This is the Pre-Published Version.
Yip, T. L., Sun, X. Y., & Liu, J. J. (2011). Group and individual heterogeneity in a stochastic frontier model: Container terminal operators. European Journal of Operational Research, 213(3), 517-525

Group and Individual Heterogeneity in a Stochastic Frontier Model: Container Terminal Operators

Tsz Leung Yip*, Xin Yu Sun, and John J. Liu

Department of Logistics and Maritime Studies
Faculty of Business
The Hong Kong Polytechnic University
Hung Hom, Hong Kong

* Corresponding author. Tel.: +852 2766 4631; Fax: +852 2330 2704. E-mail address: lgttly@polyu.edu.hk

(5 January 2011)

ABSTRACT

Container ports are a major component of international trade and the global supply chain. Hence, the improvement of port efficiency can have a significant impact on the wider maritime economy. This paper deconstructs a representation in the existing literature that neglects the heterogeneity of individual and group-specific terminal operators. In its place, we present a hierarchical model to make a connection between efficiency and terminal operator group characteristics. The paper develops a stochastic frontier model that controls not only individual heterogeneity but also group-specific variations. The model decomposes the total stochastic derivation from the frontier into inefficiency, individual heterogeneity, group-specific variations, and noise components, with the estimation being performed using Markov chain Monte Carlo simulations. The validity of the model is tested with a panel of container terminal operator data from 1997-2004. Our findings show that terminal operator groups are important in promoting terminal efficiency at the global level, and that the operators with stevedore backgrounds show a higher efficiency than carriers.

Key Words

Stochastic processes; Stochastic production frontier; Markov processes; Container terminal operators; Port globalisation; Group-specific

Highlights

- We decompose individual and group-specific variations in frontier analysis.
- This study is at the terminal level rather than port level.
- Inefficiency is overestimated by a homogeneous frontier analysis.
- Terminal operator groups generate more terminal throughput.
- Terminal operator groups are more efficient than individual operators.

1. Introduction

1.1. Background

In recent years, operational research methods have gained considerable importance in econometrics. The production and cost theories in economics make it possible to estimate production and cost functions empirically, and thus to investigate changes in both the productivity and technology of a firm. The conventional stochastic frontier method for estimating a frontier assumes that all firms are successful in reaching the efficient frontier (and only deviate randomly). If, however, firms are not always at the frontier, then the conventional estimation method will not reflect the efficient production or cost frontier against which to measure efficiency. Empirical estimations for the port production function have been performed by Chang (1978) and Tongzon (1993), whereas Kim and Sachis (1986), Martínez-Budría et al. (2003), Martínez-Budría et al. (1999), and Jara-Díaz et al. (2002) estimated the cost functions of ports for both single-output and multiple-output cases. Using a single frontier function, Liu (1995), Notteboom et al. (2000), and Estache et al. (2002) estimated production frontiers or cost frontiers while recognising that some ports may not be at the efficient frontier.

Today, the port industry has a hierarchical structure. Each port has many terminals, which are operated by one or several operators. For example, the Hong Kong Port has 9 terminals operated by six operators. From Table 1, there are 1.8 operators and 5.0 terminals in a port on average. The operators are the firms to operate the terminals for their own objectives. Obviously, container terminal operators are the decision making units (DMU). As a DMU, each operator in a port makes his own operation and technical decisions. At the terminal level, the efficiency level should be different between different operators in a port. At the port level, different operators in a port should have certain similarity in production efficiency because they share the same water depth and other natural conditions. At the country level, the operators from different ports within the same country should also have certain similarity in production efficiency because they are subject to the same government regulations and legal systems. At the global level, most terminal operators in the world today belong to several major terminal operator groups such as HIT, DP world, and PSA (Figure 1). Two operators, even from different countries, should have certain similarity in production efficiency if they belong to the same terminal operator group.

However, despite the efforts of the aforementioned studies, three fundamental issues remain unaddressed. First, existing literature on efficiency does not address the group-specific effects over time, providing our motivation to examine the efficiency advancement of both groups and individuals. Instead of treating ports as the decision making unit (DMU), this paper treats container terminal operators as the DMU, which represents a divergence from previous port efficiency studies (e.g., Gonzalez and Trujillo, 2009). Second, heterogeneity is generally ignored in port efficiency studies but is mistakenly included in the stochastic error term. Unlike other industries, ports are characterised by their geographical and operational settings. Terminal operators from different groups, different locations, and different times are assumed to have different characteristics. We attempt to separate the group and individual heterogeneity from stochastic errors and this attempt leads to a substantial simulation effort.

Third, the group-specific effect has not been studied in port efficiency studies. Seaports are characterised by global competition in a number of dimensions. They have sought to exploit network effects in the containerisation era, and terminal operators have attempted to expand their line of activities through vertical and/or horizontal integrations along the transport chain (Figure 2). Currently, several terminal operator groups having derived competitive advantages are coming up against one another. However, despite the importance of the port industry, there has been little attention paid to the underlying motives of terminal operator grouping from a scientific perspective. The aim of this paper, therefore, is to explore the efficiency motives of globalisation of terminal operation.

1.2. Terminal operator group (TOG)

The structure of the container terminal industry has changed since port privatisation started in the 1990s. Governments contract out the ownership and management of ports and terminals, and today container terminals are run for commercial objectives. The port industry has a particularly global structure, and global and multinational players, TOGs (terminal operator groups), are becoming increasingly dominant. There are two generic globalisation strategies in the globalisation of terminal operation: (1) Horizontal merger initiated by leading stevedores; and (2) Vertical integration initiated by global carriers (shipping lines).

There are many reasons for the dominance of TOG, and the discussion of multinational behaviour brings together a number of economic theories (e.g., Caves, 2007). Studies of multinational service industries have received increased attention from researchers, e.g., multinational banks (Chang et al., 1998), hotels (Shang et al., 2008), insurance companies (Fenn et al., 2008), and airports (Oum et al., 2008). There has, to date, been no efficiency study concerning TOG. One reason for this is that existing port studies are based on port data, not terminal data, while terminal operator groups operate in several countries. Our terminal-based data collection makes a study into terminal operation globalisation feasible. However, more fundamentally, we argue that TOGs are more efficient because operators actually create terminal globalisation in order to improve the efficiency of their operations.

In summary, the port industry is different from other service industries. The clients of terminal operators are shipping lines (carriers), and container handling is highly standardised worldwide. Terminal operators provide more-or-less the same container handling services to carriers. In particular, operator groups prefer market standardised services worldwide so as to reap maximum benefits from the economies of scale that underlie their learning curve. Within this context, in this paper we develop a rigorous econometric model to test the effects of globalisation on the efficiency of terminal operation.

2. Literature Review

A deep knowledge of port firms' productivity results is essential not only to decide where, when, and how much to invest, but also to suggest optimal tariff structures. There have been numerous productivity or efficiency studies of ports. Wanhill (1974) suggested that the productivity of ports depends on the right trade-off between the costs of providing infrastructure (berth) and the time costs of the ship's stay in the port. The manual on port planning prepared by the UNCTAD (1978) for developing countries followed the same line of work as Wahnill's (1974) study. It relied on Monte-Carlo simulation techniques to calculate the costs of different types of terminals according to terminal features and ships' stay in port. Similar works include Jansson and Shneerson (1982), Shneerson (1981, 1983), and Fernández et al. (1999), all of whom adopted a queuing model as the basic form of port service production function and assumed ships' arrival is random and follows a Poisson distribution. Such studies are helpful for individual port planning.

As logistics and supply chains have evolved into the artery of the global economy, the efficiency of ports has become an important factor affecting a nation's international competitiveness. Thus, monitoring and comparing one's ports with others in terms of overall efficiency has become an essential part of many countries' microeconomic reform programs. The pressure to boost port competitiveness has triggered an increasing number of port benchmarking studies, especially for leading container ports.

