北京交通大学 经济管理学院

# 运输经济与物流评论

第二辑（Volume 2）




经济科学出版社
Economic Science Press

# 运输经济与物流评论

第二辑（Volume 2）

经济科学出版社

责任编辑：王长廷　袁　溦
责任校对：王肖楠
版式设计：齐　杰
技术编辑：邱　天

**图书在版编目（CIP）数据**

运输经济与物流评论．第2辑/北京交通大学经济管理学院主办．—北京：经济科学出版社，2011.9
ISBN 978-7-5141-0833-0

Ⅰ.①运…　Ⅱ.①北…　Ⅲ.①运输经济-文集②物流-文集　Ⅳ.①F50-53②F252-53

中国版本图书馆CIP数据核字（2011）第131730号

**运输经济与物流评论·第二辑**
北京交通大学经济管理学院　主办
经济科学出版社出版、发行　新华书店经销
社址：北京市海淀区阜成路甲28号　邮编：100142
总编部电话：88191217　发行部电话：88191540
网址：www.esp.com.cn
电子信箱：esp@esp.com.cn
北京密兴印刷有限公司印装
710×1000　16开　12印张　220000字
2011年9月第1版　2011年9月第1次印刷
ISBN 978-7-5141-0833-0　定价：38.00元
（图书出现印装问题，本社负责调换）

**刊　　名：**运输经济与物流评论
**顾　　问：**徐寿波　荣朝和
**主　　编：**刘延平
**执行主编：**欧国立　叶　龙



**主办单位：**北京交通大学经济管理学院
《北京交通大学学报》（社会科学版）编辑部

运输经济与物流评论
第二辑 2011 年 9 月

# 目 录

## CONTENTS

Vol. 2 September, 2011

# Location and Efficiency of Public Passenger Transport in Larger Cities: European Perspective

Olli-Pekka Hilmola*

(Lappeenranta University of Technology, Kouvola Research Unit, Prikaatintie 9, FIN -45100 Kouvda, Finland)

**[Abstract]** Four different DEA super-efficiency models are used to evaluate whether location has any role in the public short-distance passenger transportation efficiency. Our results show that location does not play any role with the entire efficiency data, but with small-scale exclusion we could build regression models having 30% -40% explanation power (using both latitude and longitude as independent factors). Models argue that best area for public transportation systems in the larger cities is located in European north-east, and correspondingly worst situation is in place in south-west. This is interesting finding, since harsh winter conditions in north, and legacy of soviet times in east are playing against of higher efficiency standards.

**[Key Words]** large cities; public passenger transport; super-efficiency; location

## 1. Introduction

Typically public passenger transport is significantly dependent on the amount of potential users in its sphere of influence (Lao & Liu, 2009; Karathodorou et al., 2010; Karttunen et al., 2010), and therefore it is not surprise that mega-cities (Jain et al., 2008) or larger entities (Odeck, 2008) have been analyzed to be the most efficient in previous benchmarking studies. Although, privatization and deregulation processes are catching up in global scale in transportation industry overall, e. g. rail based passenger transport is still

---

* 奥利佩卡·希尔莫拉（Olli-Pekka Hilmola），芬兰拉普兰塔理工大学（Lappeenranta University of Technology）教授，瑞典汇武德大学（University of Skövde）客座教授。主要研究领域：运输系统、供应链管理、工业管理、系统动力学等。

having rather minor share of private companies operating ( approx. 13% from produced Vol. based on Amos & Thompson, 2007). Thus, research findings are giving their support for more deregulated and privatized transportation systems (Jain et al. 2008; Cowie 1999), particularly in bus industry (Cowie and Asenova, 1999; Odeck, 2008). Encouraging results are also reported from former Eastern Europe, where lack of investment funding, and requirement for low fares (or even fares without a cost for special groups) is reality-even in this environment privatized bus companies e. g. in Albania have shown healthy profitability, and local government saved millions in terms of support funding (Pojani, 2010; Everis, 2010).

The main motivation behind this study is to examine, whether the efficiency of public passenger transportation in cities is driven by factor of location. As data of our study is mostly from Europe, and in latter parts data treatment leads us into deeper analysis of European situation, we are limited in one continent in our evaluation. Location studies have been classical in macro-economics from the very beginning (e. g. Ricardo, 1821; Smith, 1776), and it has been argued that some individuals, firms or countries (typically the last mentioned has been under interest of economics) hold "comparative advantage" over the others in some industrial branches (e. g. Samuelson, 1951; Krugman & Obstfeld, 1988). Nobel prize in year 2008 was devoted into this issue alone.

