production management, Vienna University of Economics and Business
Welthandelsplatz 1, 1020 Vienna, Austria

Alfred Taudes(Corresponding author)

Institute for production management, Vienna University of Economics and Business
Welthandelsplatz 1, Building D2, 1020 Vienna, Austria
Tel: (+43 1) 31336 5611 Fax: (+ 43 1) 31336 5610

Email: alfred.taudes@wu.ac.at

Guozhi Dong
Computational Science Center, University of Vienna
Oskar-Morgenstern-Platz 1,1090 Vienna, Austria

1 Introduction

The term “Freight Village” (FV) refers to a defined area organized for carrying out all activities related to transport, logistics and distribution for both national and international transit (Ballis 2006). Initially, it was established in response to the challenges posed by regional population and freight growth; however, with ongoing increase in globalized trade, FVs are widely used in the process of trade and transportation in the world (Wu and Haasis 2013). Spurred by changes in freight and logistics processes, FV has emerged around the world not only as a logistical interconnection point within a logistics network, but also a “business generator”, which contributes to supply chain efficiency improvement, regional economic growth and environmental protection (Meidute 2005). In the face of growing globalization of business activities and escalating demand in smoothing the flow of supply chain, FV management becomes a daunting task, which has become an important topic in supply chain management and industrial cluster.

To achieve profitability and survive in the market, all enterprises are required to perform activities in an efficient way (Andrejić et al. 2013). The FV is no exception. From a wider perspective, the FV serves as the backbone of the logistics system, affecting the performance of the entire transportation network and supply chain (Cezar-Gabriel 2010). Particularly, the intelligent multimodal transport chains that are implemented by FVs contribute to an efficient logistics network (Winkler and Seebacher 2011). Looking at it from another angle, measuring and improving FVs performance has significant implications for a certain number of stakeholders. For example, it assists 3PL companies and other related companies (warehouse operators, transportation operators) in identifying and selecting the most efficient FV at which to base their operations. Also, it aids governments making effective decisions in the FV development programs. For a FV per se, benchmarking its own efficiency against that of comparable ones is a feasible method for managers to ensure competitiveness.

Due to sizable investment, operating and maintenance costs associated with FV infrastructure and their regional economic ramifications, there were a number of research efforts to evaluate FV performance, especially in China. Notable examples of such efforts include: Wang(2009), Luo(2013) and Manfred et al.(2008). Unfortunately, there are too many papers that replicate previous research, while offering scant methodological and theoretical improvements. For instance, the majority of Chinese papers tend to construct logistics park performance frameworks, along with similar methods and procedures such as AHP, Fuzzy Evaluation Method. Particularly, the lack of explication on the variables in term of selection process and implication brings to question their usefulness as a framework to guide further study. Consistent with this, Liu et al. (2010) underscored the need to enhance the implementability of performance indicators. In addition, a distinctive characteristic of researches on FV appraisal is that it lacks standard methodologies or decision criteria (Kapros et al. 2005). It is observed that recent studies have attempted to evaluate the relative efficiency of FVs using Data Envelopment Analysis (DEA) models. However, this stream of research is still in an early stage. For example, de Carvalho and Lima Jr (2010) measured and compared the efficiency of six logistics platforms in Europe with DEA to guide the development of new logistics platform. Haralambides and Gujar (2012) proposed a new eco-DEA model and applied it to sixteen dry ports in India. One particular study was conducted by Liu et al. (2013) who treated the employee as a dual variable and utilized the dual variable DEA model to measure the efficiency performance of logistics parks in Neimenggu Province, China. It is evident that researchers introduced DEA as a possible technique for efficiency measurement and performance comparison in FV; however, they did not exhibit the application procedures systematically when taking the number of indicators and DMUs into account. Accordingly, extending previous studies by showing how DEA can be applied as a benchmarking tool for FV operations is of great importance to enrich evaluation research on FVs.

Apart from research gaps, another motivation derives from the integration and comparison of performance measurement studies of FVs from practice and academia. To assess the development level of European FVs, EUROPLATFORM EEIG and DGG carried out a large-scale benchmark study in 2010, in which 78 FVs were assessed and ranked by SWOT analysis. For the whole research, readers can refer to Koch et al. (2010). Given different benchmark methods, it is of interest to make a comparison between this study and our approach. As a consequence, this paper aims to examine the performance of sampled FVs in Europe at the macro-level, bringing forth scopes of improvement through DEA application and shedding light on efficiency measurement.

The remainder of the paper is organized as follows: Section 2 describes the DEA models used in the present study. Section 3 reports the description of the data and the specification of input and output variables. Section 4 illustrates the empirical analysis results, including relative efficiency scores derived from CCR and BCC models, a complete sensitivity analysis of the appropriateness of the chosen input and output variables and the robustness of efficiency scores, a benchmark share measurement and two hypothesis testing. The final section outlines the most relevant conclusions, along with a scope for future research.

2 Research Methodology

Data Envelopment Analysis (DEA) a mathematical programming approach for evaluating the relative efficiency of decision-making units (DMUs) (Malekmohammadi et al. 2011). We argue DEA is particularly useful in the efficiency measurement of FVs based on following reasons. Firstly, since the production process of FVs is quite complicated and knowledge of the production function is unknown, DEA allows one to gauge FVs' efficiency and performance without opening the "black box" (the operational process and mechanism of FVs). Secondly, in contrast with evaluation methods that based on PI indicators, the DEA technique captures the performance of FVs comprehensively by taking multiple inputs and outputs into account. In particular, DEA tends to identify "best practice" from a large number of FVs, rather than concerning only one FV, which thus solves the problem of generalization and applicability when several FVs are involved simultaneously. Thirdly, DEA is less data demanding for it works fine with small sample size (Sufian 2005), which can be regarded as another notable strength of DEA in the measure of FVS as gathering data from FVs is a daunting task.

2.1 Data envelopment analysis

Data Envelopment Analysis seeks to identify top performing units in a particular sector and develop possible ways to improve DMU's performances for those units that are far away from "best-practice frontier" (Liang et al. 2008). Although there is a wealth of literature on both basic and applied research in DEA, the most widely used models for DEA are the CCR and the BCC (Ho and Zhu 2004). The CCR model was initially proposed by Charnes et al. (1978) under the assumption of constant returns to scale; while the BCC model, revised on the foundation of CCR model by Banker et al. (1984) allowing variable returns to scale. For the sake of brevity, specific formula of CCR and BCC model are not given here. Readers can refer to Cooper et al. (2007) for the discussion of the standard DEA model and the Mathematical Appendix.

In this study, a FV is viewed as a DMU and its operating efficiency will be broken down into aggregate (mix), technical, and scale efficiencies, which can be measured by CRS model, VRS model and the ratio of CRS (CCR) and VRS score (BCC), respectively. Technical efficiency reflects the ability of a FV to obtain the maximum outputs given a set of inputs, while scale efficiency reflects the ability of a FV to increase its productivity by achieving its optimal size. It should be noted if scale-inefficient exists, it is of interest to determine whether IRS or DRS is the primary cause of scale it. A detailed discussion of this problem is given in the paper of Zhu and Shen (1995). In addition, based on these efficiency measures, the root cause of inefficiency and the projection to be efficient can be investigated, based on top performing FVs that can be references for inefficient ones.

The DEA method can be measured in input-side or output-side. The former pursue minimal possible reduction of usage in inputs when remain the output levels, while the latter seek maximal feasible expansion in outputs without changing the input quantities(Yu and Chen 2011). Within the context of the FVs, both orientations are useful. Managers who are concerned with “how to fully and efficiently use resources” might prefer input-oriented models. On the contrary, output-side models (vs. input-side) are more associated with planning and strategy formulation (Cullinane et al. 2006). However, the choice should be made according to prevailing circumstances (Golany and Roll 1989). In this study, output-oriented models were chosen, because (i) outputs in our model are more controllable than inputs. FVs are normally associated with long-lived infrastructures and facilities and with a long-term planning horizon, thus adjusting a facility in the short-term is impossible once it has been built (e.g., the size); (ii) an output-oriented model can provide information for managers on the capacity utilization of a FV, indicating whether output has been maximized given the input, which, in turn provide reference for further expansion planning.

2.2 Sensitivity analysis

DEA, a data-based analysis method, is sensitive to data and measurement error (Singh and Bajpai 2013). Stated differently, different parameters (inputs or outputs) or fewer parameters for evaluation might results in different outcomes. To evaluate the robustness of efficiency scores, a sensitivity analysis is conducted from two perspectives: the removal of variables and a jack-knifing analysis.

