Non-Parametric Performance Evaluation of Container Ports

By:

Susila Munisamy
and
Wang Danxia

Paper presented at the ***XII European Workshop on Efficiency and Productivity Analysis Conference*** held on 22-24 June 2011 in Verona, Italy

Non-parametric Performance Evaluation of Container Ports

Susila Munisamy
University of Malaya, Kuala Lumpur, Malaysia
susila@um.edu.my

Wang Danxia
University of Malaya, Kuala Lumpur, Malaysia
danxiawong@alumni.nus.edu.sg

Abstract

This paper analyses the performance of major container ports in Asia using the non-parametric approach. Because of the known sensitivity of the traditional non-parametric approaches such as Data Envelopment Analysis (DEA) and Free Disposal Hull (FDH) to extreme values or outliers, small sample size and slow convergence rates of efficiency estimates, this study employs an alternative robust non-parametric technique called order-m estimator following Cazals et al. (2002) to evaluate the efficiency of container ports. In this study, we conduct two analyses comprising order-m and Free Disposal Hull (FDH), and rank the container ports with the order-m efficiency estimates. The estimates are robust with respect to outliers in the data.

Keywords

Robust estimates, nonparametric, efficiency analysis, container port

1. Introduction

Since the 1980s, globalization expedited the growth of international trade. The spin-off of increased exports and imports has created highly differentiated demands in goods, which made cargo unitization, or containerization, increasingly popular. Such innovation is followed by intermodal freight transport systems which initiated a new market of global networking and freight transportation chains. Wang et al. (2005) observed that port services providers are obliged to adopt the most sophisticated practices, ranging from acquiring higher specifications of quayside cranes, yard gantries, to yard space optimization and automation systems, to maintain efficiency and competitiveness. Hence, container ports have evolved to become highly capital intensive in the race of acquiring new technology. Furthermore, port authorities have frequently come under pressure to improve competitiveness due ports' macroeconomic significance. As the competition amongst container ports evolve at the expense of capital intensity, indicators of plausible improvements in port management and operational planning becomes significantly important. As such, the efficiency benchmarking process against competing ports provides port operators with *ex ante* advantage in investment and planning decisions.

The theory of efficiency analysis began with the works of Koopmans (1951), Debreu (1951) and Farrell (1957) who made the first attempts at estimating efficiencies for a set of observed production units. Among the non-parametric approach for estimating efficiency, the Free Disposal Hull (FDH), and Data Envelopment Analysis (DEA) are based on the envelopment technique. The idea of free disposability of the production set was developed by Deprins et al. (1984), which led to the Free Disposal Hull (FDH) estimator. Other convexity assumptions made on the production possibility set produced the Data Envelopment estimator, which had been developed by Charnes et al. (1978), who imposed constant returns to scale using the convex cone; and subsequently Banker et al. (1984) introduced the convex hull on the production possibility frontier to enable variable returns to scale assumption. Although the FDH estimate is proven by Korostelev et al. (1995a, 1995b) to be the most efficient estimator, the estimate is highly susceptible to extreme values, or outliers (Simar and Wilson (2000); Kneip, Simar and Wilson (2008)) which can strongly influence the non-parametric estimation of efficiency. In a multivariate input and output case, the convergence in probability of the FDH estimator is shown to be slower than the root-*n* rate of the standard parametric estimate benchmark. Therefore, as the number of inputs and outputs increase, the rate of convergence reduces causing the *curse of dimensionality*. The more recent work of Cazals et al. (2002) introduced the order-m estimator which is related to the FDH estimator but can address the issues of convergence rate, and susceptibility of the estimates to extreme points. In addition, for large number of inputs and outputs, the order-m estimator requires far less data to provide a more efficient estimate than FDH (Wheelock and Wilson, 2003).

The main non-parametric methods used in previous empirical studies to estimate efficiency of ports are DEA and FDH. In this paper, we employ the robust alternative to the standard DEA/FDH estimators i.e. the order-m technique proposed by Cazals et al. (2002). This technique converges in root-*n* consistency, achieves asymptotic normality, and is more robust to outliers and extreme points since it does not envelop all units. We formally investigate the efficiency performance of 71 major container ports in the Asian region for the year 2007. We conduct two analyses comprising FDH and order-m and compare the results. We identify outliers following (Simar, 2003). Subsequently we rank the container ports with the order-m estimates which are robust with respect to outliers in the data. The paper is organized into five sections. Section 2 describes the main framework of our approach and methodology. In Section 3 we describe the container port operations and the data, while we discuss the results in Section 4. Section 5 concludes.