This manuscript is structured as follows: In Section 2 literature research is completed from the increasing need of efficient public passenger transport in larger cities. Thereafter, in Section 3 used research methodology, different data sources, and DEA performance measurement used in this study are introduced. Section 4 reports our results from two main DEA models (space and service), and their relationship with location data (latitude and/or longitude). Main results of different models are altogether discussed in Section 5, where poorest and highest performing areas of Europe are being analyzed further. In Section 6 we conclude our work, and propose further areas for the research in this topic area.

## 2. Literature Review: Short-Distance Public Passenger Transport in Larger Cities

Major concern of larger cities, but also countries within passenger transport, is the increasing popularity of private car based road transports, and the future emission scenarios arising from them (Meyer et al., 2007; Hensher, 2008). For example, if China does not bother to do nothing with this respect, then in year 2030 it will have 400 million passenger cars on roads, hungry for gasoline (Hu et al., 2009). In the city level situation is showing similar frightening growth potential, based on Kenworthy (2002) in USA per capita consumption of energy for private cars is 60000 MJ (in cities), while in China it is only 2500 MJ. Interestingly in this respect, in public transport large Asian cities (e. g. Beijing, Singapore and Hong Kong) base their short-distance passenger transport on road, namely using buses (Cameron et al., 2004; Yuanzhou et al., 2010). However, future does not look promising with regard of high income class-study from France shows that higher income results on larger use of private cars and air transport, leading eventually to considerably higher $CO_2$ emissions than in other income classes (Nicolas & David, 2009).

Only strict policies and regulation have been able to constrain this development; in Hong Kong and Singapore amount of private cars is 5 – 10 times lower compared to cities in Europe and USA (Cameron et al., 2004). But these only due to very unfavourable cost implications of owning and driving private car-similar plans also exist in Australia to reduce transport emissions of Sydney (Hensher, 2008). Among policies and regulation, careful planning of urban areas and closeness of people living besides each other as well as services, decreases throughout the world traveling by private car (Cameron et al., 2003; Karathodorou et al., 2010). Problem is that when cars become smaller, more affordable and less polluting, their total usage increases, which in the end results on higher overall emissions (Hensher, 2007; Mayer et al., 2007). For example, De Jong & de Rier (2008) argue that only hope lies in the increasing amounts of older population in "old west", and their inability to use private cars, but this potential's positive direction could easily be substituted by smaller sized households in younger age groups, and their wider use of cars.

Earlier research has shown (Lao & Liu, 2009; Karttunen et al., 2010) that inside city there exist some routes, which are extremely popular, and even could be profitable to operate. However, this is not complete solution and package, since numerous other routes need to be operated too in order to satisfy transportation needs, and their popularity and profitability could be questionable. Efficiency is not the ultimate answer here, in other words efficiency of actor level (or route level, like George & Rangaraj, 2008 illustrated); most interesting is how efficient public passenger transportation systems are from system's perspective (city level). This efficiency concerns both objectives: Public passenger transportation system services used, and space needed to build this system to operate. These objectives could have tradeoffs as well, where favouring efficiency of another one will have decreasing efficiency in other respect. However, problem with coverage of public transportation system, and structure of large cities is constant challenge, and especially for old cities, which have considerably being enlarged with sub-urban areas within last six decades time period (Pojani, 2010; Nicolas & David, 2009).

## 3. Research Methodology and Used Data

International Association of Public Transport (UITP, 2005) maintains database concerning public transportation in larger cities. Latest version is from year 2001 and contains 52 cities around the world. However, most of these are from Europe (47 of total). This database has been used in earlier scientific studies (e. g. Kenworthy, 2002; Cameron et al., 2004; Karathodorou et al., 2010; Albalate & Bel, 2010), and could be assumed to have needed reliability of indicators gathered. From most recent studies Karathodorou et al. (2010), and Albalate & Bel (2010) both used UITP database as main source within their large-scale statistical analysis, not only describing the large cities public transport environment, which has been the case in the earlier studies (Kenworthy, 2002; Cameron et al., 2004).

In Figures 1 and 2 is presented two DEA efficiency measurement models of this research work; one concerning space utilization efficiency and second one service usage efficiency. We altered these models in a manner that amounts of inputs were either two (population and urban population density) or all five. In all of the following efficiency measurement models outputs are the same in each case (space and service used per se), four of these outputs are rail related and one concerns busses. Of course walking and biking should be part of public passenger transportation system, but we were forced to leave them out of measurement due to lack of data.