In all efficiency studies, efficiency is measured by comparing observed and optimum costs, production, revenue, or whatever parameter the organisation is assumed to pursue, subject to the constraints on quantities and prices. The optimal quantity is termed the frontier, and the efficiency is then calculated as the distance between the observed quantity and the frontier. In empirical research, two methods are widely used to calculate or estimate the frontier functions and thereby measure efficiency: data envelope analysis (DEA) and stochastic frontier analysis (SFA). DEA is a deterministic method based on linear programming and was first introduced by Charnes et al. (1978). In contrast, SFA is an econometric method accounting for random shocks and measurement errors, and was first proposed by Aigner et al. (1977) and Meeusen and van den Broeck (1977). Cullinane et al. (2006) compared the results from DEA and SFA on port efficiencies and found high correlations between the results from the two approaches. In port literature, Hung, Lu, and Wang (2010) introduced scale efficiency into DEA analysis, and Sharma and Yu (2010) studied container terminals using the DEA model. Dowd and Leschine (1990) and Talley

(1994) used index number approaches, which allow comparisons to be made of the efficiency among various ports and throughout time for a single port, in order to study ports' productivity. Index number procedures generally construct a ratio-type productivity/efficiency measure, without the need for statistical estimation of a production or cost function. Many studies which have used these two methods to study port efficiency assume the homogeneity of the global port industry. Instead, we consider the heterogeneity of terminal operators and examine factors that explain why terminal operator groups play their leading role in the port industry.

Heterogeneity of DMUs often exists in the presence of geographical features. Banker et al. (1986) first discussed the idea of categorical variables which is a clustering technique to solve the heterogeneity variations in DMUs. Cook et al. (1998) pointed out that clustering of DMUs may appear at different levels. Doyle and Green (1994) introduced the cross efficiency evaluation method to address the heterogeneity of DMUs. Recently, Lee (2010) attempted the group effect with parametric models.

Previous studies have compared the temporal variations between ports (e.g., Gonzalez and Trujillo 2009). The terminal operators should be treated as independent decision-makers, but the existing literature considers ports as the decision-makers. Previous efficiency studies are at the port level, but this one is at terminal level since port globalisation exists due to terminal operators rather than port operators.

3. Methodology

3.1. Empirical analysis

In container terminals, the output is the number of containers handled, and the input is the equipment and manpower used to handle the containers. The level of terminal output is an important indicator of a terminal's efficiency. As the data for container throughputs at terminals are reliable and well documented, throughput in TEU (twenty-foot equivalent unit) is the most frequently used indicator for container terminals, although there are alternative indicators. The model we use incorporates the necessary physical characteristics of container terminals, such as quay cranes, yard equipment, and the number of berths, as inputs to container terminal production (Table 1).

The production frontier of terminal operator i at time t is parameterised as the Cobb-Douglas function:

$$y_{it} = \alpha_{it} + X_{it}B - \Delta_{it} + \varepsilon_{it} \tag{1}$$

$$\alpha_{it} = \overline{\alpha} + \Pi_i\Theta + r_1 t + r_2 t^2 + \nu_i , \tag{2}$$

where

$y_{it} =$ the logarithm of observed output (herein Container Throughput in TEUs) of the i-th operator at year t

$X_{it} =$ the matrix of the logarithm of observed inputs (e.g., cargo handling, terminal infrastructure, and storage facilities)

$\Delta_{it} =$ the positive random deviation from the frontier (which means inefficiency)

$\varepsilon_{it} =$ the time-varying measurement error. It represents the measurement error with a normal distribution independent of both operator and year, i.e., $\varepsilon_{it} \sim N(0, \sigma_{\varepsilon}^2)$. Thus, ε_{it} is the time-varying error.

$\Pi_i =$ the matrix of observed terminal characteristics (e.g., terminal factors, port factors, and country factors)

$t =$ time

$\overline{\alpha} =$ constant

Equation (2) shows that the intercept of the logarithm production frontier varies across both individual terminal operators and time (with time trend terms). Part of the variation across individual operators can be explained by individual profiles denoted as $\boldsymbol{\Pi}_i$ and part of the variation is unobserved and is thus modelled as random denoted as ν_i , whose distribution is parameterised in as normal with zero mean, i.e., $\nu_i \sim N(0, \sigma_{\nu}^2)$.

Our special feature is to measure the efficiencies after controlling for the individual heterogeneity. We achieve this by randomising the intercept of the log production frontier and letting the distribution be conditional on observable individual profiles.

In the port industry, three levels of individual profiles can be observed -- terminal, port and country. The unobserved factors are specified by random component v_i to the intercept. Ignoring v_i will bias the estimates of parameters, because the uncontrolled individual effects will cause observations from the same individual to be correlated.

Under our model specification, the efficiency level of an individual *i* at time *t* is defined as $h_{it} = \exp(-\Delta_{it})$, which can be understood as the percentage achieving the production frontier. In our case, we not only want to know the inefficiency, but also want to investigate the determinants of productivity. Many authors have previously presented a two-step method (Zheng et al., 1998; Shao and Lin, 2002). In the first step, the productivity measure is computed by SFA or DEA. In the second step, analysis is conducted to examine the relationship between inefficiency and various influential factors. However, such a method will be statistically biased since efficiency is also an estimator. Our specification augments efficiency as a function of determinants by $\Delta_{it} = \exp(Z_{it} g_i)$, where Z_{it} are the vectors of variables measuring port observable characteristics including terminal operator group, and g_i are coefficients of inefficiency parameters. g_i are distributed with $N(\bar{g}, \Sigma)$, where $\Sigma = diag\left(\sigma_{g1}^2, ..., \sigma_{gN}^2\right)$ is the diagonal matrix of variances. We thus use one step to estimate this hierarchical model.

In contrast to previous models, our model is enhanced to quantify the effects of operator globalisation by adding dummy variables in order to control operator group effects (see Table 1). A recent trend in the container port industry is that global stevedore and global carriers are investing in several terminals in the same country and also in different countries. The operator group effects are observable only if the individual units are container terminals, rather than ports. For the specification of our model, the key feature is that the distribution of Δ_{it} is conditional on Z_{it}, the vector of variables measuring observable port characteristics and time affecting the production inefficiency. To analyse the two generic globalisation strategies in the port industry, we use two dummy variables which refer to a terminal operator's background as carrier or stevedore. We therefore create a connection between inefficiency and terminal characteristics. Therefore, the model contains a linear trend *t*, quadratic trend t^2, carrier dummy, and stevedore dummy. We include linear and quadratic time trends in our model in order to account for technological change over time.

A typical criticism of port efficiency studies is that there are no credible data on the labour inputs of the port or of terminal operators. Tongzon (2001) counted the number of stevedores and other employees that work in terminals. However, because the direct counting of labour inputs is not possible in most terminals, Yan et al. (2009) ignored labour inputs by assuming little variation across terminal operators. Hui et al. (2010) used housing price as a proxy variable of local labour costs of port operation. In our study, because of the difficulty in collecting labour cost data, we include the GDP per capita at the country level to proxy the labour costs of terminal operations.

3.2. Estimation procedure with MCMC simulation

The Markov chain Monte Carlo (MCMC) simulation is used to determine the unknown parameters. The MCMC approach is a class of algorithms for sampling from probability distributions based on constructing a Markov chain that has the same desired distribution as its equilibrium distribution. The state of the chain after a large number of steps is then used as a sample from the desired distribution, with the quality of the sample improving as a function of the number of steps. Usually it is not difficult to construct a Markov chain with desired properties. A more difficult problem is to determine the number of steps needed to converge to the stationary distribution within an acceptable error. A good Markov chain will have rapid mixing, where the stationary distribution is reached quickly starting from an arbitrary state. The MCMC simulation can augment and filter imperfect panel data of differential characteristics. One application of this critical feature is to interpolate and extrapolate statistically missing and censored diffusion data. It is inevitable that panel data collected from industry contain missing and erroneous data entries, which must be augmented and corrected, respectively.