2. Methodology

In a framework based on the theory of the firm, a container port production can be formulated as a set of *n* ports of $S_n = \{(x_i,y_i)\}_{i=1}^{n}$ consisting of $x \in R_{+}^{p}$ inputs used to produce $y \in R_{+}^{q}$ outputs, which makes a production possibility set of Ψ:

$$\Psi = \{(x,y) \in \mathbb{R}_{+}^{p+q} \mid x \text{ can produce } y\}. \tag{1}$$

Based on this production possibility definition, the Farrell measure of output-orientated efficiency for a port operating at (x,y) is defined as:

$$\lambda(x,y) = sup\{\lambda \mid (x, \lambda y) \in \Psi\} \tag{2}$$

where $\lambda(x,y) \geq 1$ represents the proportionate expansion in outputs without altering inputs. In this framework, $\lambda(x,y) = 1$ indicates the container port is on the efficient frontier which serves as the yardstick to benchmark other ports.

A common estimator of Ψ is the free disposal hull FDH of S_n suggested by Deprins et al. (1984). The Ψ_{FDH} assumes free disposability. The free disposal of Ψ is defined as follows:

$$\hat{\Psi}_{FDH} = \{(x,y) \in \mathbb{R}_{+}^{p+q} \mid y \leq y_i,\ x \geq x_i,\ i = 1, \ldots, n\}. \tag{3}$$

Subsequently, the FDH estimates are arrived as below:

$$\hat{\lambda}(x,y)_{FDH} = sup\{\lambda \mid (x, \lambda y) \in \hat{\Psi}_{FDH}\}. \tag{4}$$

The FDH estimator can cause problems in efficiency performance analysis due to the slow convergence rates, curse of dimensionality, and extreme sensitivity to outliers and small sample size. For these matters, the order-*m* estimator is more robust and overcome the drawback of the FDH estimator.

2.1 *Order-m Efficiency Estimate*

As an alternative to traditional non-parametric estimators, Cazals et al. (2002) proposed an estimator based on the expected maximum output (or minimum input) frontier of order-*m*. This estimator is robust to outliers while converging in probability at the standard parametric estimation rate of root-n, by characterizing stochasticity through probabilistic framework in the production possibility set. In this non-traditional framework, the production process is described by the joint probability measure of (X,Y) on $\mathbb{R}_{+}^{p} \times \mathbb{R}_{+}^{q}$:

$$H_{XY}(x,y) = Prob(Y \geq y \mid X \leq x) Prob(X \leq x). \tag{5}$$

where $S_Y(y|x) = Prob(Y \geq y \mid X \leq x)$ is the conditional distribution on Y, arrived using the Bayes' rule. Hence, from here, we can say that the attainable set of Ψ and the Farrel output efficiency estimate with the support of (X,Y) under the free disposability assumption is given as follows:

$$\lambda(x, y) = sup\{\lambda \mid S_Y(\lambda y \mid x)\}. \tag{6}$$

Empirically, $S_Y(y \mid x)$ can be arrived by the following formulation:

$$\hat{S}_Y(y \mid x) = \frac{\sum_{i=1}^{n} \mathbb{I}(X_i \leq x, Y_i \geq y)}{\sum_{i=1}^{n} \mathbb{I}(X_i \leq x)}, \tag{7}$$

where $\pi(\cdot)$ is the indicator function. Hence, the FDH output efficiency estimate in Equation (5) can be empirically expressed as below:

$$\hat{\lambda}_n(x, y) = sup\{\lambda \mid \hat{S}_Y(\lambda y \mid x)\}. \tag{8}$$

The order-*m* output efficiency is defined as follows. Given an input level x in the interior of the support of X, consider *m* i.i.d random variables Y_i, i = 1,… ,m generated from the conditional *q*-variate distribution function $S_Y(y \mid x) = \text{Prob}(Y \leq y \mid X \leq x)$ and define the production possibility set:

$$\Psi_m(x) = \{(x', y) \in \mathbb{R}_+^{p+q} \mid x' \leq x, Y_i \leq y, i = 1, \ldots, m\}. \tag{9}$$