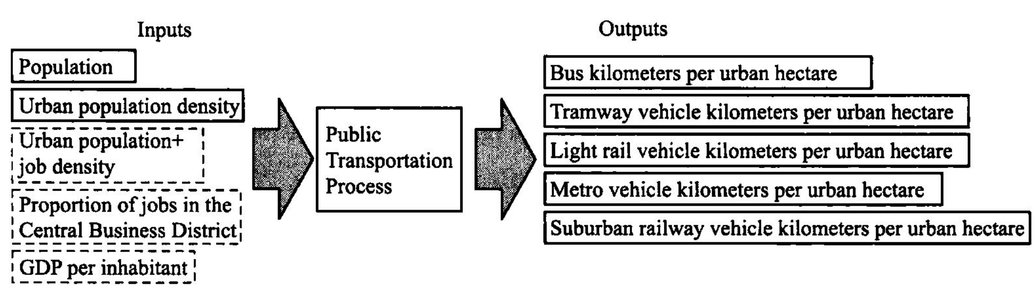


**Figure 1 DEA measurement model of public passenger transport concerning super-efficiency of space being used**

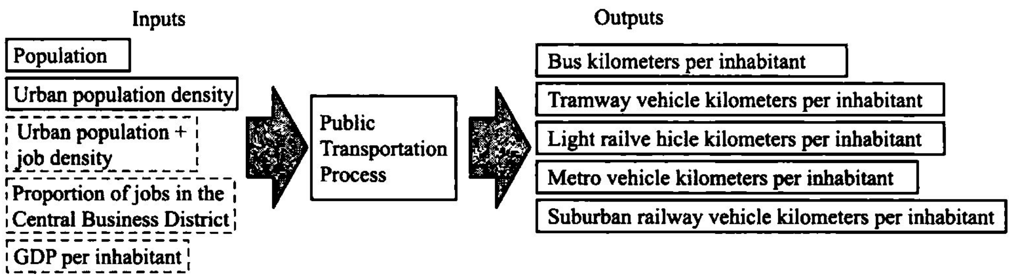


**Figure 2 DEA measurement model of public passenger transport concerning super-efficiency of transportation services used per inhabitant**

As in larger efficiency measurement models five inputs and five outputs were being used, some cities were not included in the efficiency analysis. Mostly reason for excluding was the lack of data concerning inputs of "urban population plus job density" and "proportion of jobs in the central business district". Therefore, our analysis in the larger models consist 43 cities. However, in smaller models all of the database cities were involved.

For latitude and longitude information, we used Google Maps (2010), and its geographical coordinate feature. This information was matched with UITP (2005) railway database, and each large city was thereafter having its latitude and longitude data. For the purposes of the following analysis, we have scaled latitude in a manner that countries below equator (south) are having negative latitude values and above (north) positive ones. Similarly longitude information is scaled in a manner that countries belonging to west longitude (from Greenwich) are having negative values as east longitude countries are marked as positive ones. Therefore, in theory our scale in latitude is ranging from −90° (South Pole) to +90° (North Pole), and in longitude from −180° to +180°.

Typically DEA efficiency measurement technique is used to fit scale curves (constant or variable) on given input-output models. Initial introduction to the idea was made during 1950's (Farrell, 1957), while its real down to earth linear programming application was developed during late 1970's (Charnes et al., 1978). In the beginning constant returns on scale were assumed, and fitted on data. Thereafter, other scale curves, like variable returns on scale (favouring smaller decision making units) were applied (Banker et al., 1984). Problematic part in DEA models is that its measurement scale is fixed, and efficiency values could range from 0 to 100% (latter number called as frontier; Barros et al., 2009; Odeck, 2009; Mavi et al., 2010). As one option to avoid erroneous interpretations in using regression models with DEA results, has been to calculate super-efficiencies (e.g. Nahra et al., 2009). As we apply in the following super-efficiency models, scale of efficiency measurement is not having maximum value, and it could be basically anything from 100% onwards, giving us justification to use ordinary regression analysis to make interpretation from possible causalities.

# 4. Does Location Drive Super-Efficiency of Short-Distance Public Passenger Transports in Large Cities?

## 4.1 Location and Small DEA Super-Efficiency Measurement Model

Surprising finding was that small DEA super-efficiency measurement model results were not driven by location (see Table 1 in below). Latitude (model 1) in space and services used model was slightly better than longitude (model 2) in explaining efficiency performance, but explanation power of 1% –2% does not give any indication even nearly statistically significant relationship. As both location factors are taken into account in model 3, explanation power improves a bit, but is still around 4%, and not even near of statistical significance.