Our estimation is performed based on the MCMC estimator developed in Bayesian statistics. The unknown parameters can be represented as $\Psi \equiv \left(\overline{\alpha}, \Theta, \gamma_1, \gamma_2, \mathrm{B}, \sigma_v^2, \overline{g}, \Sigma, \sigma_\varepsilon^2\right)$. The MCMC estimator draws from the joint posterior distribution $p\left(\Psi,\{\nu_i, g_i\}_i \middle| Data\right)$, with $\{\nu_i, g_i\}_i$ representing the instruments of endogenous individual effects in stochastic frontier equations. The data posterior is expressed as:

$$p\left(\Psi,\{\nu_i, g_i\}_i \middle| Data\right) \propto \rho(\Psi)\cdot\prod_{i=1}^{N}\left\{\phi(g_i;\overline{g},\Sigma)\cdot\phi\left(\nu_i;\sigma_\nu^2\right)\cdot\prod_{t=1}^{T_i}\phi\left(y_{it};\overline{\alpha},\Theta,\tau,\mathrm{B},\sigma_\varepsilon^2,g_i,\nu_i\right)\right\}. \quad (3)$$

Since the functional form of the data augmented posterior in Eq. (3) is complicated, it is impossible to derive analytical properties of it. We use the Monte Carlo simulation to take random draws from the posterior and the empirical properties of the draws will be used to approximate the theoretical ones. Appendix of the paper presents the details of the MCMC algorithm to take random draws from the augmented posterior in Eq. (3).

3.3. Data collection

We consider a panel data set for the terminal operators which come from the container ports ranked in the top 100 in 2005, with the period covered being from 1997 to 2004. For each container terminal operator of these ports, we collected data on output, terminal inputs, port characteristics and country features (Table 1). Most of the information can be found in the Containerisation International Yearbooks. We also used a subscription database, Containerisation Intelligence Online, to obtain the addresses of the websites of each terminal from which further information was gained. Additional useful information was also obtained from the websites of port authorities and government agencies. The country data is obtained from the World Bank.

After we removed some missing data, a set of unbalanced panel data was created with 597 observations in total. The data covered 141 terminal operators from 78 container ports.

3.4. Model Validation

In implementation, we employ non-informative priors on the parameters. As shown in Figure 3, the variance of the posterior is smaller than that of prior. To confirm the convergence, we run the Gibbs sampler from different starting values of the parameters. For each of the runs, we plot the time-series of the generated variables such as Figure 4. The draws for parameters converge very fast and have much better mixing properties compared with these second-layer parameters. In general, the Gibbs draws converge after about 10,000 draws. Performing many different runs from diverse starting points and changing priors had certain but not substantial effects on estimates of inefficiency.

4. Empirical results and discussion

In Table 2, we show posterior means and standard deviations of coefficients of the stochastic production frontier. The heterogeneous model is based on our model specification and estimated by MCMC. The conventional model is that all the operators face all the common frontiers and to ignore their heterogeneity.

4.1. Homogeneous versus heterogeneous frontiers

A conventional model does not distinguish individual heterogeneity and inefficiency. The wide variation in terminal operation across countries introduces a considerable amount of terminal heterogeneity. Conventional models (Figure 5) then overestimate inefficiency by including heterogeneity in inefficiency. In our model whose results are showed in Figure 6, we randomise the intercept of the production frontier to account for individual heterogeneity (the so-called true random effects model in Greene, 2008). When the operators differ in their adopted technologies, and such differences are not well controlled, the estimated inefficiency absorbs both the heterogeneity and inefficiency and the estimation is thus inevitably biased. Unlike the conventional model using a convenient one-parameter distribution form (such as a half-normal or exponential distribution) to model random inefficiency, we used a more flexible log-normal distribution with two parameters. This specification enabled us to interact some observable managerial inputs with inefficiency, in order to seek policy implications.

Figures 5 and 6 plot the estimated distribution of efficiency level in different years for the two models, respectively. Figure 5 refers to Eq. (1) by assuming the operators are homogeneous *(i.e.,* α_{it} = constant), while Figure 6 is based on Eq. (1) and (2) in which the operators are heterogeneous. The different patterns of Figure 5 and Figure 6 are due to the different assumptions of homogeneous versus heterogeneous frontiers.

Inefficiency is significantly overestimated by a conventional SFA, as shown in Figure 5 in which the individual heterogeneity in production frontiers is controlled only by the observables. This suggests that the estimated inefficiency from the conventional model does not allow for individual heterogeneity across terminals. In fact, the conventional model absorbs the individual heterogeneity in the production frontier, and thus the distribution of individual level efficiency shifts to the left. All the terminals are measured against "the most efficient terminal" in terms of efficiency when using the conventional model.

After controlling for the observed heterogeneity, the estimated value of σ_v^2 is still significant with large magnitude, indicating the heterogeneity in the adopted technologies caused by many unobserved or omitted variables. The range of v_i is from -2.04 to 9.04 to show there is significant heterogeneity in operators in Figure 7. Obviously the homogeneous stochastic frontier model cannot meet the situation of the port industry. Every operator should have its own frontier to reveal its real inefficiency.

Figure 8 shows four examples of individual efficiency change of the heterogeneous model and conventional model. The efficiency in the heterogeneous model is higher than that in the conventional model. But the difference is not constant. The efficiency difference in Singapore PSA terminal and Hong Kong ModernTerminals terminal is very small compared with other two terminal comparisons. This is because Singapore and Hong Kong terminals are well known due to their efficiency. In the conventional model, the efficiency is measured based on the most efficient terminals such as Singapore and Hong Kong. Therefore, the efficiency difference is huge between poorly performing terminal operators in the heterogeneous model and conventional model.

4.2. Trend effect

Modelling time varying inefficiency is also possible based on the model specification used. The sign of the estimated coefficient of the linear trend effect t is negative (Table 2); i.e., there is a declining linear effect on terminal operator throughput over time. This reflects the fact that the productivity of equipment and technologies declines as time passes. However, as the sign of the quadratic trend t^2 is positive, there is a non-linear inference of terminal productivity over time. Within the global port market, there is constant pressure to enhance the efficiency of terminal operation. Effective use of technology enhances efficiency and technological advancement is significant in terminals. New equipment is purchased, and methods of cargo handling are changed to effectively handle the increased amount of throughput. The linear and quadratic effects together suggest that the large productivity enhancement is a result of the effect of the non-linear component of the trend.

4.3. Horizontal merger and vertical integration

As shown in Table 2, the heterogeneous model identifies that the terminal operator group, Carrier Group (g_4) and Stevedore Group (g_5), in general improves terminal efficiencies. The model provides strong evidence to explain why terminal operators are moving towards globalisation. The model further shows that a global stevedore background is more efficient than that of a global carrier. This is because there are inevitable conflicts of interests between terminal and carrier operations. It is well known that the most efficient terminal operation occurs at a higher economic scale of throughput than that of carrier operation (e.g., *McConville 1999, Chapter 13*). If the terminal is operated by a carrier, terminal efficiency will likely be compromised. Findings show that global-carrier-based terminal operators on average outperform local operators, while there are no negative impacts of the involvement of carriers in the terminal business.

Small or local terminal operators, however, do not enjoy the same advantages as terminal operator groups. As operator groups increase their market share, the small operators are losing some of the container traffic that used to flow through them.

4.4. Group versus individual

Our model shows that group operators (Carrier Group g_4 and Stevedore Group g_5), are advancing in efficiency more rapidly than individual operators. A possible reason behind this observation is that groups may learn more quickly than individuals, because experience could be shared within the group and efficiency improved by benchmarking performance against multiple terminals. However, further research is needed into the relation of efficiency to learning curve effect.