2.2.1 Removal of variables

As stated by Ramanathan (2003), it is possible for a DMU turn to be efficient if it achieves extraordinarily better results in terms of one input, but performs below average in other inputs. Correspondingly, to test how efficiency scores vary with changes in inputs and outputs, one variable is removed at a time from the variables set. Then, the impact of different criteria on the efficiency score is evaluated by comparing DEA efficiencies with the structurally perturbed models. To maintain the same degree of freedom, the removed variable is returned before the next round of analysis (Singh and Bajpai 2013).

2.2.2 Jack-knifing analysis

Jack-knifing is an iterative technique that produces a distribution of estimates by systematically dropping one observation at a time (Ondrich and Ruggiero 2002). Following Charles et al. (2012), the observations to be discarded are efficient units that construct the frontier, not each DMU. This analysis specifically operates by observing the change of efficiency scores after dropping the efficiency unit. If significant shifting is experienced when removing one efficient unit, then possible outliers may exist and further analysis should be followed. Otherwise, one can argue that no outlier can be identified and the efficiency result is not sensitive to the efficient unit.

2.3 Benchmark share measure

The benchmark share measure, a ranking measure by combining the factor-specific measure and variable RTS, aims at distinguishing the most important variables (inputs/outputs) and identifying those efficient DMUs which can be treated as benchmarks (Zhu 2000). Specifically, this method consists of two steps: (i) applying specific modes (input/output specific model) for each inefficient DMU to determine the maximal possible decrease in a certain input (or increase in a certain output) without adjusting the remaining inputs and outputs; (ii) calculating each efficient DMU's benchmark share. The bigger the benchmark share measure, the more important an efficient DMU is in the benchmarking. The zero benchmark-share indicated that an efficient DMU does not act as a reference set for any inefficient units. With limited space, the benchmark-share model is shown in Appendix and we refer readers to Zhu (2000) for the details of the estimation algorithm.

3 Data and variable construction

Regarding the application of DEA method in FVs, data availability is particularly important and might be a bottleneck. Due to strictly confidential, only a few variables regarding FVs are in the public domain, indicating that "getting" data directly from publications such as annual report or statistical report is impossible. Besides, it is difficult to identify appropriate responders for the survey, as many operators in FVs possess the first-hand data, rather than the FV management company. For convenience and information transparency, we attempted to survey FVs in Europe. Techniques like (i) list potential variables as complete as possible (ii) serving data availability as the ultimate criterion for variable selection and construction are also adopted aid in the data and variable construction.

Most notably, the unit of analysis is the FV itself, rather than a specific internal facility, such as a warehouse or intermodal terminal, due to the fact that FVs is a broad concept with varying size and functions. Additionally, since all of our sampled FVs have participated in the study "Ranking of the European Freight Village locations-benchmarking of the European experiences" and have already complied with the homogeneous criteria, we can assume that they are comparable. For more details on the selection of comparable FVs, please refer to Koch et al.(2010).

3.1 Identification of input and output variables

For DEA assessment, choosing the input and output variables is the most important stage as DEA results are highly influenced by this choice (De Witte and Marques 2010). However, DEA itself does not provide guidance for the specification of the input and output variables (Nataraja and Johnson 2011). Basically, literature survey and data availability assist in identifying suitable indicators (Bhanot and Singh 2014). Theoretically, the identification of variables should base upon the operational process of FV to ensure precise and complete analysis results. Until now, few literature analysed the operational process of FVs in a systematically way. Cassone and Gattuso(2010) analysed the FVs from the functions perspective, where he classified primary elements of FVs and visualized their relations among the areas(Fig.1). Obviously, the classification and analysis was conducted on an aggregated and Marco level because the FV, however, is a highly complex system with a large number of entities, a wide variety of services and complicated relationships among processes. Alternatively, we start with analysing indicators generally used in production approaches and then take the main functions of a typical FV into account from a broader perspective.

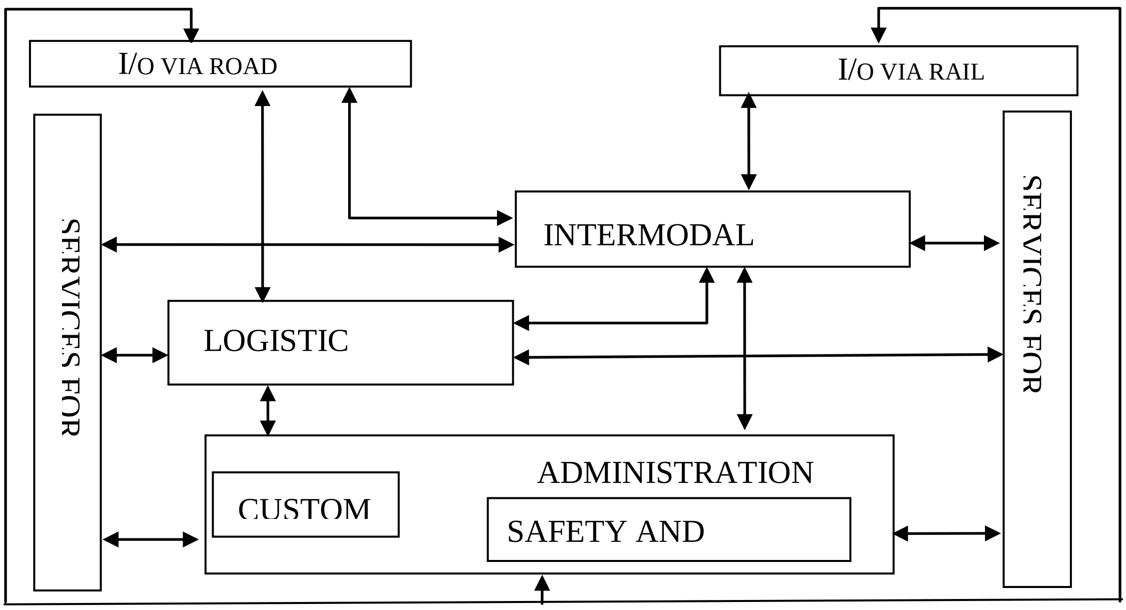

Fig.1. Relations between the areas of a Freight Village (Cassone and Gattuso, 2010).

Essentially, the inputs are various resources consumed by DMUs for operation, while the outputs represent a set of quantitative measures of results expected from operation (George and Rangaraj 2008). In general, resource input can include any combination of labour, equipment, capital and/or information; outputs can be categorized as aggregate revenue, profits, quality, utilization and customer satisfaction (Ross and Droge 2004). Accordingly, we list the following variables by considering all these variables suggested in previous studies(e.g., Chakraborty et al. 2011; Haralambides and Gujar 2012).

- *Area of the FV.* In many cases, this item is used to measure the input of land. The total area of a FV in hectares, however, it is somewhat subjective as some FVs report a gross area including an expanding area that is not yet defined. Since the undeveloped area does not have a strong influence on the current output levels, the already-developed area in hectares is more appropriate.
- *Total amount of investment.* Like total area, this variable is generally used in the production approach. As an aggregated concept, this indicator includes investment in land, equipment and infrastructure. Due to the huge investment and diversity of shareholders, it seems difficult to collect crisp data on this item. For the sake of simplicity, the measurement unit of investment is one million Euros.
- *Intermodal terminal and warehouse.* Warehouses and the intermodal terminal are the most important infrastructures inside a logistics center (Europlatforms 2004). The intermodal terminal is the heart of the FV and multimodal trans-shipment enables the consolidation of transport, creating strong logistics processes by boosting efficiency (Ballis and Golias 2002). Accordingly, it is essential to take these two facilities as substantial resources to support the output, which can be measured in both area and capacity.
- *Number of employees.* As a rule, the number of employees is a proxy variable of labor input; however, as argued by Liu et al. (2013) that it is reasonable to treat it as input and/or output in the FV context. We investigated employees from both management companies and operator companies on site, which stand for the input indicator and the output indicator, respectively.
- *Number of companies attracted on site.* This indicator is a typical output in existing studies, reflecting the development and utilization of FVs. The higher the efficiency of a FV, the more possible it is to attract related companies settled on site; correspondingly, the FV can be operated in a healthy manner. Without the support of companies on site, the FV does not perform.
- *Annual load handling.* Similarly, this indicator was treated as an output. It is a positive variable: the more goods are handled, the better the relative performance of this FV. Since the majority of FVs can provide intermodal transportation, this variable is the sum amount of load handling in FVs, including road, rail and water.
- *Annual turnover.* This is a monetary indicator for measuring the operational profit and the sustainable development of a FV. Similarly to “total investment”, this factor is very sensitive to the financial situation.