Subsequently, for any *y*, we define

$$\begin{aligned}\tilde{\lambda}_m(x, y) &= sup\{\lambda \mid (x, \lambda y) \in \Psi_m(x)\} \\ &= \max_{i=1,\ldots,m} \left\{ \min_{j=1,\ldots,q} \left(\frac{Y_i^j}{y^j} \right) \right\}\end{aligned} \tag{10}$$

The (expected) order-m output efficiency, is defined for all *x* in the interior of the support of *X* as follows:

$$\lambda_m(x, y) \equiv E(\tilde{\lambda}_m(x, y) \mid X \leq x), \tag{11}$$

where the expectation is assumed to exist. This is a less extreme, and non-traditional benchmark from which the efficiency scores is the expectation of the maximal output that is technically feasible with the given inputs of the unit (x,y) when compared to any *m* peers who are potential competitors, randomly drawn from the population of units utilizing input levels of less than *x*. In this study, 4,000 samples with size *m* are drawn with replacement. The 4,000-time Monte Carlo replication is done to ensure a stable mean of the expected maximum output level of order-m. The order-m efficiency estimate can then be computed as

$$\lambda_m(x, y) = \int_0^\infty [1 - (1 - S_Y(uy \mid x))^m] du = \lambda(x, y) - \int_0^{\lambda(x,y)} (1 - S_Y(uy \mid x))^m du. \tag{12}$$

The non-parametric estimator of $\lambda_m(x,y)$ is obtained by substituting the empirical $S_{Y|X,n}(y \mid x)$ into the above equation:

$$\hat{\lambda}_m(x,y) = \int_0^{\infty} [1-(1-\hat{S}_Y(uy \mid x))^m]du = \hat{\lambda}(x,y) - \int_0^{\hat{\lambda}(x,y)} (1-\hat{S}_Y(uy \mid x))^m du. \tag{13}$$

For a given sample, the order-*m* estimate converges to the FDH estimate as $m \to \infty$, i.e. $\lim_{m\to\infty} \hat{\lambda}_m(x,y) = \hat{\lambda}_{FDH}(x,y)$. But for a finite *m*, the order-*m* estimator creates a partial frontier that does not envelop all observed data points as the FDH estimate does, thus, making it robust to extreme values or outliers. The method results in scores greater or less than unity since the unit analyzed is not necessarily part of the order-*m* sample nor will there necessarily be any other units dominating the unit analyzed in the output.

3. Container Port Operations and Data

Vis and de Koster (2003) defines a container port as a transshipment gateway – where containers are transferred from ships to barges, trucks and trains, and vice versa. After an import container is taken off from the cargo ship using quayside cranes, the container is transferred using straddle carriers or vehicles such as prime movers or trailers to the stack. After arrangements by parties at the receiving end, the container at the stack can be transported by the vehicles to other transportation modes like barges, deep sea ships, trains or are merely collected by trucks. An export container encounters a reverse of this process. Such is the typical operation of a container port with throughput greater than 100,000 TEUs a year. On the other hand, at a container port with throughput less than 100,000 TEUs throughput, mobile cranes may be used instead of quayside cranes, reachstackers instead of straddle carriers. There may be no use of gantry cranes for transport of containers within the stack, while more common uses of utility trailers and prime movers. To model the port operations, we capture the main resources used by the ports (inputs) for acquiring the main goods and services produced (outputs). Under the traditional microeconomic framework, capital and labour are necessarily the input for production. A common issue in the empirical studies of container port efficiency performance is finding a proxy to reflect labour or the number of workers. According to Notteboom et al. (2000), expert analysis shows that there is a stable relationship between the number of yard gantries with the number of dock workers. Wang et al. (2005) goes to show that the average number of workers per crane is six. Hence, we take the total yard equipments, i.e. sum of straddle carriers, yard gantries, reachstackers, front-end handlers, and forklifts, as an input factor, reflecting the labour that is required. We enlist another four inputs encompassing berth length, terminal area, total reefer points, and total quayside cranes (and/or mobile cranes) – to reflect the capital inputs in the industry. The single output used is the total throughput of the container port. We obtain secondary data from Containerisation International Yearbook 2007. The analyses in the next section are performed using routines in the FEAR v1.12 package (see