**Table 1 Linear regression model fits in small DEA super-efficiency evaluation**

| Services used | $R^2$ | Intercept (p-value) | Latitude (p-value) | Longitude (p-value) |
|---|---|---|---|---|
| Model 1 | 0.027 | 0.432<br>(0.174) | 0.008<br>(0.244) | |
| Model 2 | 0.001 | 0.762<br>(0.000) | | 0.001<br>(0.813) |
| Model 3 | 0.041 | 0.268<br>(0.471) | 0.011<br>(0.160) | 0.003<br>(0.402) |
| Model 4 | 0.088 | -1.042<br>(0.372) | 0.039<br>(0.101) | 0.012<br>(0.407) |
| Model 5 | 0.312 | -1.404<br>(0.012) | 0.043<br>(0.000) | 0.006<br>(0.450) |
| Model 5<br>(Without intercept) | 0.243 | | 0.013<br>(0.000) | 0.010<br>(0.194) |

| Space used | $R^2$ | Intercept (p-value) | Latitude (p-value) | Longitude (p-value) |
|---|---|---|---|---|
| Model 1 | 0.012 | 0.581<br>(0.064) | 0.005<br>(0.435) | |
| Model 2 | 0.008 | 0.776<br>(0.000) | | 0.002<br>(0.521) |
| Model 3 | 0.037 | 0.372<br>(0.306) | 0.009<br>(0.236) | 0.004<br>(0.272) |
| Model 4 | 0.018 | -0.026<br>(0.982) | 0.017<br>(0.462) | 0.005<br>(0.739) |
| Model 5 | 0.296 | -0.344<br>(0.350) | 0.020<br>(0.014) | 0.011<br>(0.029) |
| Model 5<br>(without intercept) | 0.285 | | 0.013<br>(0.000) | 0.012<br>(0.014) |

Denotation: model 1 is relation with latitude, model 2 is relation with longitude, model 3 is relation with latitude and longitude, and model 4 is relation with latitude and longitude when top 15 population cities are excluded, and model 5 is relation with latitude and longitude when in all models non-European cities, and Bern are excluded.

Even though these results were not convincing, we were able to develop linear regression model having high explanation power, and also statistical significance. In smaller super-efficiency models it is exclusion model, where non-European cities are not included, but also very high performing Bern is taken away (named as model 5). Explanation power in this case is nearly 30%, and in services used model latitude is having statistical significance, while in space used both latitude and longitude do so. We also calculated linear regression models without intercept (as latter was not statistically significant), and in both cases statistical significance performance improved for latitude and longitude (in services used latter one was still out of acceptable limits). However, we do want to emphasize that exclusion in large-scale alone is not sufficient to improve explanation power of used regression models-model 4 shows situation, where TOP15 cities (by population) were excluded from the analysis, and it very marginally improves the situation in services used model as compared to model 3, but in space used decreases below 2%.

As minor exclusion seems to provide better explanation power, we did explore its effect on other models (model 3 in Table 2 is model 5 in Table 1). As could be noted from Table 2, even single factor co-ordinate models are able to predict at best case super-efficiency performance by 20% -30%. This again in the models, where latitude is explaining the performance; longitude is having non-significant situation still in the services used model. If exclusion is continued further by not only incorporating non-European cities and Bern, but also excluding TOP15 by population, we do reach in services used model explanation power of nearly 40%.

**Table 2　Linear regression model fits in small DEA super-efficiency evaluation, when exclusion is effective in all models**

| Services used | $R^2$ | Intercept (p-value) | Latitude (p-value) | Longitude (p-value) |
|---|---|---|---|---|
| Model 1 | 0. 302 | -1. 50 (0. 006) | 0. 05 (0. 000) | |
| Model 2 | 0. 076 | 0. 588 (0. 000) | | 0. 015 (0. 067) |
| Model 3 | 0. 312 | -1. 404 (0. 012) | 0. 043 (0. 000) | 0. 006 (0. 450) |
| Model 4 | 0. 395 | -1. 478 (0. 008) | 0. 044 (0. 000) | 0. 015 (0. 028) |

| Space used | $R^2$ | Intercept (p-value) | Latitude (p-value) | Longitude (p-value) |
|---|---|---|---|---|
| Model 1 | 0. 210 | -0. 533 (0. 157) | 0. 026 (0. 002) | |
| Model 1 (without intercept) | 0. 210 | | 0. 015 (0. 000) | |
| Model 2 | 0. 184 | 0. 581 (0. 000) | | 0. 016 (0. 003) |
| Model 3 | 0. 296 | -0. 344 (0. 350) | 0. 020 (0. 014) | 0. 011 (0. 029) |
| Model 4 | 0. 241 | -0. 498 (0. 220) | 0. 023 (0. 006) | 0. 008 (0. 112) |
| Model 4 (without intercept) | 0. 236 | | 0. 014 (0. 000) | 0. 006 (0. 193) |