3.2 Sampling and data collection

A survey was carried out to find FVs’ efficiencies. Typically, the questionnaire is developed in light of Koch et al.(2010) and based on following assumptions: (i)the data acquired by the previous benchmark study is available in reality; (ii) following a similar pattern to the benchmarking survey that had already been carried out in FVs might contribute to the development of variables and questions as well as the enhancement of response rate.

The survey was created on Survey Monkey and sent to 150 FVs. The first survey was followed by two reminder emails and follow-up calls. As the survey involves nine different countries in Europe, the survey language could be expected to affect the response rate. Thus, the questionnaire and invitation letter were translated from English into the respective languages (such as Italian and German) and follow-up calls were made by a native-speaker to further explain the survey’s purpose and to increase credibility. In addition, company brochures or annual reports were suggested to provide as supplementary material. Once feedback became available, we requested the individual’s help in recommending other respondents. In total, this survey was carried out over three months (March 2014-June 2014).

Despite the use of a covering letter assuring data confidentiality, the response rate is quite low (12 responses) and, as expected, respondents skipped some questions for a certain reasons such as lack of accurate statistic data, data confidentiality. In this case, data availability was serving as the selection criteria of specific indicators and additional data was drawn from secondary sources such as the FV’s websites, brochures and research reports. Ultimately, out of 150 FVs in the original sample, 20 FVs were selected for further analysis. The available data are summarized in Table 1. To lessen the impact of large differences in data magnitudes (scaling difficulties), the normalized data is thus suggested to execute before the efficiency value calculation (Sarkis and Talluri, 2004).

Table 1 Summary statistics of the dataset.

Variables	Mean	Median	St. deviation	Minimum	Maximum
Input variables					
Total area	253.21	212.00	273.04	22.00	1311.80
Intermodal area	25.80	12.00	40.92	0.03	180.00
Warehouse area	56.67	33.50	80.61	0.25	315.00
Amount of investment	1190.50	149.00	3374.72	18.00	14376.00
Output variables					
Number of jobs	3131.60	1750.00	3863.28	22.00	13000.00
Amount of goods handled	18.35	6.00	25.97	0.10	80.00
No. companies attracted	101.75	96.50	90.70	2.00	270.00

3.3 The determining of variables

To further confirm whether the selection of input and output variables is able to fully explain the effect on efficiency, "isotonicity" principle-the increase of an input will not decrease output of another item-need to be verified (Liu 2008). Yadav et al. (2011) suggested executing Pearson's correlation analysis to test and verify "isotonicity". That is, if the correlation of the selected input and output is positive, the factors are isotonically related and can be included in the analysis; otherwise, the variable should be omitted. The result of the Pearson's correlation coefficient and its significance level is shown in Table 2. The correlations of the total amount of investment with all of the output variables are found to be negative and thus should be excluded.

Table 2 Correlation result among variables.

Items	x_1	x_2	x_3	x_4	y_1	y_2	y_3
x_1	1						
x_2	0.509^{*}(.022)	1					
x_3	0.856^{**}(.000)	0.545^{*}(.013)	1				
x_4	0.01(.000)	-0.01(.099)	0.01(.000)	1			
y_1	0.716^{**}(.000)	0.483^{*}(.031)	0.609^{**}(.004)	-0.05(.014)	1		
y_2	0.247(.294)	0.534^{*}(.015)	0.114(.632)	-0.13(.468)	0.345(.136)	1	
y_3	0.645^{**}(.002)	0.319(.171)	0.474^{*}(.035)	-0.10(.000)	0.62^{**}(.004)	0.545^{*}(.013)	1

* Correlation significant at the 0.05 level. ** Correlation significant at the 0.01 level.

Significantly, the correlation coefficients among inputs and outputs are relatively high (r>0.5); for instance, total area and intermodal area, number of companies and employees. This might be questioned by researchers(e.g., Lau 2012) who advocated that variables highly correlated with existing model variables are merely redundant and thus should be removed. However, the use of pairwise correlation should only be seen as a tool for the identification of candidate inputs and outputs and the actual decision should be based on much broader consideration (Dyson et al. 2001; Podinovski and Thanassoulis 2007). In this study, except 'the amounts of investment', we retained the rest of variables taking the following reasons into account (i) the correlation results derived from small sample (twenty FVs) cannot serve for wider reference (ii) in reality, the size of a FV does not always positively associate with intermodal terminal and warehouse (iii) managers may wish to investigate the roles warehouses and intermodal terminals in FV performance. Table 3 defines these variables and provides the corresponding explanation.

Table 3 The definition of variables.

Items	Variables	Description	units
Inputs	Total area	Total area already currently developed , not including the area for further expansion	Hectares
	Intermodal terminal area	The total area of intermodal terminal	Hectares
	Warehouse area	The total area of warehouse	Hectares
Outputs	Number of employees	The number of employees of companies that rented facilities are working in FV	Number
	Annual load handling	Annual load traffic generated by the facilities offered by FV	Million Tons
	Number of companies settled	Number of companies on site	Number

With respect to the sample size, as rule of thumb, some researchers suggested such following relationships among the number of DMUs (n), inputs (m) and outputs (s) to to obtain sufficient discrimination power: $n \geq 2(m+s)$ (Golany and Roll 1989), $n \geq 2m \times s$ (Dyson et al. 2001), $n \geq \max\{m \times s;\ 3(\mathrm{m+s})\}$ (Cooper et al. 2007). Given $m=3$ and $s=3$, the sample size ($n=20$) used in this study exceeds the desirable size and thus the rule of thumb works well.

4 Empirical results and analysis

The data were evaluated using MaxDEA Pro 6.3 (Chen and Qian 2010), as well as Matlab 2014 and SPSS 21. The observations of 20 European FVs are taken in 2013, the latest available period of observation. Both CCR and BCC models were applied for the lack of precise information on the returns to scale of the FV production function.

4.1 Efficiency value analysis

Table 4 shows the results obtained from the CCR and BCC model to determine the efficiency of FVs under study. As previously noted, the BCC model identifies technical efficiency (TE) alone, while the CCR model measures overall efficiency (OE) which is the combination of technical efficiency (TE) and scale efficiency (SE). Hence, the BCC model, as expected yields higher values than the CCR model, with respective average values of 0.840 and 0.710. With a closer look, the CCR efficiency scores range from 0.2631 to 1, with an overall mean and standard deviation of 0.71 and 0.27, respectively. Among them, 35% of FVs present at overall efficiency, with efficiency scores equal to one, whilst 84.61% has OE scores below the mean score(0.71). For inefficient units, they can improve efficiency by enhancing outputs while maintaining the same proportions of input. In particular, the BCC model is applied to determine the sources of inefficiencies

Table 4 The results of the CCR and BCC efficiency model.

No.	FVs	Aggregate efficiency	Technical efficiency	Scale efficiency	$\sum\lambda^*$	Returns to scale
1	Eurocentre Toulouse	**1**	**1**	1	1	CR
2	GVZ Berlin Süd Großbeeren	0.3105	0.3534	0.8785	1.1759	DR
3	GVZ Bremen	0.834	**1**	0.834	1.7163	DR
4	GVZ Dresden	0.6456	**1**	0.6456	0.0988	IR
5	GVZ Europark	0.5889	0.5959	0.9882	0.8722	IR
6	GVZ Nürnberg	**1**	**1**	1	1	CR
7	Interporto Bologna	0.66	0.6769	0.975	0.695	IR
8	Interporto Novara	0.3607	0.4088	0.8824	0.2863	IR
9	Interporto Padova	0.6808	0.6961	0.978	0.68	IR
10	Interporto Parma	0.5049	0.5144	0.9816	0.7011	IR
11	Interporto Rovigo	**1**	**1**	1	1	CR
12	Interporto Venezia	0.5208	**1**	0.5208	0.0838	IR
13	Interporto Verona	**1**	**1**	1	1	CR
14	Interporto Marche spa	**1**	**1**	1	1	CR
15	Interporto Nola Campano	0.6667	0.6701	0.9949	1.0119	DR
16	Interporto Quadrante Europa	**1**	**1**	1	1	CR
17	Interporto Rivalta Scrivia	**1**	**1**	1	1	CR
18	Interporto Torino	0.8818	0.8897	0.9911	0.9178	IR
19	PLAZA	0.2821	**1**	0.2821	4.8398	DR
20	TVT	0.2632	**1**	0.2632	0.3473	IR
	Mean	0.71	0.8403	0.8608	1.0213	
	SD	0.2714	0.2264	0.2388		

Notes: IR-increasing returns to scale; CR-constant returns to scale; DR-decreasing returns to scale; $\sum\lambda^*$ sum of optimized value of λ

present in the CCR efficiency. Of the twenty FVs, 60% are found to be technically efficient, while the remaining eight are identified as technically inefficient and their efficiency score lies between 0.3534 and 0.8897.