Wilson, 2008). The descriptive statistics for variables used to calculate the order-m and FDH efficiency estimates for the 71 major container ports in Asia are as shown in Table 1 below:

Table 1: Major Asian Container Ports Year 2007: Descriptive Statistics for Inputs and Outputs

	Inputs					Output
	Berth Length (m)	Terminal Area (m^2)	Total Reefer Points	Total Quayside Cranes	Total Yard Equipments	Total Throughput (TEU)
Min	100	3200	0	1	4	20,700
Max	12,610	6,169,837	7,422	131	674	27,935,500
Mean	2,525.90	892,461	1,130.28	19.23	88.80	3,239,158.84
St.Dev.	2,648.60	1,194,399.88	1,795.83	24.51	119.51	5,690,010.84

4. Empirical Results

Table 2 shows the results of the output-oriented order-m and FDH efficiency estimates for 71 container ports in Asia in year 2007. The rightmost column gives the FDH estimates ($\hat{\lambda}_{FDH}$). The FDH estimates indicate 30 of the 71 ports (42 percent) lie on the frontier, with a score of 1, reflecting the curse of dimensionality. The third to seventh columns give the order-m estimates ($\hat{\lambda}_m$) for all ports for 5 values of *m* (25, 50, 75, 100, 150), each replicated under a Monte-Carlo simulation of 4,000 times to produce stable estimates. Recall that, the order-m estimates for any given port is derived by comparing the actual output of that port with the expected maximum output of the port, where the expected maximum output is obtained by drawing 4,000 samples of m ports that use no more than the inputs of the given port. In the output orientation, a value less than 1 indicates that the port produces more than the expected maximum, whereas a value larger than 1 indicates the port produces less than the expected maximum. Generally, in the output orientation, the order-*m* estimates are larger for larger values of m. The table shows some $\hat{\lambda}$,m values are less than 1, which indicates that the container port unit is situated outside the order-m frontier i.e. with input values smaller then the given port. Shanghai (11) for instance, has a $\hat{\lambda}_{25} = 0.3655$, which means that Shanghai produces 2.7 times more cargo throughput than the expected value of the maximal level of output of 25 other container ports drawn from the population of firms using less than its inputs. Thus, Shanghai is a potential super-output-efficient outlier. Looking at the geometric averages at the bottom of the table, as expected, it is observed that for an increasing *m*, the estimates approximate the FDH estimate. We also observed that as the value of *m* increases, the percentage of points that lie outside of the order-m frontier reduces, as presented in Table 2 and Figure 1.

Table 2: Output oriented order-m and FDH efficiency estimates of major container ports in Asia