Denotation: model 1 is relation with latitude, model 2 is relation with longitude, model 3 is relation with latitude and longitude, and model 4 is relation with latitude and longitude when top 15 population cities are excluded (in all models non-European cities, and Bern is excluded)

So, based on our regression analysis of super-efficiency models of small DEA of public passenger transport, we can argue that overall they do not have any statistically significant relationship with location. However, if some parts of observations are being excluded, we can find statistically significant models, and these typically explain 30% from efficiency performance. Frequently we do find that latitude is the mostly explaining the efficiency, and some cases it could have longitude besides of it, giving a bit better explanation power.

## 4. 2 Location and Large DEA Super-Efficiency Measurement Model

In base analysis large DEA super-efficiency models do not bring any further change on the results of smaller DEA analysis. As Table 3 highlights, all data incorporating models from one to three have very low explanation power (ranging from 0 to 9%), except of model 3 for space used DEA. This last mentioned is showing some potential significance from longitude, which is contrarian finding on previous analysis.

**Table 3 Linear regression model fits in large DEA super-efficiency evaluation**

| Services used | $R^2$ | Intercept (p-value) | Latitude (p-value) | Longitude (p-value) |
|---|---|---|---|---|
| Model 1 | 0.003 | 1.142<br>(0.008) | 0.003<br>(0.748) | |
| Model 2 | 0.023 | 1.202<br>(0.000) | | 0.005<br>(0.333) |
| Model 3 | 0.041 | 0.812<br>(0.102) | 0.009<br>(0.392) | 0.007<br>(0.214) |
| Model 4 | 0.088 | -1.071<br>(0.492) | 0.045<br>(0.147) | 0.017<br>(0.362) |
| Model 5 | 0.372 | -1.176<br>(0.109) | 0.042<br>(0.008) | 0.025<br>(0.021) |
| Model 5<br>(without intercept) | 0.33 | | 0.018<br>(0.000) | 0.028<br>(0.01) |

| Space used | $R^2$ | Intercept (p-value) | Latitude (p-value) | Longitude (p-value) |
|---|---|---|---|---|
| Model 1 | 0.003 | 1.048<br>(0.209) | 0.007<br>(0.711) | |
| Model 2 | 0.088 | 1.088<br>(0.003) | | 0.018<br>(0.053) |
| Model 3 | 0.135 | -0.171<br>(0.853) | 0.027<br>(0.147) | 0.025<br>(0.018) |
| Model 4 | 0.008 | 0.353<br>(0.829) | 0.013<br>(0.682) | 0.005<br>(0.771) |
| Model 5 | 0.222 | 0.069<br>(0.908) | 0.012<br>(0.328) | 0.021<br>(0.019) |
| Model 5<br>(without intercept) | 0.222 | | 0.014<br>(0.000) | 0.021<br>(0.017) |

Denotation: model 1 is relation with latitude, model 2 is relation with longitude, model 3 is relation with latitude and longitude, and model 4 is relation with latitude and longitude when top 15 population cities are excluded, and model 5 is relation with latitude and longitude when in non-European cities, and Bern and Moscow are excluded.

Exclusion in large DEA models yields results, where model 4 (non-Europeans plus Bern and Moscow excluded) shows strikingly high explanation power (ranging between 20% -40%). Highest performance recorded in the model of services used, where regression model is having intercept (which is not statistically significant). Interestingly in this exclusion case we do not face any problems to find statistical significance for longitude (which was weak parameter in small DEA), and both of the location parameters have needed high statistical significance, if without intercept models are also taken into account. Word of caution over in exclusion approach-model 4, where all TOP15 (by population) cities are excluded from the regression analysis, we face very low and insignificant explanation power still.