Surprisingly, a number of FVs that far away from the CCR frontier are now observed to be efficient in the BCC model. Purely technically efficient FVs, such as Bremen, Dresden, Venezia, PLAZA and TVT increased to become efficient ones. This suggests that the inefficiencies assigned to these five FVs, with respect to CRS assumption, are purely scaled-based inefficiencies. Particularly, seven FVs with a remarkable efficiency score equal to one reveal themselves to be overall, technically and scale efficient. This consistency reflects that the operation of these FVs is at the most productive scale size and has efficient operations. In addition, eight overall inefficient FVs are ranked as such mainly due to their technical inefficiency because their TE scores are lower than their SE scores. The scale efficiency of FVs indicates that almost half of the FVs (45%) are characterized by IRS followed by CRS (35%). And only 20% of them operate at DRS. In sum, 13 FVs are found to be scaled inefficiently, implying that 65% FVs present an unbalanced status of scale. It can be identified from Table 4 that the lowest scale efficiency is calculated for the TVT (0.2632), followed by PLAZA (0.2821). The results imply that several FVs are technically inefficient and relative scales of these operations have unbalanced status and require attention for efficiency improvement. For instance, FVs found to be operating under an IRS may prefer to expand their operations in the future. On the contrary, for those operating at DRS, their scale sizes need to be decreased for efficiency improvement. The results of ANOVA ($F=1.437$, $p=0.05$, critical value=3.03) analysis and Spearman's rank order correlation coefficient ($r=0.486$) further confirm our statement. As a consequence, the choice of these two methodologies applied in our study has no apparent impact on the estimated average efficiency scores.

4.2 Slack analysis

Slacks provide the vital information pertaining to the areas which an inefficient DMU needs to improve its drive towards attaining the status of an efficient one (Kumar and Gulati 2008, p.558). In this study, slack analyses are executed under CCR assumptions to obtain the long-term improvement directions for the inefficient FVs.

According to Table 5, the slack value of CCR demonstrates that most of the FVs are inefficient due to poor annual goods load handling from the output side, while from the input side the intermodal terminal and warehouse could be greatly reduced. Overall, eleven FVs have non-zero slacks for "intermodal area", while nine have non-zero slacks for "warehouse area" and one has non-zero slack for "total area". Specifically, only Europark has to decrease "total area" by 137.367. Bremen and PLAZA have the greatest excesses in the input variable "intermodal area" and "warehouse area". With respect to output slacks, only 50% of FVs have an "annual goods load handling" slacks equal to zero. This indicates that the other 50% of FVs does not obtain satisfying results in this aspect. In particular, PLAZA, the largest platform in Europe, requires the greatest increase of 157.84 in "annual goods load handling". In addition, Europark, Novara and

Padova need to increase their output standards for "number of job creation", while Bremen, Dresden and Novara should attract more companies settled in their campus.

As a whole, for most of FVs, "total area", "companies settled" and "number of employees" are three variables that do not require much adjustment. However, in general, utilization is poor for the "intermodal area" and "warehouse". In terms of output factors, augmenting the level of goods handed could enable most inefficient FVs to move to the efficiency frontier. Eurocentre Toulouse, Nürnberg, Rovigo, Verona, Quadrante Europa, Rivalta Scrivia perform well with both input and output slack variables of zero.

Table 5 CCR slack analysis of inefficient FVs.

Freight Villages	CCR	slack values: Size	Intermodal area	Warehouse area	Employees	Goods handled	Companies settled
BerlinSüd	0.310	0	-24.340	-203.700	0	40.523	0
Bremen	0.834	0	-158.345	-64.552	0	0	14.748
Dresden	0.646	0	-5.449	0	0	0	8.287
Europark	0.589	-137.37	0	0	2323.837	8.613	0
Bologna	0.660	0	-1.958	-26.508	0	30.695	0
Novara	0.361	0	-11.106	0	691.898	0	40.245
Padova	0.681	0	-20.720	-10	277.440	47.952	0
Parma	0.505	0	-6.173	-42.243	0	35.396	0
Venezia	0.521	0	-7.583	-9.225	0	1.630	0
NolaCampano	0.667	0	-2.321	-15.702	0	32.095	0
Torino	0.882	0	0	-28.446	0	14.436	0
PLAZA	0.282	0	-27.996	-233.242	0	157.843	0
TVT	0.263	0	-7.005	0	0	1.290	0
No. DMUs with slacks		1	11	9	3	10	3

Note： Negative value means suggest reduction of input parameters.

4.3 Sensitivity analysis

This section reports the sensitivity analysis results. In order to avoid redundancy, only BCC efficiency scores were scrutinized.

4.3.1 Removal of variables

According to the rules in Section 2.2, sensitivity analysis was conducted (see Table 6). The three FVs Quadrante Europa, Nürnberg and Eurocentre Toulouse received identical TE value across different criteria; while the others experienced sort of variation. Specifically, without “total area”, half of the efficiency score was reduced. Notably, Venezia and TVT reacted significantly, with efficiency scores dropping to 91.7% and 74%, respectively. Interestingly, the same situation occurred when removing the “number of companies”. That means that the “total area” and “number of companies” are critical to those two FVs. GVZ Dresden is similar. After taking away the “intermodal area”, four FVs changed their efficiency scores. Among them, three decreased slightly, the notable exception being Rivalta Scrivia, which changed from one to 0.3823, with a dropping rate of 61.77%. Notably, some FVs were sensitive to variable change and rapidly become inefficient by changing a few variables. For instance, when “warehouse area” was excluded from the input list, the TE score of Marche reduced from 1 to 0.6556. In regard to outputs, the efficiency values ranged from 0.2037 to 1 and 65% of FVs retained their efficiency values when we removed “employees”. The influence of “goods handled” on the TE value was not apparent, because only three FVs changed their efficiency values, and Verona was more sensitive to this change in comparison to the other two.

Overall, the “number of companies” variable heavily influences the TE score for most FVs, with a changing rate of up to 70%, followed by “total area” (50%). Less than 50% of sampled FVs shifted their efficiency value after removing the remaining variables. In particular, if we delete “number of companies”, six FVs (Toulouse, Dresden, Rovigo, Venezia, Rivalta Scrivia, TVT) change from full efficiency status to non-efficiency, particularly Interporto Venezia, which experienced the greatest variation; on the contrary, four relatively inefficient FVs (Berlin, Süd Großbeeren and Europark) turn out to be efficient. The Person correlation coefficient between the full BCC model and changed models ranges from 0.591 to 0.991, implying that the results are robust for these different efficiency scores. However, one exception is the scenario of removing “number of companies”, which has a positive but low coefficient of 0.052 with full BCC. This result also confirms that this variable heavily influences the BCC efficiency score.

Table 6 Sensitivity analysis results by removal of variables.

DMUS	Full BCC	Efficiency value without input: Total area	Intermodal area	Warehouse area	Efficiency value without output: Employees	Goods handed	Companies
Eurocentre Toulouse	1	1	1	1	1	1	**0.9231**
Berlin Süd	.3534	0.3224	0.3534	**0.3534**	**0.2037**	0.3534	**1**
Bremen	1	1	1	1	**0.9625**	**0.7981**	1
Dresden	1	**0.5589**	1	1	1	1	**0.3077**
Europark	0.5959	0.5959	**0.5711**	**0.1679**	0.5959	0.5959	**1**
Nürnberg	1	1	1	1	1	1	1
Bologna	0.6769	0.4436	**0.6729**	0.6769	**0.6367**	0.6769	**0.1910**
Novara	0.4088	0.3293	0.4088	**0.3597**	0.4088	**0.1912**	**0.4719**
Padova	0.6961	0.4630	0.6961	0.6961	0.6961	0.6961	**0.3210**
Parma	0.5144	0.3337	0.5144	0.5144	**0.4789**	0.5144	**0.4088**
Rovigo	1	1	1	**0.0746**	1	1	**0.1426**
Venezia	1	**0.0826**	1	1	1	1	**0.2130**
Verona	1	1	1	1	1	**0.5931**	1
Marche	1	1	1	**0.6556**	1	1	1
Nola Campano	0.6701	0.6567	0.6701	0.6701	**0.6481**	0.6701	**1**
Quadrante Europa	1	1	1	1	1	1	1
Rivalta Scrivia	1	1	**(0.3823)**	1	1	1	**0.4486**
Torino	0.8897	0.8152	**0.8505**	0.8897	**0.8737**	0.8897	**1**
PLAZA	1	1	1	1	**0.9259**	1	1
TVT	1	**0.2593**	1	1	1	1	**0.3846**
Average	0.8403	0.6930	0.8060	0.7529	0.8215	0.7989	0.6906
Efficient DMUs	12	9	11	10	10	10	10
Changing rate		50%	20%	20%	35%	15%	70%