DMU	Port	λ_{25}	λ_{50}	λ_{75}	λ_{100}	λ_{150}	λ_{FDH}
1	Chittagong	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	Muara	6.0892	6.1112	6.1115	6.1115	6.1115	6.1115
3	Sihanoukville	2.3720	2.3761	2.3761	2.3761	2.3761	2.3761
4	Dalian	1.4836	1.8896	2.0224	2.0589	2.0774	2.0824
5	Fuzhou	0.9048	0.9855	0.9954	0.9991	0.9999	1.0000
6	Guangzhou	0.5592	0.8436	0.9685	1.0049	1.0154	1.0174
7	Hong Kong	0.5450	0.8023	0.9484	1.0404	1.1252	1.1641
8	Lianyungang	0.8878	0.9954	1.0000	1.0000	1.0000	1.0000
9	Ningbo	0.3131	0.6111	0.8088	0.8977	0.9826	1.0000
10	Qingdao	0.7406	0.9220	0.9762	0.9925	0.9987	1.0000
11	Shanghai	0.3655	0.5381	0.6635	0.7490	0.8818	1.0000
12	Shantou	1.2688	1.2752	1.2752	1.2752	1.2752	1.2752
13	Shekou	1.3952	2.0196	2.3606	2.5098	2.7004	2.7773
14	Tianjin	0.4924	0.8011	0.9222	0.9677	0.9933	1.0000
15	Xiamen	0.4127	0.7066	0.8980	0.9691	0.9941	1.0000
16	Yantai	0.6885	0.9201	0.9757	0.9948	0.9997	1.0000
17	Yantian	0.6216	0.8828	0.9650	0.9912	0.9987	1.0000
18	Zhangjiagang	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
19	Chennai	0.9598	0.9957	0.9996	0.9999	1.0000	1.0000
20	Jawaharlal Nehru	0.8039	1.0385	1.1027	1.1273	1.1366	1.1397
21	Kochi	2.4528	2.4770	2.4777	2.4777	2.4777	2.4777
22	Mumbai	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
23	Mundra	0.9915	0.9995	0.9999	1.0000	1.0000	1.0000
24	New Mangalore	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
25	Pipavav	5.7906	5.8009	5.8009	5.8009	5.8009	5.8009
26	Tuticorin	1.3335	1.3362	1.3362	1.3362	1.3362	1.3362
27	Visakhapatnam	10.3659	10.3659	10.3659	10.3659	10.3659	10.3659
28	Belawan	1.0351	1.0351	1.0351	1.0351	1.0351	1.0351
29	Tg Priok	0.8136	1.0981	1.1941	1.2261	1.2487	1.2540
30	Hakata	1.0703	1.1139	1.1195	1.1208	1.1212	1.1212
31	Kawasaki	13.9042	13.9042	13.9042	13.9042	13.9042	13.9042
32	Kitakyushu	0.9876	0.9990	0.9998	1.0000	1.0000	1.0000
33	Kobe	3.2277	3.8530	4.0271	4.0791	4.1223	4.1478
34	Mitajiri	13.4514	13.4514	13.4514	13.4514	13.4514	13.4514
35	Mizushima	2.8624	2.8624	2.8624	2.8624	2.8624	2.8624
36	Nagoya	1.4478	2.2462	2.6617	2.9554	3.1337	3.2318
37	Niigata	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
38	Osaka	2.6003	3.6294	3.9541	4.0474	4.1133	4.1237
39	Shimizu	1.3877	1.3964	1.3968	1.3969	1.3969	1.3969
40	Shimonoseki	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
41	Tokyo	1.4645	1.8871	2.0612	2.1695	2.2331	2.2697
42	Yokkaichi	3.5988	3.6254	3.6261	3.6262	3.6262	3.6262
43	Yokohama	1.9619	2.4457	2.6381	2.7254	2.7704	2.7785
44	Bintulu	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
45	Kuantan	4.7092	4.7160	4.7163	4.7163	4.7163	4.7163
46	Pasir Gudang	0.9415	0.9933	0.9991	1.0000	1.0000	1.0000
47	Penang	1.5622	1.9151	2.0551	2.1229	2.1566	2.1609
48	Port Klang	0.9003	1.1358	1.2174	1.2565	1.2927	1.3148
49	Tg Pelepas	0.8412	1.3067	1.5057	1.5959	1.6709	1.7018
50	Karachi	0.9133	0.9792	0.9935	0.9976	0.9997	1.0000
51	Port Mohammad bin Qasim	0.9829	0.9987	0.9999	0.9999	1.0000	1.0000

52	Cebu	1.1930	1.1932	1.1932	1.1932	1.1932	1.1932
53	Davao	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
54	General Santos	5.8102	5.8102	5.8102	5.8102	5.8102	5.8102
55	Iloilo	1.8983	1.8992	1.8992	1.8992	1.8992	1.8992
56	Manila	2.3543	3.0407	3.2170	3.2846	3.3146	3.3195
57	Subic Bay	16.7287	16.7287	16.7287	16.7287	16.7287	16.728
58	Zamboanga	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
59	Jurong	1.1534	1.2313	1.2520	1.2604	1.2650	1.2656
60	PSA International Ltd	0.3414	0.5163	0.6428	0.7363	0.8704	1.0000
61	Busan	0.7240	0.8558	0.9037	0.9380	0.9688	1.0000
62	Inchon	0.7313	0.9343	0.9882	0.9969	1.0000	1.0000
63	Kwangyang	2.4999	3.9923	4.8655	5.1840	5.4287	5.4856
64	Colombo	0.6695	0.8518	0.9317	0.9616	0.9919	1.0000
65	Kaoshiung	0.7171	0.9219	0.9705	0.9857	0.9957	1.0000
66	Keelung	0.6030	0.8614	0.9582	0.9881	1.0000	1.0000
67	Taichung	2.0778	2.8894	3.2525	3.4770	3.6480	3.7016
68	Bangkok	1.6944	2.2064	2.5338	2.7214	2.8839	2.9689
69	Laem Chabang	1.3509	1.7748	1.9262	1.9932	2.0453	2.0520
70	Da Nang	11.3390	11.4598	11.4598	11.4598	11.4598	11.4598
71	Qui Nhon	9.7338	9.7338	9.7338	9.7338	9.7338	9.7338
	Geometric Average	1.4388	1.6804	1.7710	1.8101	1.8407	1.8556
	Max	16.7288	16.7288	16.7288	16.7288	16.7288	16.7288
	Min	0.3131	0.5163	0.6428	0.7363	0.8704	1.0000
	Std. Deviation	3.4076	3.3720	3.3631	3.3600	3.3579	3.3559
	No. of ports outside order-m frontier (percentage)	27 (38.03)	23 (32.39)	22 (31.00)	18 (25.35)	13 (18.31)	