As promising path of excluding non-European cities and Bern as well as Moscow showed promising opportunities for location significance in large DEA, we performed further analysis with this regard shown in Table 4. This action improved situation especially with services used DEA model, where latitude and longitude alone were able to give more than 20% explanation power (models 1 and 2). Similarly longitude (model 2) in space used model nearly reached 20% level. Latitude performed somewhat poorly in space used model (only 7% explanation power), but was still statistically significant. As regression model and data in model 3 in Table 4 is the same as model 5 of Table 3, and we can only conclude that it is still having highest explanation power, even if the variety of exclusion models is taking into account (in Table 4).

**Table 4 Linear regression model fits in large DEA super-efficiency evaluation, when exclusion is effective in all models**

| Services used | $R^2$ | Intercept (p-value) | Latitude (p-value) | Longitude (p-value) |
|---|---|---|---|---|
| Model 1 | 0.258 | -1.487<br>(0.057) | 0.052<br>(0.002) | |
| Model 1<br>(without intercept) | 0.258 | | 0.022<br>(0.000) | |
| Model 2 | 0.217 | 0.808<br>(0.000) | | 0.033<br>(0.005) |
| Model 3 | 0.372 | -1.177<br>(0.109) | 0.042<br>(0.008) | 0.025<br>(0.021) |
| Model 4 | 0.345 | -1.673<br>(0.048) | 0.053<br>(0.002) | 0.021<br>(0.035) |

| Space used | $R^2$ | Intercept (p-value) | Latitude (p-value) | Longitude (p-value) |
|---|---|---|---|---|
| Model 1 | 0.074 | -0.193<br>(0.759) | 0.020<br>(0.114) | |
| Model 1<br>(without intercept) | 0.074 | | 0.017<br>(0.000) | |
| Model 2 | 0.198 | 0.650<br>(0.000) | | 0.024<br>(0.007) |
| Model 3 | 0.222 | 0.070<br>(0.908) | 0.012<br>(0.328) | 0.021<br>(0.019) |
| Model 4 | 0.122 | -0.318<br>(0.646) | 0.022<br>(0.110) | 0.010<br>(0.205) |
| Model 4<br>(without intercept) | 0.12 | | 0.016<br>(0.000) | 0.009<br>(0.228) |

Denotation: model 1 is relation with latitude, model 2 is relation with longitude, model 3 is relation with latitude and longitude, and model 4 is relation with latitude and longitude when top 15 population cities are excluded. (in all models non-European cities, and in small models Bern, while in large models Bern and Moscow are excluded)

Interestingly, in large DEA model TOP15 (by population) cities exclusion yields also good results on regression analyses (also non-European cities and Bern as well as Moscow excluded). This is especially the case in services used model (where regression model is able to explain nearly with the same precision with model 3). In space used model this large-scale exclusion does not yield this level of impressive results, but still in situation where regression model is not having intercept, latitude is highly significant in statistical terms.

In large DEA regression analysis we found that longitude is not problematic location parameter, even if it did so in small DEA. However, most of the results are very similar to earlier sub-section findings. So, we could argue still that some exclusion for super-efficiency DEAs is needed, and then both longitude and latitude are needed jointly in the models to explain efficiency formation.

# 5. Discussion

As our findings and regression models yielded significance only in European context, and in a case that Bern is excluded, it is worthwhile to discuss and analyze further performance of north-east and south-west pubic passenger transportation performance (please see Tables 5 and 6 in the below). However, we would like to emphasize that two European cities, Bern and Moscow (former excluded in both small and large, while latter in only large), have exceptional performance in the exclusion cases. For example, Bern sets frontier with approx. 600% efficiency in smaller model, and in larger model performance remains in the level. Moscow in larger model dominates space used model with enormous efficiency level of just above 1 200%, and in services used it is just below 500%. So, these exceptional cities should be taken into account as we analyze difference in performance of north-east and south-west of Europe.

**Table 5 Five example larger cities from Europe concerning north-east (left-side) and south-west (right-side) classification, and their respective super-efficiency performance in two small DEA models**

| City | Space DEA | Services DEA | City | Space DEA | Services DEA |
|---|---|---|---|---|---|
| Krakow | 135.1% | 97.5% | Lisbon | 40.0% | 46.8% |
| Warsaw | 94.8% | 81.6% | Sevilla | 17.3% | 22.2% |
| Tallinn | 99.1% | 121.0% | Madrid | 53.3% | 22.6% |
| Helsinki | 75.3% | 109.0% | Bilbao | 36.7% | 37.7% |
| Moscow | 113.9% | 19.2% | Valencia | 24.9% | 22.5% |
| Average | 103.7% | 85.7% | Average | 34.4% | 30.4% |