Note: Changing rate=Number of changing DMUs (compared to the basic BCC model)/Total number of DMUs (20)*100%;

4.3.2 Removal of efficient DMUs

Twelve additional DEA analyses were performed on the basis of VRS assumption to test the robustness of the DEA results with regard to stability of reference set and outliers. The results in Table 7 show that the average TE scores vary between 0.7886 and 0.8873 with a standard deviation range of 2.0267 to 2.6491. Although deleting Rovigo and Quadrante Europa shifts the mean value relatively significantly, the overall fluctuation is not apparent. For this reason we argue that removing efficient units does not shift the average TE score significantly, and thus none of the efficient FVs in the DEA analysis is extreme. In terms of the reference set, in 11 out of 12 cases the reference set remains unaltered. Further, Spearman's rank correlation coefficient was used to gauge the similarity of efficiency ranking between the model with full DMUs and those based on removing each efficient DMU at a time. Table 7 shows that these coefficients range from 0.828 to 1.0 and are significant at 99%. The high rank correlation coefficient indicates that rankings are stable in regard to efficiently FVs defining the efficient frontier, further confirming the robustness of the efficiency analysis.

Table 7 Results of the jack-knifing analysis.

FVs removed from analysis	Mean TE	SD.	NE DMUs	Coefficient	New DMUs in the reference set
Eurocentre Toulouse	0.8794	2.4345	12	0.910**	None
Bremen	0.8319	2.2295	11	1.000**	None
Dresden	0.8336	2.2289	11	1.000**	None
Nürnberg	0.8382	2.4456	12	0.963**	None
Rovigo	0.7886	2.0267	10	0.987**	None
Venezia	0.8325	2.2293	11	0.987**	None
Verona	0.8322	2.2294	11	1.000**	None
Marche	0.8336	2.229	11	1.000**	None
Quadrante Europa	0.8873	2.6491	13	0.828**	Torino, Europark
Rivalta Scrivia	0.8319	2.2295	11	1.000**	None
PLAZA	0.8321	2.2294	11	0.963**	None
TVT	0.8319	2.2295	11	1.000**	None
Full BCC model	0.8403	0.2264	12		

Note: (i) NE: the number of efficient DMUs; (ii) ** Correlation is significant at the 0.01 level (2-tailed)

4.4 Benchmark analysis

In this part, we investigate the role that an efficient FVs plays in benchmarking inefficiency FVs. In doing so, Zhu (2000) recommended two possible approaches: (i) count the number of times a particular efficient unit acts as referent DMU; (ii) benchmark share measure.

4.4.1 Number of peer count

The peer count number measures the extent to which the performance of an efficient units can be useful for the non-efficient ones (Mostafa 2007). A FV that frequently appears in the reference set is likely to be a genuinely efficient unit and is probably an exemplary operating performer. On the other hand, those seldom appearing in the reference set of other FVs are likely to possess a very uncommon input/output mix and are thus not suitable examples of other inefficient ones.

By accounting the reference frequencies of the efficient FVs when both CCR and BCC models are applied (show in Fig.2), 15 FVs are regarded as the reference set for inefficient ones. In particular, Quadrante Europa appears most frequently as a peer in both CCR (13 times) and BCC (10 times), followed by Eurocentre Toulouse in CCR (9 times) and BCC (3 times). Seven FVs are treated as a reference set in both the BCC and CCR models: Quadrante Europa, Bologna, Padova, Marche, Verona, Nürnberg, and Eurocentre Toulouse. Here, it is worth noting that although some FVs have an efficiency score equal to one, there is no reference from a unit other than itself, such as GVZ Bremen, TVT, which might because models employed in this paper are based on self-appraisal rather than peer assessment.

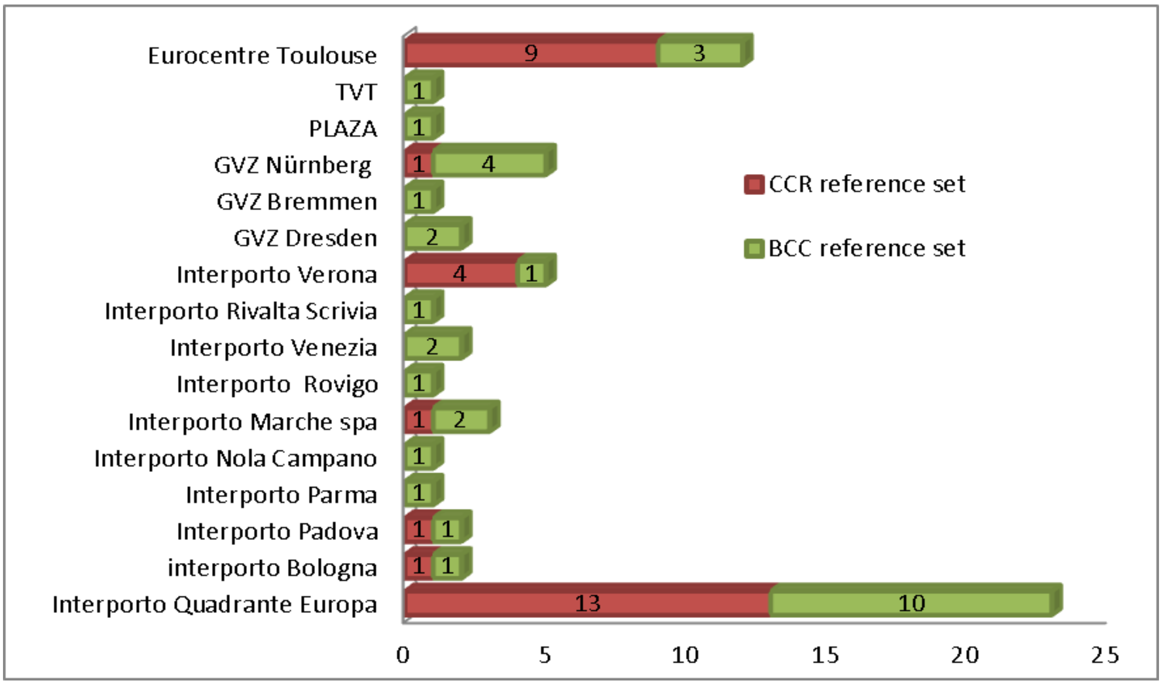

Fig.2. Reference set frequencies under the CCR and BCC models

4.4.2 Benchmark share measure

The benchmark share measure is developed to further characterize the performance of efficient units (Yadav et al. 2011). Table 8 summarizes the benchmark share of the technically efficient FVs, with the ranking mentioned in parentheses and ordered by the average rank of the efficient units.

As presented in Table 8, of the total 72 benchmark share measures, 18 are greater than 10% and 4 in particular are greater than 50%. Appropriately, 45% of benchmark share measures show no effect on inefficient FVs. In particular, Quadrante Europa, which is a highly technically efficient FV, has the biggest benchmark share in job creation (67.76%). In addition, Eurocentre Toulouse also has outstanding benchmark shares in terms of "goods handled" and "number of companies", with benchmark shares of 56.89% and 57.62%, respectively. As far as input "total area" is concerned, Venezia contributes the highest benchmark shares for other inefficient FVs. Rivalta Scrivia (36.75%) and Marche (58.06%) occupy the first rank in terms of benchmark shares for input "intermodal terminal" and "warehouse", respectively. All the above stated FVs with the highest share measure are overall efficient except Venezia (52.08%). These benchmarks may offer a first guideline for the performance improvement of other FVs.

Table 8 Benchmark shares of 12 efficient FVs.