Another important feature newly identified by the model is that Stevedore Groups show a steady improvement in efficiency, while the Carrier Groups present a higher fluctuation in efficiency. The stevedore operators serve multiple carriers, while carrier operators are dedicated to one particular carrier. Thus, the stevedore operators can easily develop a portfolio of operation for different carriers so that efficiency is steadily improved. In contrast, carrier operators streamline their operations according to their carriers and efficiency inevitably depends on particular carriers' operations. Although carrier operators can still take advantage of multiple terminal operations and perform better than individual operators, carrier operators face a more fluctuating efficiency. As carrier operators are more sensitive to the market, they may quit the terminal market and later re-enter the terminal market depending on the carrier's strategy.

4.5. Other implications

The results of the study contribute to the quest for multiple terminal management concepts in port research, in line with the integration that has been taking place at the global level, and the integration of port management into the carrier. The study identifies certain parameters perceived to be instrumental in the integration of ports in global supply chains. Further research is required to ratify the development of a measurement instrument for assessing vertical and horizontal integrations in globalisation, since the lack of such an instrument has hindered research in the area. Without a valid and reliable instrument measuring port integration in supply chains, generalisable implications and strategies are difficult to generate.

This study identifies a positive relationship between terminal integration in the supply chain and terminal performance. It is important for this association to be replicated empirically using different ports and different contexts and performance measures. It is also important for measures of competitiveness to be incorporated into the measurement of terminal performance. The privatisation of ports and terminals, together with the quest for competitiveness, means that conventional performance measures such as market share, sales growth, and even profitability have become legitimate performance measures for ports. A container terminal with high efficiency indicators may not be necessarily competitive due to the higher costs involved in becoming more efficient.

It is likely the market share of the terminal operator groups will increase, as the port sector is particularly sensitive to economies of scale. Container terminals are part of a capital-intensive industry and require a volume business to sustain their business investment.

5. Concluding remarks

In this paper, we considered a stochastic frontier model that allows for group-specific temporal patterns, which assumes that container terminal operators from the same operator group have the same group-specific parameter. We further studied how operator groups compete with other groups on the basis of productivity efficiency. The results contribute to an understanding of general group-based competition, such as multinational companies and supermarket groups. The involvement of international companies in the operation of container terminals has been a major factor in boosting productivity. As the port sector becomes more internationalised, improvements in efficiency are expected to continue.

This paper is one of the first attempts to appear in the literature to empirically investigate whether terminal operator groups are more efficient than individual terminal operators. As the heterogeneity of terminal operation cannot be ignored, the heterogeneous model is developed to separate the heterogeneity from inefficiency. The results confirm that terminal operator groups generate 35.9% and 45.4% more throughput for carrier groups and stevedore groups, respectively. The results also confirm the improvement of efficiency due to global grouping, and the efficiency advancement should be a strong motive behind observed globalisation of terminal operation. By confirming the efficiency of terminal operator groups, this study not only highlights the contribution to terminal grouping, but also provides empirical evidence to policy makers who design and seek to implement more efficient terminal operation. By presenting the terminal efficiency associated with different operator backgrounds, the study emphasises the importance of establishing regulatory actions to regulate the globalisation of terminal operator groups.

In the era of globalisation, governments have found it necessary to open up terminal services to terminal operator groups. In general, this has resulted in increased throughput and efficiency of container handling. Terminal operator groups have advanced their efficiency faster than individual operators. Between the two types of terminal operator groups, stevedore operators show more steady advancement than carrier operators. However, to sustain and encourage efficiency achievements, further research is needed to sustain the terminal efficiency improvements. Regarding the economic policies, this study should be extended to consider the terminal concession effect and how the regulations shape terminal efficiency.

Acknowledgement

This research was funded by the Hong Kong Polytechnic University under PPRI Grant #J-BB7E.

Appendix - MCMC algorithm

In order to conduct the Bayesian inference, we need to complement the likelihood function in (3) with a prior distribution on the parameters. We choose a proper prior distribution with the following product structure:

$$\rho(\Psi) = \rho(\bar{\alpha}, \Theta, B, r_1, r_2) \cdot \rho(\bar{g}) \cdot \rho(\Sigma) \cdot \rho(\sigma_\nu^2) \cdot \rho(\sigma_\varepsilon^2). \qquad \text{(A1)}$$

We briefly specify the priors for parameters. $(\bar{\alpha}, \Theta, B, r_1, r_2) \sim N(\Lambda_0, V_\Lambda)$,

$\bar{g} \sim N(\bar{g}_0, V_{\bar{g}})$, $\mathbf{\Sigma} \sim IW(r, rS)$, $\sigma_\nu^2 \sim IG\left(\frac{d_\nu}{2}, \frac{d_\nu c_\nu^2}{2}\right)$, and $\sigma_\varepsilon^2 \sim IG\left(\frac{d_\varepsilon}{2}, \frac{d_\varepsilon c_\varepsilon^2}{2}\right)$, where

$IW(r, rS)$ and $IG\left(\frac{d_\nu}{2}, \frac{d_\nu c_\nu^2}{2}\right)$ denote the inverted Wishart distribution and scaled

inverted Chi square distribution, respectively.

We separate several blocks to sample, $(\bar{\alpha}, \Theta, B, r_1, r_2)$, σ_ε^2, $\{g_i\}_i$, $\{\nu_i\}_i$, $\bar{g}$, $\mathbf{\Sigma}$, and σ_ν^2. Firstly, we sample ν_i and g_i for each operator:

(a). Sampling v_i for each i from $N\left(\dfrac{\sum_{t=1}^{T_i} \hat{y}_{it} \Big/ \sigma_\varepsilon^2}{T_i \big/ \sigma_\varepsilon^2 + 1 \big/ \sigma_v^2}, \dfrac{1}{T_i \big/ \sigma_\varepsilon^2 + 1 \big/ \sigma_v^2}\right)$,

where $\hat{y}_{it} = y_{it} - \alpha_{it} - X_{it}B + \exp(Z_{it}g_i)$.

(b). Sampling g_i for each i from

$$p(g_i|\bullet, Data) \propto \exp\left\{-\frac{1}{2}\left[(g_i - \bar{g})' \Sigma^{-1}(g_i - \bar{g}) + \frac{1}{\sigma_\varepsilon^2}(Y_i + \exp(Z_i g_i))'(Y_i + \exp(Z_i g_i))\right]\right\}, \tag{A2}$$

where $Y_i = \{y_{it} - \alpha_{it} - v_i - X_{it}B\}_t$. We cannot directly draw a sample from Eq. (A2), being non-standard distributions. From our experience, a simple random walk Metropolis-Hastings algorithm works very well and we choose it for our applications. Other parameters are sampled for all operators once.

(c). Sampling $\Lambda \equiv (\bar{\alpha}, \Theta, B, r_1, r_2)$ from $N(Dd, D)$, with $D = \left(\dfrac{1}{\sigma_\varepsilon^2}\tilde{X}'\tilde{X} + V_\Lambda^{-1}\right)^{-1}$,

and $d = \dfrac{1}{\sigma_\varepsilon^2}\tilde{X}'\tilde{Y} + V_\Lambda^{-1}\mathbf{\Lambda}_0$,

where $\tilde{Y} = \{y_{it} - v_i + \exp(Z_{it}g_i)\}_{i,t}$, and $\tilde{X} = \{1, \mathbf{\Pi}_i, t, X_{it}\}_{i,t}$.