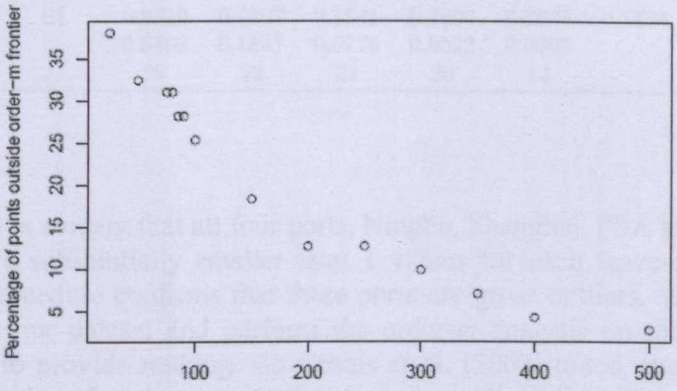

Figure 1: Percentages of points outside the order-m frontier as a function of *m*

Simar (2003) highlighted that if the sample observation remained outside the order-*m* frontier as *m* increases, then such an observation is a likely outlier. From Table 2, we observe that the container ports of Ningbo (9), Shanghai (11), PSA International Ltd. (60) and Busan (61) are potential outlier as it has the furthest distance from 1, as the value of *m* increases. However, this does not automatically indicate they are outliers. In this regard, Simar (2003) proposed an outlier detection strategy using the order-*m* partial frontier approach. Simar (2003) suggested a leave-one-out order-*m* estimate that allows for the detection of super-output-efficient DMUs. Here, for each port, its order-*m* output efficiency is calculated after deleting itself from the reference set. As such, we conduct a leave-one-out order-*m* efficiency $\hat{\lambda}_m^{(i)}$ for $i = 9;\ 11;\ 60;\ 61$. Table 3 below shows the results:

Table 3: Leave-one-out order-m output efficiency measures for container ports Ningbo (9), Shanghai (11), PSA International Ltd. (60), and Busan (61). For each container port, the first row is the output-oriented efficiency estimate, the second row shows the Monte-Carlo standard deviation (B=4000) and the third row indicates the number of container ports situated outside the respective order-*m* frontier.

Unit(i)	$\hat{\lambda}_{25}^{(i)}$	$\hat{\lambda}_{50}^{(i)}$	$\hat{\lambda}_{75}^{(i)}$	$\hat{\lambda}_{100}^{(i)}$	$\hat{\lambda}_{150}^{(i)}$	$\hat{\lambda}_{FDH}^{(i)}$
9	0.1622	0.2044	0.2234	0.2311	0.2359	0.2366
	2.6348	1.5366	0.9603	0.6122	0.2200	
	27	25	23	22	16	
11	0.2739	0.3294	0.3520	0.3590	0.3633	0.3642
	1.3488	0.6890	0.3484	0.2045	0.0636	
	27	23	23	20	12	
60	0.2551	0.3068	0.3282	0.3356	0.3401	0.3410
	1.4663	0.7468	0.4016	0.2333	0.0679	
	27	23	22	20	14	
61	0.6319	0.7297	0.7531	0.7605	0.7678	0.7735
	0.5199	0.1847	0.0775	0.0522	0.0305	
	27	23	23	20	14	