	Output factors			Input factors			*Average*
DMUs	Y1 (%)	Y2 (%)	Y3 (%)	X1(%)	X2(%)	X3(%)	*rank*
Quadrante Europa	**67.76 (1)**	3.55 (4)	1.77 (6)	13.11 (4)	30.78 (2)	23.41 (2)	3.17
Marche	0.00 (10)	0.00 (10.5)	2.86 (5)	0.00 (9.5)	0.00 (9.5)	**58.06 (1)**	**7.58**
Rovigo	1.71 (5)	1.09 (6)	1.55 (7)	0.00 (9.5)	1.64 (5)	0.39 (5)	6.25
Venezia	7.93 (3)	0.00(10.5)	17.91 (2)	**39.20 (1)**	0.00 (9.5)	0.00 (9)	5.83
Rivalta Scrivia	0.00 (10)	0.48 (7)	0.29 (8)	0.00 (9.5)	**36.75 (1)**	0.00 (9)	7.42
Verona	0.00 (10)	25.47 (2)	0.00 (10.5)	0.38 (6)	0.00 (9.5)	0.00 (9)	7.83
Dresden	0.12 (7)	9.92 (3)	3.39 (4)	25.43 (2)	13.98 (4)	0.00 (9)	4.83
Bremen	0.00 (10)	0.12 (8)	0.00 (10.5)	0.00 (9.5)	0.00 (9.5)	0.00 (9)	9.42
Nürnberg	17.56 (2)	2.47 (5)	14.60 (3)	3.17 (5)	15.27 (3)	5.93 (4)	3.67
PLAZA	0.30 (6)	0.00 (10.5)	0.00 (10.5)	0.00 (9.5)	0.00 (9.5)	0.00 (9)	9.17
TVT	0.00 (10)	0.00 (10.5)	0.00 (10.5)	0.00 (9.5)	0.00 (9.5)	0.00 (9)	9.83
Eurocentre Toulouse	4.62 (4)	**56.89 (1)**	**57.62 (1)**	18.70 (3)	1.57 (6)	12.21 (3)	3.00
Total	100	100	100	100	100	100	

To take a closer look at each output variable, such as "number of companies", Fig.3 uses a pie-diagram to show the benchmark share of technically efficient FVs. Accordingly, Eurocentre Toulouse alone refers to over half of the potential improvement in attracting companies on site (57.62%). Interporto Quadrante Europa and GVZ

Nürnberg have benchmark shares of more than 10%. By contrast, the remaining efficient ones cannot exert much influence on inefficient units. The similar case also presents in two other output variables.

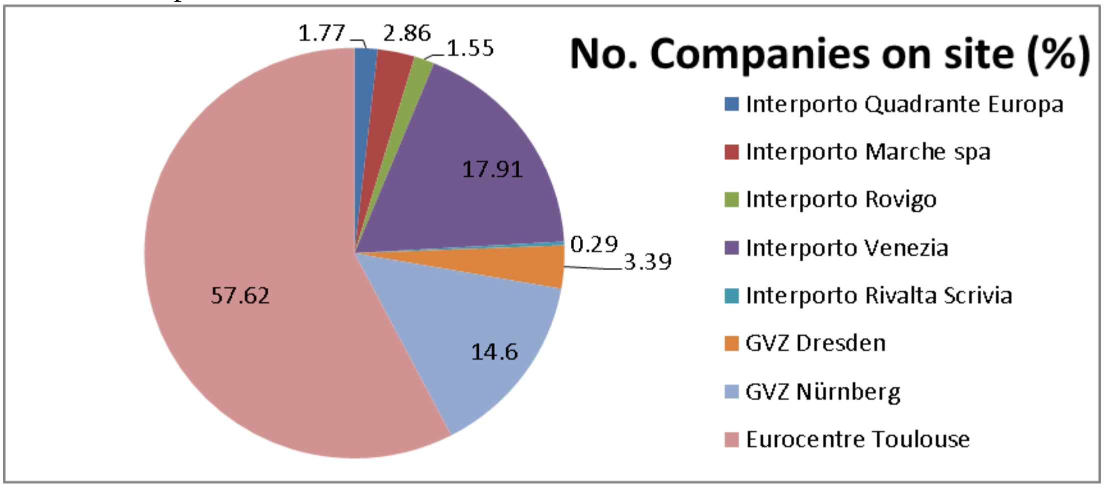

Fig.3. Share of efficient FVs for efficiency improvement.

On the other hand, FVs like PLAZA, TVT, Bremen and Rovigo are poorly benchmarked by inefficient units with benchmark shares below 10%. For most of the inputs and outputs the share is 0%. Taking a closer look, these FVs are technically efficient and are termed as self-evaluators, which share similar conclusion with section 4.4.1.

4.5 Hypothesis testing

In this section, two additional analyses are conducted (i) to identify whether the efficiency scores depend on the FV's region and (ii) to test whether the size of FV will affect the efficiency score. The CCR-efficient scores are chosen for analysis because CCR model can discriminate more adequately among the units analysed than the BCC model.

4.5.1 Regional differences in efficiency scores

The sampled FVs in our study come from different countries and 80% of them play a leading role in their countries. In previous studies, given relatively small samples, no analysis based on a subgroup was conducted. However, we remain interested in investigating whether the efficiency scores vary across different countries. Accordingly, the first hypothesis is proposed: there is no difference in efficiency scores of FVs from different countries.

Twenty FVs were grouped into five subgroups according to their locations. The number of FVs from different countries and the average efficiency score for each group is presented in Table 9. While the focus can only be put on the scrutiny subgroup of

Germany and Italy, the remaining category is not representative due to only one FV included. As the tested efficiency scores are not normally distributed, two non-parametrical tests, Kruskal-Wallis and Mann-Whitney, were applied to test whether the efficiency scores differ between subgroups. According to Table 10, the p-values of 0.389 indicate that there are no reasons for rejecting the null hypothesis with a significant level of =0.05.

Table 9 Average efficiency scores according to countries.

Countries	France	Germany	Italy	Spain	Portgal
Number of units	1	5	12	1	1
Average effiicney score	**1**	0.6758	0.77298	0.2821	0.2632

Table 10 Results of the Kruskal-Wallis and Mann-Whitney tests for differences between Germany and Italy.

Kruskal-Wallis test (α=0.05)	Results
Chi-square	0.839
df	1
p-value	0.389
Mann-Whitney test (α=0.05)	**Results**
U	21.5
Z	-0.916
Exact Sig. (2-tailed)	0.389

4.5.2 Freight villages' size and efficiency score

Previous studies showed that there are differences in efficiency scores between small and large distribution systems or warehouses (Andrejić et al. 2013; Banaszewska et al. 2012; Hamdan and Rogers 2008). Accordingly, we are interested to investigate whether the efficiency scores differ among groups. As there is no standard classification about FV size, two approaches are applied to classify FV size (i) FVs less than 150ha are small and those over than 150ha are large (ii) small FVs are less than 100ha, large FVs are over as 250ha and the size between 100 and 250 are considered as medium. The average efficiency scores and number of units are presented in Table 11.

Table 11 Average efficiency scores according to the size of Freight Villages.

	Testing approach				
	2 groups (ha)		3 groups (ha)		
Group	Small	Large	Small	Medium	Large
Criteira	<150	>150	<100	[100,250]	>250
Number of units	7	13	6	8	6
Average Efficiency	0,675326	0,760637	0,727761	0,768854	0,856227

Since the efficiency scores do not fit within a standard normal distribution, the Mann–Whitney U-test is adapted in the context of two groups. Obviously, with a p-value much larger than 0.05, we cannot reject the null hypothesis and state that there is no significant difference between large and small FVs. In terms of the three groups, as presented in Table 12, the Kruskal-Wallis test was run, and with significance at 0.05 levels, we also cannot confirm there is significant difference among three subgroups of FVs.

Table 12 Results of the Mann-Whitney and Kruskal-Wallis tests for differences between FVs of different size.

Two groups Mann-Whitney test(α=0.05)		Three groups Kruskal-Wallis Test(α=0.05)	
U	42	Chi-Square	0.763
Z	-0.283	df	2
Exact Sig. (2-tailed)	0.813	P-value	0.683

Asymp. Sig. (2-tailed) =0.777

5 Discussion and Conclusions

This study attempts to provide a compelling answer to the problem of assessing the relative efficiency levels of FVs in Europe. With the application of DEA model, twenty FVs have been estimated in terms of relative efficiency scores, slack analysis, sensitivity analysis, benchmark analysis and hypothesis testing.

The results of analysis demonstrate that seven FVs are observed to be inefficient in both the CCR and BCC models, while eight FVs suffer from technical and scale inefficiency. The mean technical efficiency score is found to be 84.03%, and twelve FVs are technically efficient. Only 35% FVs operate at constant returns to scale and the rest need to adjust their operating scale for efficiency improvement. The slack analysis shows that most of the inefficient FVs need to reduce their use of intermodal area and warehouse and augment the amount of goods load handling to move closer to the efficiency frontier. Based on the reliability test, our results are stable across all criteria and none of the efficient units has been observed to be extreme. The composition of the reference set remained unaltered in the most cases. A benchmark analysis was conducted to further identify important variables and efficient FVs as a way for inefficient ones to arrive at the efficient frontier. Comprehensively, Interporto Quadrante Europa and Eurocentre Toulouse dominate the benchmark share and are frequently referenced by inefficient ones. In the last step, statistical tests were applied to investigate whether differences in efficiency scores exist among countries and the size of FVs. Nonetheless, it should be noted that the conclusion derived from the hypothesis testing is based on a relatively small sample.