(d). Sampling σ_ε^2 from $IG\left(\dfrac{d_\varepsilon + \sum_{i=1}^{N} T_i}{2}, \dfrac{d_\varepsilon c_\varepsilon^2 + \sum_{i=1}^{N}\sum_{t=1}^{T_i} e_{it}^2}{2}\right)$,

with $e_{it} = y_{it} - \alpha_{it} - X_{it}B - v_i + \exp(Z_{it}g_i)$.

(e). Sampling σ_v^2 from $IG\left(\dfrac{d_v + N}{2}, \dfrac{d_v c_v^2 + \sum_{i=1}^{N} v_i^2}{2}\right)$.

In our hierarchical model, the above parameters are upper level in our model. However, $\bar{g}$ and $\mathbf{\Sigma}$ are the lower level parameters since they are sampled based on the high level parameters.

(f). Sampling $\bar{g}$ from $N(Dd, D)$, with $D = \left(N\Sigma^{-1} + V_{\bar{g}}^{-1}\right)^{-1}$, $d = \sum_{i=1}^{N} \Sigma^{-1} g_i + V_{\bar{g}}^{-1} \bar{g}_0$.

(g). Sampling $\mathbf{\Sigma}$ from $IW\left(r + N, rS + (g_i - \bar{g})(g_i - \bar{g})' \right)$.

The MCMC simulation is implemented by a Metropolis sampling within Gibbs algorithm. The above steps (a) to (g) are repeated many times after the "burn-in" period, and the draws from the convergent distribution are used to construct the estimates for the unknown parameters. Although we do not model directly the individual heterogeneity across terminal operators in their cost functions, the MCMC method is flexible enough to incorporate individual heterogeneity in the parameters. For example, we can easily extend the model by modelling Ψ to vary across terminal operators following a joint normal distribution, which is conditional on the ports' characteristics. To estimate the extended model, we only need modify steps (a) to (b) to draw for each port. Then, conditional on the draws of Ψ, we can estimate the hyper-parameters governing the conditional distribution of Ψ as a simple multivariate regression model. This kind of hierarchical Bayesian approach has already shown flexibility and computational advantages in estimating models with random parameters in recent econometrics literature.

References

Aigner A, Lovell CAK, Schmidt P. Formulation and estimation of stochastic frontier production models. *Journal of Econometrics* 1977; 6(1); 21-37.

Caves RE. *Multinational Enterprise and Economic Analysis* (3rd edition). Cambridge University Press; 2007.

Chang S. Production function and capacity utilization of the port of Mobile. *Maritime Policy and Management* 1978; 5(4); 297-305.

Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. *European Journal of Operational Research* 1978; 2(6); 429-444.

Chang CE, Hasan I, Hunter WC. Efficiency of multinational banks: An empirical investigation. *Applied Financial Economics* 1998; 8(6); 689-696.

Cullinane K, Wang TF, Song DW, Ji P. The technical efficiency of container ports: Comparing data envelopment analysis and stochastic frontier analysis. *Transportation Research Part A* 2006; 40(4); 354-374.

Dowd TJ, Leschine TM. Container terminal productivity: A perspective. *Maritime Policy and Management* 1990; 17(2); 107-112.

Drewry. *Annual Review of Global Container Terminal Operators*. Drewry Shipping Consultants: London; 2002.

Drewry. *Annual Review of Global Container Terminal Operators*. Drewry Shipping Consultants: London; 2005.

Drewry. *Annual Review of Global Container Terminal Operators*. Drewry Shipping Consultants: London; 2007.

Estache A, Gonzalez M, Trujillo L. Efficiency gains from port reform and the potential for yardstick competition: Lessons from Mexico. *World Development* 2002; 30(4); 545-560.

Fenn P, Vencappa D, Diacon S, Klumpes P, O'Brien C. Market structure and the efficiency of European insurance companies: A stochastic frontier analysis. *Journal of Banking and Finance* 2008; 32(1); 86-100.

Fernández JE, de Cea J, Fernández JM. Port privatization in developing countries: The case of container terminals. *International Journal of Transport Economics* 1999; 26(2); 292-313.

Gonzalez MM, Trujillo L. Efficiency measurement in the port industry: A survey of the empirical evidence. *Journal of Transport Economics and Policy* 2009; 43(2); 157-192.

Greene WH. *Econometric Analysis* (6th edition). Pearson Prentice-Hill: Upper Saddle River, New Jersey; 2008.

Hui ECM, Ng MH, Xu JJ, Yip TL. The cargo throughput response to factor cost differentials: An analysis for the port of Hong Kong. *Transportmetrica* 2010 (in press).

Hung SW, Lu WM, Wang TP. Benchmarking the operating efficiency of Asia container ports. *European Journal of Operational Research* 2010; 203(3); 706-713.

Jansson JO, Shneerson D. *Port Economics*. MIT Press: Cambridge, Mississippi; 1982.

Jara-Díaz S, Martinez-Budría E, Cortes C, Basso L. A multioutput cost function for the services of Spanish ports' infrastructure. *Transportation* 2002; 29(4); 419-437.

Kim M, Sachish A. The structure of production, technical change and productivity in a port. *Journal of Industrial Economics* 1986; 35(2); 209-223.

Lee YH. Group-specific stochastic production frontier models with parametric specifications. *European Journal of Operational Research* 2010; 200(2); 508-517.

Liu Z. The comparative performance of public and private enterprises: The case of British ports. *Journal of Transportation Economics and Policy* 1995; 29(3); 263-274.

Martínez-Budría E, Díaz-Armas R, Navarro-Ibañez M, Ravelo-Mesa T. A study of the efficiency of Spanish port authorities using data envelopment analysis. *International Journal of Transport Economics* 1999; 26(2); 237-253.

Martínez-Budría E, Jara-Díaz SR, Ramos F. Adapting productivity theory to the quadratic cost function: An application to the Spanish electric sector. *Journal of Productivity Analysis* 2003; 20(2); 213-229.

McConville J. *Economics of Maritime Transport: Theory and Practice*. Witherby: London; 1999.

Meeusen W, van den Broeck J. Efficiency estimation from Cobb-Douglas production function with composed error. *International Economic Review* 1977; 18(2); 435-444.

Notteboom T, Coeck C, van den Broeck J. Measuring and explaining the relative efficiency of container terminals by means of Bayesian stochastic frontier models. *International Journal of Maritime Economics* 2000; 2(2); 83-106.

Oum TH, Yan J, Yu C. Ownership forms matter for airport efficiency: A stochastic frontier investigation of worldwide airports. *Journal of Urban Economics* 2008; 64(2); 422-435.

Shang JK, Hung WT, Lo CF, Wang FC. Ecommerce and hotel performance: Three stage DEA analysis. *Service Industries Journal* 2007; 28(4); 529-540.

Shao BM, Lin WT. Technical efficiency analysis of information technology investments: A two-stage empirical investigation. *Information and Management* 2002; 39(5); 391–401.

Sharma M, Yu SJ. Benchmark optimization and attribute identification for improvement of container terminals. *European Journal of Operational Research* 2010; 201(2); 568-580.

Shneerson D. Investment in port systems: A case study of the Nigerian ports. *Journal of Transport Economics and Policy* 1981; 15(3); 201-216.

Shneerson D. Short term planning for a ports system. *Maritime Policy and Management* 1983; 10(4); 217-250.

Talley WK. Performance indicators and port performance evaluation. *Logistics and Transportation Review* 1994; 30(4); 339-352.

Tongzon JL. The Port of Melbourne Authority's pricing policy: Its efficiency and distribution implications. *Maritime Policy and Management* 1993; 20(3); 197-203.

Tongzon J. Efficiency measurement of selected Australian and other international ports using data envelopment analysis. *Transportation Research Part A* 2001; 35(2); 113-128.

UNCTAD. Port development. A handbook for planners in developing countries. Second edition. UNCTAD: Geneva; 1978.

UNCTAD. *Port Marketing and the Challenge of the Third Generation Port*. UNCTAD: Geneva; 1990.