From Table 3, it is evident that all four ports, Ningbo, Shanghai, PSA International Ltd. and Busan, have substantially smaller than 1 values for each leave-one-out order-*m* analyses. The procedure confirms that these ports are gross outliers. We remove these four ports from our dataset and perform the order-*m* analysis on the rest of the 67 container ports to provide ranking. As Cazals et al. (2002) noted, the choice of *m* is arbitrary. The choice of *m* depends on the researcher's discretion. The rule of thumb is that the choice of *m* will determine the level of robustness of the analysis. We conducted an $m = 75$ frontier for the remaining 67 container ports, each replicated under a Monte-Carlo simulation of 4,000 times. At $m = 75$ we attempt to strike a balance between being sufficiently restrictive to lessen the effect of outliers, but allow for meaningful comparisons across the ports at the same time. The results are displayed in Table 4. With $m = 75$, 28:35 percent of the remainder ports lie beyond the frontier. Thus, 71.65 percent were used to determine the maximal output estimate of order-*m* while 28.35percent were left out, with 19 ports out of 67 that lying outside the frontier. Though these ports may be considered as outliers, they are not gross outliers, and lie near the order-*m* frontier, considering a threshold of 0.4 deviations from the frontier. We rearrange the order-*m* estimates for value $m = 75$ to depict the most efficient container ports, while maintaining consideration that the ports Mundra, Chittagong, Lianyungang, Zhangjiagang, Mumbai, New Mangalore, Niigata, Shimonoseki, Bintulu, Davao, and Zamboanga as ports that lie on the frontier. It can be seen that Subic Bay of the Philippines is the most inefficient port, producing an output of 5 percent of the cargo throughput of the maximum expected value of 75 other container ports, while Hong Kong is the most efficient container port producing 1.45 times of the maximal expected output. The Asian container port's average efficiency in 2007 was estimated to be around 1.7502, denoting the group of ports only produced 0.57 times the maximal expected output.

Table 4: Container Ports Rank using Output oriented order-*m* estimates

Rank	Port	$\hat{\lambda}_{75}$	Rank	Port	$\hat{\lambda}_{75}$
1	Hong Kong	0.6880	36	Jurong	1.2538
2	Guangzhou	0.8065	37	Shantou	1.2752
3	Yantian	0.8503	38	Shekou	1.3361
4	Xiamen	0.8951	39	Tuticorin	1.3362
5	Qingdao	0.9038	40	Shimizu	1.3969
6	Tianjin	0.9252	41	Nagoya	1.4389
7	Colombo	0.9275	42	Tokyo	1.5474
8	Kaoshiung	0.9445	43	Laem Chabang	1.6343
9	Tg Pelepas	0.9448	44	Dalian	1.8149
10	Keelung	0.9651	45	Iloilo	1.8992
11	Port Klang	0.9686	46	Penang	2.0735
12	Yantai	0.9747	47	Yokohama	2.3183
13	Inchon	0.9869	48	Sihanoukville	2.3761
14	Karachi	0.9934	49	Kwangyang	2.4464
15	Fuzhou	0.9960	50	Kochi	2.4777
16	Chennai	0.9993	51	Bangkok	2.5219
17	Pasir Gudang	0.9996	52	Mizushima	2.8624
18	Port Mohammad bin Qasim	0.9998	53	Manila	2.9460
19	Kitakyushu	0.9998	54	Taichung	3.2745
20	Mundra	1.0000	55	Osaka	3.3750
21	Chittagong	1.0000	56	Yokkaichi	3.6261
22	Lianyungang	1.0000	57	Kobe	3.9446
23	Zhangjiagang	1.0000	58	Kuantan	4.7163
24	Mumbai	1.0000	59	Pipavav	5.8009
25	New Mangalore	1.0000	60	General Santos	5.8102
26	Niigata	1.0000	61	Muara	6.1114
27	Shimonoseki	1.0000	62	Qui Nhon	9.7338
28	Bintulu	1.0000	63	Visakhapatnam	10.3659
29	Davao	1.0000	64	Da Nang	11.4598
30	Zamboanga	1.0000	65	Mitajiri	13.4514
31	Belawan	1.0351	66	Kawasaki	13.9042
32	Jawaharlal Nehru	1.1037	67	Subic Bay	16.7287
33	Hakata	1.1199		Geometric Mean	1.7502
34	Cebu	1.1932		Std. Deviation	3.4387
35	Tg Priok	1.1946		No. of ports outside frontier (percentage)	19 (28.35)

5. Conclusion

This study analyses the efficiency of major container ports in Asia using the robust non-parametric order-*m* estimator introduced by Cazals et al. (2002). The order-*m* estimator provides another way of looking at efficiency that is less sensitive to extreme values or outliers using small samples. The arrived estimates do not suffer from the curse of dimensionality. The estimates also reflect more realistic benchmark and define an output level that is best expected among any *m* ports chosen randomly from the population using no larger inputs than the given port. The leave-one-out version of the order-*m* estimator proposed by Simar (2003) is a useful tool for outlier detection.