As far as the management application is concerned, we are interested in comparing the SWOT-based benchmark study with present research. With limited space, the comparison is executed from a broad perspective. Based on different research perspectives (practical and academic), both study are expected to draw general inferences for the development of FVs. For research purpose, it seems worthless to compare the rank of FVs; instead, our attention is restricted to FVs presenting significant differences between two studies. For example, GVZ Bremen is the oldest and largest example of a FV developed in Germany, having absorbed 8000 employees – an outstanding figure for Europe. However, this advantage is not reflected in our study, as this FV revels scale inefficiency and operates in decreasing returns to scale. This demonstrates that, by handing multiple inputs and outputs, the DEA can provide more information on performance assessment and improvement. Interporto Bologna, one of the leading FVs in Italy also deserves more attention; nevertheless, it is identified as inefficient in both the CRS and VRS assumptions and is operates on IRS. According to our research, it should expand its scale for future efficiency operation. In reality, over 200 hectares of land are to be developed for future expansion of Interporto Bologna. In this case, we might expect our study to provide useful references for the strategic planning of FVs.

The contribution of this paper is to enrich the body of previous FV performance assessments. At the first time, this paper introduces the DEA method in the context of FVs for efficiency measurement in a systematically manner by (i) the extension of input and output variables and sample size (ii) providing useful insights for FV benchmarking with multiple analysis perspectives. Taking the advantage of DEA and the complexity of FVs into account, this paper showed why DEA is a feasible benchmarking approach for FVs. However, DEA is a methodology which relies on accessible information. Since internal data on management are hard to access, more effort has been put to overcoming the obstacles, such as the use of proxy items and reference existing survey. Indeed, if more data were available, FV efficiency could be more thoroughly explored and detailed. It would be extremely helpful if government standardize the data collection and openly publish data, as this would enable fair and transparent comparisons.

It should be noted that this research is an exploratory study; the purpose is not to achieve definitive results (e.g. ranking FVs) for the direct use of management. Rather, it draws attention to the value of benchmarking in an effort to measure the performance of FVs and serve as a management tool. In the future, some extensions can be envisaged. First, in view of the limited number of FVs analysed and the relatively small set of inputs and outputs used in present analysis, further studies are recommended to maximize the sample size and consider a wider range of inputs and outputs. For increased strategic relevance and reliable results, future research in FVs measurement should strive to cover longer time spans. Second, instead of output-orientation standard models, input-oriented models and other extensions can also be utilized to measure more subtleties in reality. Network-DEA would be suitable for opening the black box of FVs for further investigation, too. Third, to further confirm the comparability of FVs, future research can divide FVs into various clusters in terms of size, facilities and function, and only FVs belonging to the same cluster are included and compared. Last but not least, other decision-making tools such as AHP or techniques for dealing missing and fuzzy data should be involved to assist the application of DEA to FVs.

Acknowledgments

The authors would like to thank Dr.Claudio.di.Ciccio for the questionnaire translation and all of the participants who kindly provide useful information for our survey.

Compliance with Ethical Standards

We express our appreciation for the State Scholarship Fund of China Scholarship Council (Grant no.2011613013). We herewith state that there are no potential conflicts of interest (financial or non-financial) and that the submission complies with the Ethical Rules of the Central European Journal of Operations Research (CJOR).

Reference

Andrejić M, Bojović N, Kilibarda M (2013) Benchmarking distribution centers using Principal Component Analysis and Data Envelopment Analysis: A case study of Serbia. Expert Syst Appl 40(10):3926–3933

Ballis A (2006) Freight villages: Warehouse design and rail link aspects. Transp Res Rec 1966(1):27–33

Ballis A, Golias J (2002) Comparative evaluation of existing and innovative rail–road freight transport terminals. Transp Res Part A Policy Pract 36(7):593–611

Banaszewska A, Cruijssen F, Dullaert W, Gerdessen JC (2012) A framework for measuring efficiency levels–The case of express depots. Int J Prod Econ 139(2):484–495

Banker RD, Charnes A, Cooper WW (1984) Some models for estimating technical and scale inefficiencies in data envelopment analysis. Manag Sci 30(9):1078–1092.

Bhanot N, Singh H (2014) Benchmarking the performance indicators of Indian Railway container business using data envelopment analysis. Benchmark Int J 21(1):101–120

Cassone, G. C., & Gattuso, D. (2010). Models of intermodal node representation. Retrieved from http://www.openstarts.units.it/xmlui/handle/10077/6166

Cezar-Gabriel C (2010) Performance assessment in operating Dry Ports. Ann Faculty Econ 1(2):934–938

Chakraborty PS, Majumder G, Sarkar B (2011) Performance measurement of distribution center combining Data Envelopment Analysis and Analytic Hierarchy Process. Adv Ind Eng Manag 6(2):117–128

Charles V, Kumar M, Irene Kavitha S (2012) Measuring the efficiency of assembled printed circuit boards with undesirable outputs using data envelopment analysis. Int J Prod Econ 136(1):194–206

Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2(6):429–444

Chen G, Qian Z (2010) MaxDEA LP Manual. http://www.maxdea.cn/. Accessed 14 June 2014

Cooper WW, Seiford L M, Tone K (2007) Data envelopment analysis: A comprehensive text with models, applications, references and DEA-Solver Software. Springer, New York

Cullinane K, Wang TF, Song DW, Ji P (2006) The technical efficiency of container ports: comparing data envelopment analysis and stochastic frontier analysis. Transp Res Part A Policy Pract 40(4):354–374

de Carvalho CC, Lima Jr OF(2010) Efficient logistic platform design: the case of Campinas Platform. http://www.abepro.org.br/biblioteca/enegep2010_TI_ST_113_741_17234.pdf. Accessed 20 February 2013

De Witte K, Marques RC (2010) Designing performance incentives, an international benchmark study in the water sector. Cent Eur J Oper Res 18(2):189-220

Dyson RG, Allen R, Camanho AS, Podinovski VV, Sarrico CS, Shale EA(2001) Pitfalls and protocols in DEA. Eur J Oper Res 132(2):245–259

Europlatforms. E.E.I.G. (2004) Logistics Centres Directions for Use, Europlatforms EEIG.http://www.unece.org/fileadmin/DAM/trans/main/eatl/docs/EN-REV-What is a Freight VillageFinalcorretto.pdf. Accessed 12 March 2012

George SA, Rangaraj N (2008) A performance benchmarking study of Indian Railway zones. Benchmark Int J 15(5):599–617

Golany B, Roll Y (1989) An application Procedure for DEA. Omega 17(3):237–250

Hamdan A, Rogers KJ (2008) Evaluating the efficiency of 3PL logistics operations. Int J Prod Econ 113(1):235–244

Haralambides H, Gujar G (2012) On balancing supply chain efficiency and environmental impacts: An eco-DEA model applied to the dry port sector of India. Marit Econ Logis 14(1):122–137
Ho CT, Zhu DS (2004) Performance measurement of Taiwan's commercial banks. Int J Prod Perform Manag 53(5):425–434
Kapros S, Panou K,Tsamboulas D (2005) Multi-criteria approach to the evaluation of intermodal freight villages. Transp Res Rec 1906(1):56–63
Koch H, Münch S, Nestler S, Nobel N(2010) Ranking of the European Freight Village locations-Benchmarking of the European experience. wvb Wissenschaftlicher Verlag, Berlin.
Kumar S, Gulati, R (2008) Evaluation of technical efficiency and ranking of public sector banks in India: An analysis from cross-sectional perspective. Int J Prod Perform Manag 57(7):540–568
Lau KH (2012) Distribution network rationalization through benchmarking with DEA. Benchmark Int J 19(6):668–689
Liang L, Wu J, Cook WD, Zhu J (2008) Alternative secondary goals in DEA cross-efficiency evaluation. Int J Prod Econ 113(2):1025–103
Liu CC (2008) Evaluating the operational efficiency of major ports in the Asia-Pacific region using data envelopment analysis. Appl. Econ 40(13):1737–1743
Liu J, Dong J, Chang Q (2013) Assessing logistics park operational efficiency using dual variables DEA model. Statistics and Decision (8):60–62
Liu J, Zhou Y, Wei X (2010) International studies on logistics parks from 1999 to 2009, a literature review. Logi. Technol 29(8):43–45
Luo J (2013) Analysis on the operation effects of Logistics Park based on BP Neural Network. In: Proceedings of the 19th international conference on industrial engineering and engineering management, pp 1073–1079
Malekmohammadi N, Lotfi F H, Jaafar A B (2011) Data envelopment scenario analysis with imprecise data. Cent Eur J Oper Res 19(1):65-79
Manfred G, Martin P, Thouraya B (2008) Assessing performance indicators for inland terminals. http://simcont.net/fileadmin/simcont/downloads/Pub_PI.pdf. Accessed 3 May 2010
Meidute I (2005) Comparative analysis of the definitions of logistics centres. Transport 20(3):106–110
Mostafa M (2007) Benchmarking top Arab banks' efficiency through efficient frontier analysis. Ind Manage Data Syst 107(6):802–823
Nataraja NR, Johnson AL (2011) Guidelines for using variable selection techniques in data envelopment analysis. Eur J Oper Res 215(3):662–669
Ondrich J, Ruggiero J (2002) Outlier detection in data envelopment analysis: an analysis of jackknifing. J Oper Res Soc 53(3):342–346
Podinovski, VV, Thanassoulis E (2007) Improving discrimination in data envelopment analysis: Some practical suggestions. J Prod Anal 28(1-2):117–126
Ramanathan, R (2003) An introduction to data envelopment analysis: a tool for performance measurement. Sage Publ. India