UNCTAD. Fourth generation port. *UNCTAD Ports Newsletter* 1999; 19; 9-12.

Wanhill SRC. A study in port planning: The example of Mina Zayed. *Maritime Policy and Management* 1974; 2(1); 48-55.

Yan J, Sun X, Liu JJ. Assessing container operator efficiency with heterogeneous and time-varying production frontiers. *Transportation Research Part B* 2009; 43(1); 172-185.

Zheng J, Liu X, Bigsten A. Ownership structure and determinants of technical efficiency: An application of data envelopment analysis to Chinese enterprises (1986–1990). *Journal of Comparative Economics* 1998; 26(3); 465–484.

Table 1: Summary of Variables

Variables	Mean	Std. Dev.	Minimum	Maximum
A. Terminal Output				
TEU: Container Throughput in TEUs (000's)	936.4	1,741.7	4.6	20,600
B. Terminal Inputs				
1. Cargo Handling Equipment:				
Cargo handling capacity at quay in tonnage [a]	385.0	470.7	23.9	5,416.2
Cargo handling capacity at yard in tonnage [b]	5,116.5	7,060.9	38.6	62,731.8
2. Terminal Infrastructure:				
Number of berths	5.1	5.2	1	37
Length of quay line in meters	1,361.3	1,181.6	200	9,000
Terminal area in squared meters (000's)	604.9	844.6	7.7	8,092
3. Storage Facilities:				
Storage capacity in number of TEUs (000's)	23.2	72.4	0.6	1,200
Number of electric reefer points	480.6	539.7	4	3,768
C. Individual Characteristics				
1. Terminal and port level:				
EDI *(in fraction of total sample)*	0.3			
Depth of water in meters	13.2	3.5	4.5	32.0
Number of liners calling at the terminal	16.2	14.5	1	114
Number of operators in port	3.7	2.6	1	10
Number of terminals in port	6.8	6.2	1	31
2. Operator group dummies (in fraction of total sample):				
Global Carrier	0.09			
Global Stevedore	0.15			
Other: not belong to any of above groups	0.76			
3. Country Characteristics:				
GDP in current US$ (billion) [c]	2,240	3,270	5.4	12,500
Goods exports in US$ (billion) [c]	271	249	0.4	972
Goods imports in US$ (billion) [c]	308	365	1.8	1,670
GDP per capita in current US$ [c]	18,654.9	12,367.8	405	37,651
4. Continental Distribution (in fraction of total sample):				
Asia	0.37			
Europe	0.27			
North America	0.17			
Latin America	0.06			
Oceania	0.09			
Africa	0.04			
Period	1997-2004			
Number of Countries	39			
Number of Ports	78			
Number of Terminal Operators	141			
Number of Terminals	397			
Number of Observations	597			

[a] An aggregate of (1) Quay cranes and (2) Ship shore container gantries.
[b] An aggregate of (1) Gantry cranes, (2) Yard cranes, (3) Yard gantries, (4) Reachstackers, (5) Yard tractors, (6) Yard chassis trailers, (7) Forklifts, (8) Straddle carriers, (9) Container lifters, and (10) Mobile cranes.
[c] The country data can be found at the World Bank website: http://devdata.worldbank.org/dataonline/old-default.htm

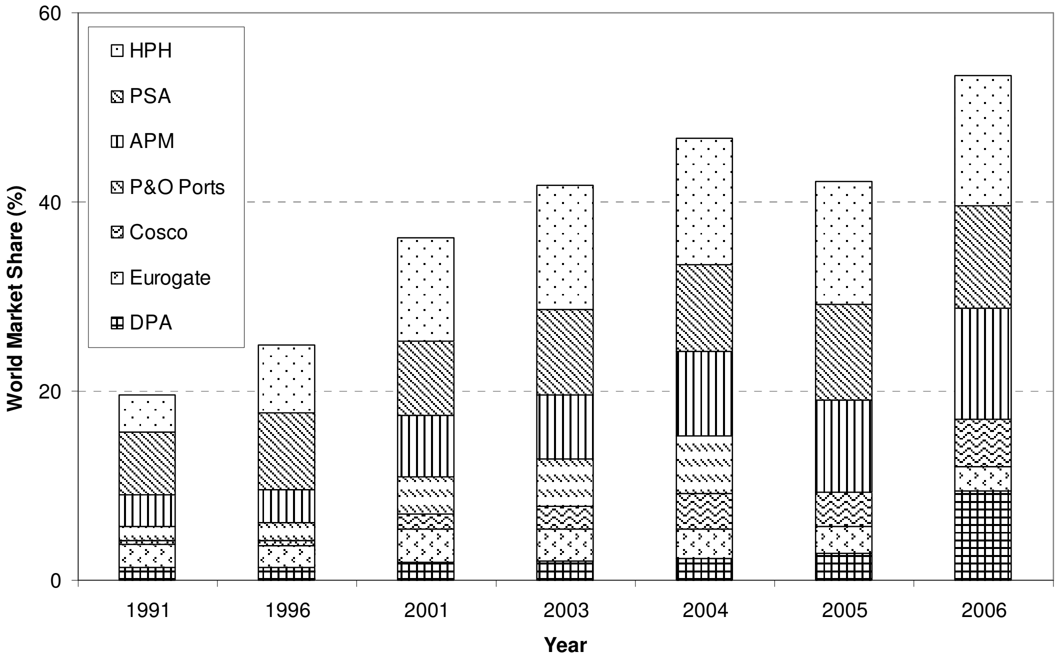
Table 2: Posterior means and standard deviations of coefficients of the Stochastic Production Frontier

Variable	Heterogeneous model	Conventional Model
1. Log inputs		
Quay superstructure (β_1)	0.1815 (0.0725)	0.2142 (0.0581)
Yard equipment (β_2)	0.0200 (0.0361)	0.0527 (0.0315)
Berth number (β_3)	0.0953 (0.0509)	0.0290 (0.0576)
Quay length (β_4)	0.0802 (0.0507)	0.1228 (0.0589)
Terminal area (β_5)	0.0268 (0.0582)	0.0912 (0.0444)
Storage capacity (β_6)	0.0087 (0.0229)	-0.0406 (0.0274)
Reefer points (β_7)	0.1400 (0.0355)	0.2125 (0.0320)
2. Individual intercept		
Constant (θ_0)	-0.7063 (0.3121)	0.8267 (0.3495)
2.1 Port characteristics		
Water depth (θ_1)	0.4184 (0.2770)	0.6644 (0.2626)
Ship calls (θ_2)	0.1322 (0.0369)	0.1550 (0.0399)
Number of operators (θ_3)	-0.0520 (0.0985)	-0.3789 (0.0761)
Number of terminals (θ_4)	-0.0219 (0.0886)	0.1992 (0.1039)
2.2 Country characteristics		
GDP (θ_5)	-0.3815 (0.0872)	0.2437 (0.0972)
Goods exports (θ_6)	0.0660 (0.1024)	-0.1715 (0.0719)
Goods imports (θ_7)	0.3088 (0.1114)	-0.0023 (0.0697)
GDP per capita (θ_8)	-0.7931 (0.2498)	-0.4889 (0.1123)
2.3 Operator Group		
Carrier (θ_9)	0.3594 (0.2344)	0.3780 (0.1626)
Stevedore (θ_{10})	0.4538 (0.1874)	1.6053 (0.3623)
2.4 Time trend		
Time (r_1)	-0.0719 (0.0870)	-0.5350 (0.2197)
Time squared (r_2)	0.1031 (0.0430)	0.1553 (0.0996)
3. Variance of the constant (σ_V^2)	0.4437 (0.0650)	
4. Inefficiency parameters		
Coeff. of Constant (g_1)	-1.8584 (0.3797)	0.3517 (0.1554)
Coeff. of Time (g_2)	-0.5967 (0.4785)	-0.4790 (0.1711)
Coeff. of Time squared (g_3)	-1.1674 (0.2398)	-0.0812 (0.0913)
Coeff. of Carrier (g_4)	-0.3961 (0.4313)	0.7357 (0.3945)
Coeff. of Stevedore (g_5)	-1.0024 (1.0824)	-0.1368 (0.2601)
Variances (Σ_{11})	2.2743 (0.7468)	0.3327 (0.0838)
Variances (Σ_{22})	2.0274 (0.8271)	0.2680 (0.0688)
Variances (Σ_{33})	1.4244 (0.4569)	0.1736 (0.0386)
Variances (Σ_{44})	1.5650 (0.9784)	0.7158 (0.4536)