The empirical results reveal that the ports of Ningbo, Shanghai, PSA International Ltd. and Busan are outliers. Removing these outliers, the order-*m* estimator was used to rank the remaining container ports. The results illustrate that the Asian container port's average efficiency in 2007 was 1.7502, denoting the group of ports only produced 57 percent of the maximal expected output of a random sample of 75 ports and could have handled 43 percent more traffic with the same resources.

References

Banker, R. D., Charnes, A. and Cooper, W. (1984), 'Some models for estimating technical and scale inefficiencies in data envelopment analysis', *Management Science,* 30(9), 1078-1092.

Cazals, C., Florens, J. and Simar, L. (2002), 'Nonparametric frontier estimation: A robust approach', *Journal of Econometrics,* 106, 1-25.

Charnes, A., Cooper, W. W. and Rhodes, E. (1978), 'Measuring the efficiency of decision making units', *European Journal of Operational Research,* 2, 429-444.

Debreu, G. (1951),'The coefficient of resource utilization'. *Econometrica*, 19(3), 273-292.

Deprins, D., Simar, L. and Tulkens, H. (1984), Measuring labour inefficiency in post offices, The Performance of Public Enterprises: Concepts and Measurements, Marchand, M., Pestieau, P. and Tulkens, H. (Eds.), Amsterdam, North-Holland.

Farrell, M. J. (1957), 'The measurement of productive efficiency', *Journal of the Royal Statistical Society A,* 120, 253-281.

Kneip, A., Simar, L. and Wilson, P. W. (2008), 'Asymptotics and consistent bootstraps for DEA estimators in nonparametric frontier models', *Economic Theory,* 24, 1663-1697.

Koopmans, T. C. (1951), An analysis of production as an efficient combination of activities, Activity Analysis of Production and Allocation, Cowles Commission for Research in Economics, Monograph 13, Koopmans, T. C. (Ed.), Wiley, New York.

Korostelev, A., Simar, L. and Tsybakov, A. (1995a), 'Efficient estimation of monotone boundaries', *The Annals of Statistics,* 23, 476-489.

Korostelev, A., Simar, L. and Tsybakov, A. (1995b), On estimation of monotone and convex boundaries, Pub. Inst. Stat. Univ. Paris XXXIX, 3-18.

Notteboom, T., Coeck, C. and van den Broeck, (2000), 'Measuring and explaining the relative efficiency of container terminals by means of Bayesian stochastic frontier models', *International Journal of Maritime Economics,* 2(2), 83-106.

Simar, L. (2003), 'Detecting outliers in frontier models: A simple approach', *Journal of Productivity Analysis*, 20, 391-424.

Simar, L. and Wilson, P. W. (2000a), 'Statistical inference in nonparametric frontier models: The state of the art', *Journal of Productivity Analysis,* 13, 49-78.

Vis, I. F. and de Koster, R. (2003) 'Transshipment of containers at a container terminal: An overview', *European Journal of Operational Research*, 147, 1-16.

Wang, T., Cullinane, K. and Song, D. (2005), Container Port Production and Economic Efficiency, Palgrave Macmillan, New York.

Wang, T., Song, D. and Cullinane, K. (2003), 'Container port production efficiency: A comparative study of DEA and FDH approaches', *Journal of the Eastern Asia Society of Transportation Studies,* 5, 698-713.

Wheelock, D. and Wilson, P. (2003), Robust nonparametric estimation of efficiency and technical change in U.S. commercial banking, Department of Economics, University of Texas, Austin, Texas 78712 USA.

Wilson, P. W. (2008), 'FEAR: A Software Package for Frontier Analysis with R', *Socio-Economic Planning Sciences*, 47, 247-254.