Ross AD, Droge C (2004) An analysis of operations efficiency in large-scale distribution systems. J Oper Manag 21(6):673–688

Sufian F (2007) Trends in the efficiency of Singapore's commercial banking groups: A non-stochastic frontier DEA window analysis approach. Int. J. Prod. Perf. Manag 56(2):99–136

Singh SK, Bajpai VK (2013) Estimation of operational efficiency and its determinants using DEA: The case of Indian coal-fired power plants. Int J Energy Sect Manag 7(4):409–429

Tsamboulas D, Dimitropoulos I (1999) Appraisal of investments in European nodal centers for goods-freight villages A comparative analysis. Transportation 26(4):381–398

Wang D, Chen JJ (2009) Analysis on Logistics Park performance appraisal based on Fuzzy Synthetic Evaluation method. Logistics Sci-Tech 10(32):79–82

Winkler H, Seebacher G (2011) Management of freight villages: findings from an exploratory study in Germany. Int J Logist-Res App 14(4):271–283

Wu J, Haasis HD (2013) Converting knowledge into sustainability performance of freight villages. Logis Rese 6(2-3):63–88

Yadav V K, Padhy NP, Gupta, HO (2011) Performance evaluation and improvement directions for an Indian electric utility. Energy Policy 39(11):7112–7120

Yu MM, Chen PC(2011) Measuring air routes performance using a fractional network data envelopment analysis model. Cent Eur J Oper Res 19(1):81-98

Zhu J (2000) Multi-factor performance measure model with an application to Fortune 500 companies. Eur J Oper Res 123(1):105–124

Zhu J, Shen ZH (1995) A discussion of testing DMUs' returns to scale. Eur J Oper Res 81(3):590–596

Appendix

(1)DEA-CCR Model

It is assumed that n DMUs are evaluated. Each DMU_j (j = 1, 2, ..., n) consumes a vector of inputs, $x_j = (x_{1j}, x_{2j}, ...x_{mj},)^{\mathrm{T}}$ to produce a vector of output $y_j = (y_{1j}, y_{2j}, ...y_{sj},)^{\mathrm{T}}$. .The superscript T represents transpose. The DMU to be evaluated is designated as $DMUo$ and its input-output vector is denoted as (x_0, y_0). The output-oriented CCR model involves two-stage DEA processes, which can be expressed as follows:

$$
\begin{aligned}
&\max \varphi + \varepsilon\left(\sum_{i=1}^{m} s_i^- + \sum_{i=1}^{m} s_i^+\right) \\
&\text{s.t.} \quad \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- = x_{io} \qquad i = 1, 2, ...m; \\
&\qquad\ \ \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ = \varphi y_{ro} \qquad r = 1, 2, ...s; \\
&\qquad\ \ \lambda_j \geq 0 \qquad\qquad\qquad\quad j = 1, 2, ...n
\end{aligned}
\tag{1}
$$

Here ε is a non-Archimedean infinitesimal, which is employed to overcome the difficulties of testing multi-optimum solutions, λ_j is the convex coefficient; s_i^- and s_r^+ represent input and output slack variables, respectively.

$DMUo$ is DEA efficient if, and only if, the following two conditions are satisfied (i) $\varphi^* = 1$, and (ii) $s_i^{-*} = s_r^{+*} = 0, \forall i, r$, where $*$ designates an optimum.

The BCC model can be yield by incorporating an additional $\sum_{j=1}^{n} \lambda_j = 1$ into the equation (1).

（2）Benchmark Share

For a particularly inefficient DMU, the factor-specific (kth input-specific and qth output specific) measure is derived via the following two linear programming problems and the existing variable RTS model's best practice frontier.

The kth input-specific DEA model is given as

$$
\begin{aligned}
&\theta_d^{k*} = \min \theta_d^{k*}, \qquad d \in N, \\
&s.t. \sum_{j \in E} \lambda_j^d = \theta_d^k x_{kd}, \qquad k \in \{1, \ldots .m\}, \\
&\sum_{j \in E} \lambda_j^d x_{ij} \le x_{id}, \quad i \ne k, \\
&\sum_{j \in E} \lambda_j^d y_{rj} \ge y_{rd}, \quad r = 1, \ldots s, \\
&\sum_{j \in E} \lambda_j^d = 1, \\
&\lambda_j^d \ge 0, \quad j \in E,
\end{aligned} \tag{2}
$$

The qth output-specific DEA model is given as

$$
\begin{aligned}
&\phi_d^{q*} = \max \phi_d^q, \qquad d \in N, \\
&s.t. \sum_{j \in E} \lambda_j^d y_{qj} = \phi_d^q y_{qj}, \qquad k \in \{1, \ldots .m\}, \\
&\sum_{j \in E} \lambda_j^d y_{rj} \le y_{rd}, \quad i \ne k, \\
&\sum_{j \in E} \lambda_j^d x_{ij} \ge x_{id}, \quad r = 1, \ldots s, \\
&\sum_{j \in E} \lambda_j^d = 1, \\
&\lambda_j^d \ge 0, \quad j \in E,
\end{aligned} \tag{3}
$$

Where $\lambda_d^{k^*}$ and $\theta_d^{k^*}$ are optimal values in (2), $\lambda_j^{d^*}$ and $\phi_j^{q^*}$ are optimal values of (3).

In this instance, E and N respectively represent the index sets for the efficient and inefficient DMUs identified by the variable returns to scale model. The factor-specific measures in Eqs. 2 and 3 determine the maximum potential decrease of an input and increase of an output without altering other inputs and outputs at current levels. These factor-specific measures are multi-factor performance measures for all related factors are considered in a single model.

The kth input-specific benchmark-share measure for each efficient FV is measured by (4).

$$
\Delta_j^k = \frac{\sum_{d \in N} \lambda_j^{d*} (1 - \theta_d^{k*}) x_{kd}}{\sum_{d \in N} (1 - \theta_d^{k*}) x_{kd}} \tag{4}
$$

The qth output-specific benchmark-share efficient FV is calculated by (5).

$$
\prod_j^q = \frac{\sum_{d \in N} \lambda_j^{d*} (\phi_d^{q*} - 1) y_{qd}}{\sum_{d \in N} (\phi_d^{q*} - 1) y_{qd}} \tag{5}
$$

The benchmark share Δ_j^k (or $\prod_j^q$) measures the contribution of efficient units to the potential input (output) improvement in inefficient units and depends on the value of $\lambda_j^{d^*}$ and $\theta_d^{k^*}$ (or $\lambda_j^{d^*}$ and $\phi_d^{k^*}$).

The normalized weights are expressed as

$$\left[\frac{(1-\theta_d^{k*})x_{kd}}{\sum_{d\in N}(1-\theta_d^{k*})x_{kd}}\right] \text{and} \left\{\frac{\left[1-(1/\phi_d^{k*})\right]y_{qd}}{\sum_{d\in N}\left[1-(1/\phi_d^{k*})\right]y_{qd}}\right\}$$

Here, $(1-\theta_d^{k*})x_{kd}$ and $\left[1-(1/\phi_q^{d*})\right]y_{qd}$ describe the potential decrease in the kth input and increase in the qth output, respectively, and the value of $\sum_{j\in E}\Delta_j^k=1$ and $\sum_{j\in E}\prod_j^q=1$.