Variances (Σ_{55})	1.9605 (1.1155)	0.5996 (0.2482)
5. Other parameters		
Variance of noise (σ_{ε}^{2})	0.0321 (0.0029)	0.0371 (0.0034)

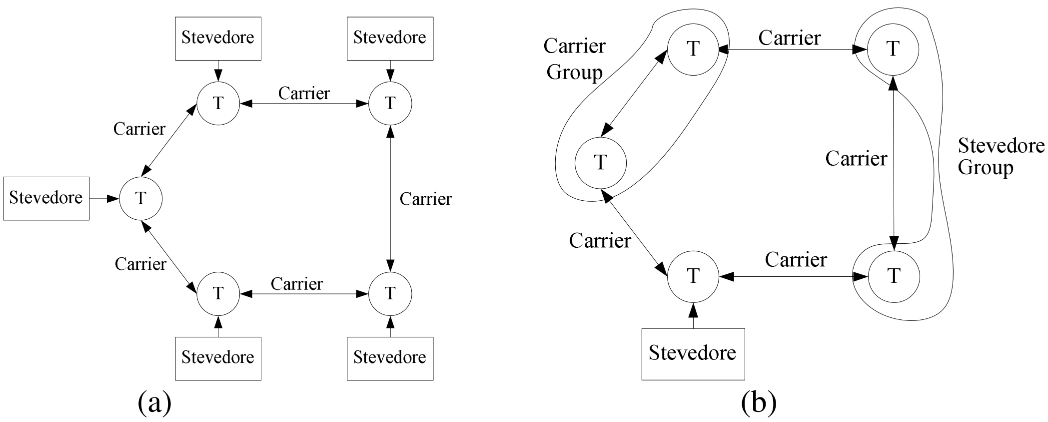
Numbers in parentheses are the posterior standard deviations.
All the input and output variables are normalised with respect to their sample means before taking log.

Figure 1: Market share of the leading terminal operator groups (2004 ranking)



Remarks: Dubai Port World acquired P&O Ports in 2006.
PSA acquired 20% of HPH in 2006.

Figure 2: Container terminals, stevedores, and carriers



Remarks: T in the circle means Terminal. (a) represents the original structure of port industry. The Stevedore handles the cargo on the terminal and Carrier transports them from Terminal to Terminal. (b) reveals the current tendency. Some Carrier Group can integrate the stevedore function. In the meanwhile, Stevedore Group purchase many terminals to operate. Eventually, there are Carrier Group, Stevedore Group and individual operators in terminal running.

Figure 3: Distributions of Prior and Posterior of variable β_1

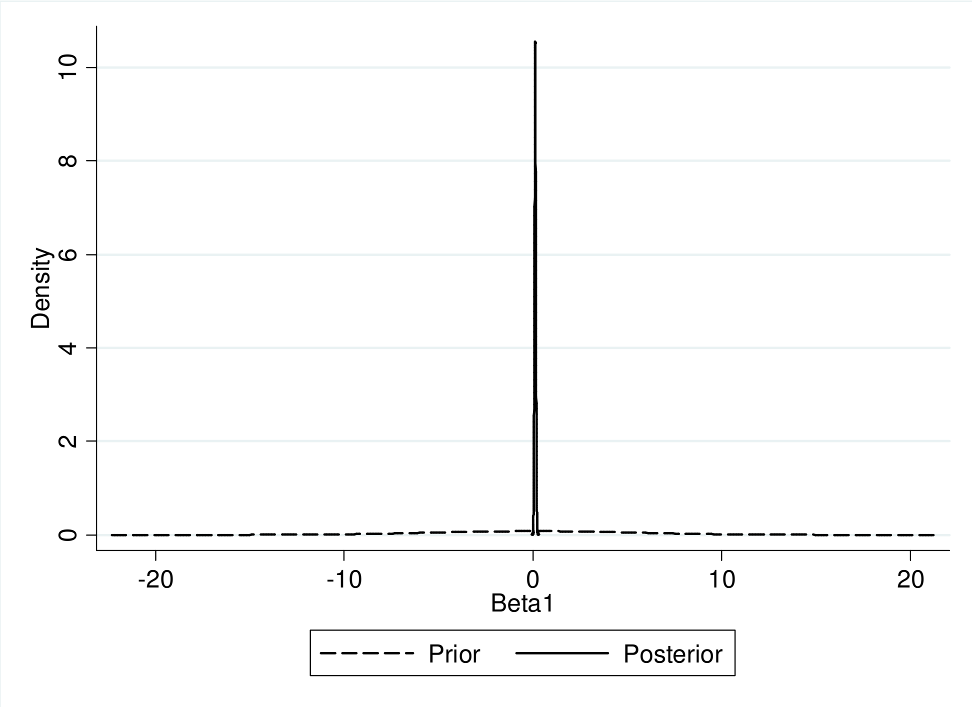


Figure 4: The converge analysis of variable β_1

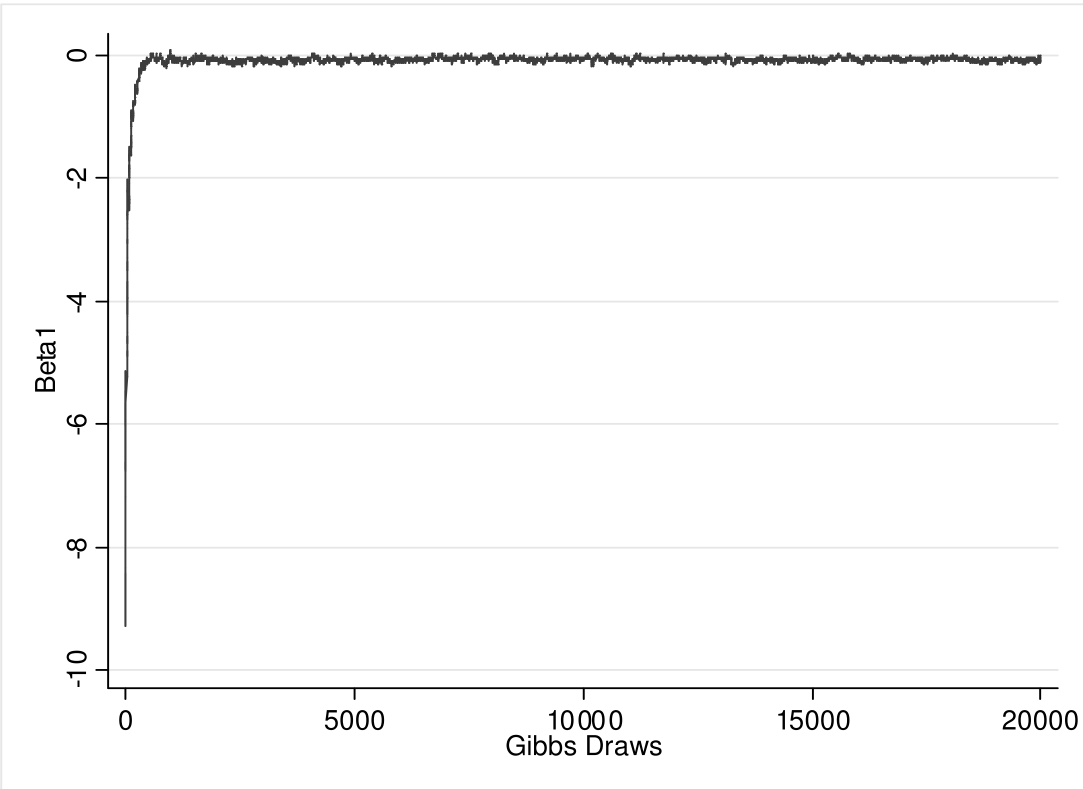


Figure 5: Conventional Stochastic Frontier Analysis Model

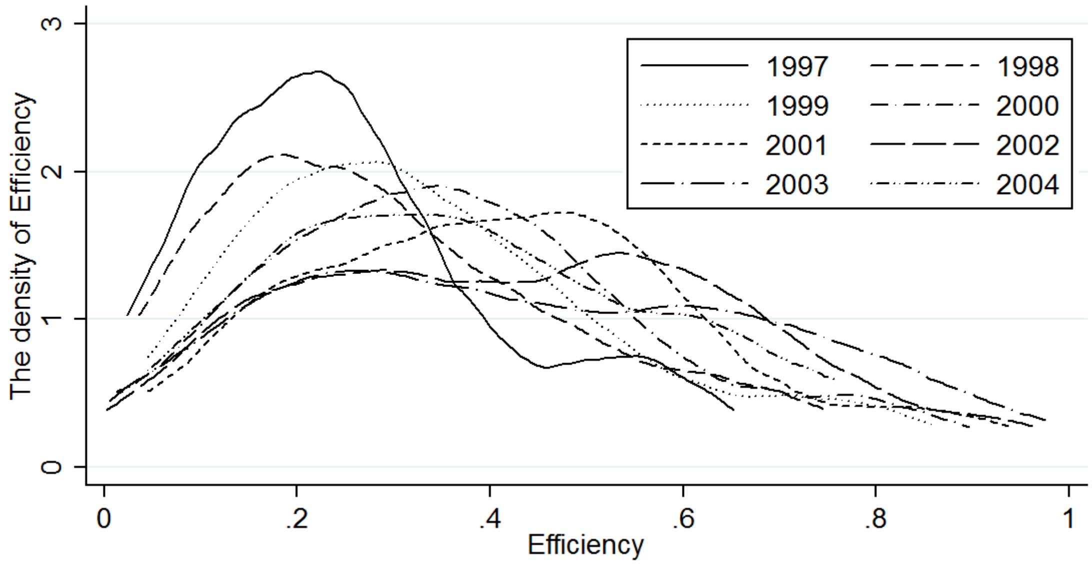


Figure 6: Heterogeneous Stochastic Frontier Analysis Model

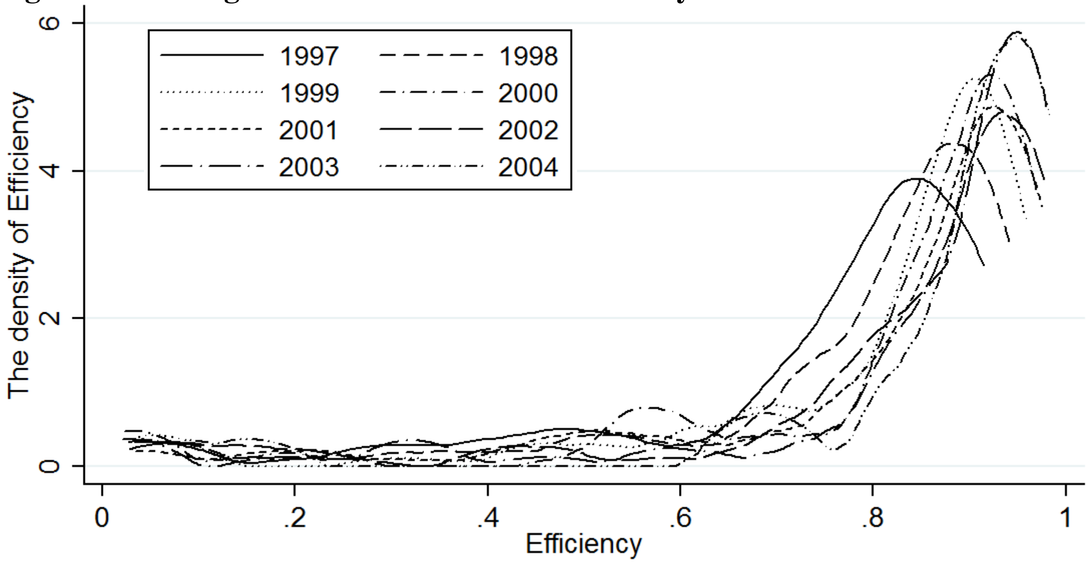


Figure 7: The distribution of random variable v_i

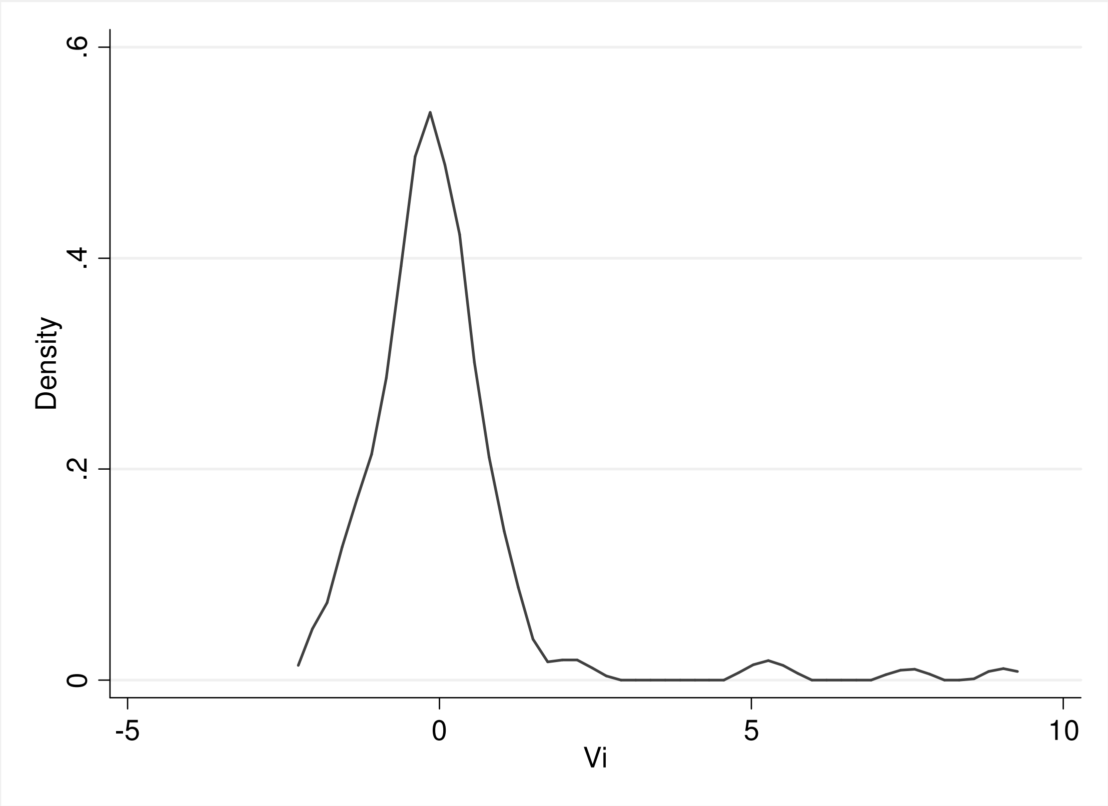
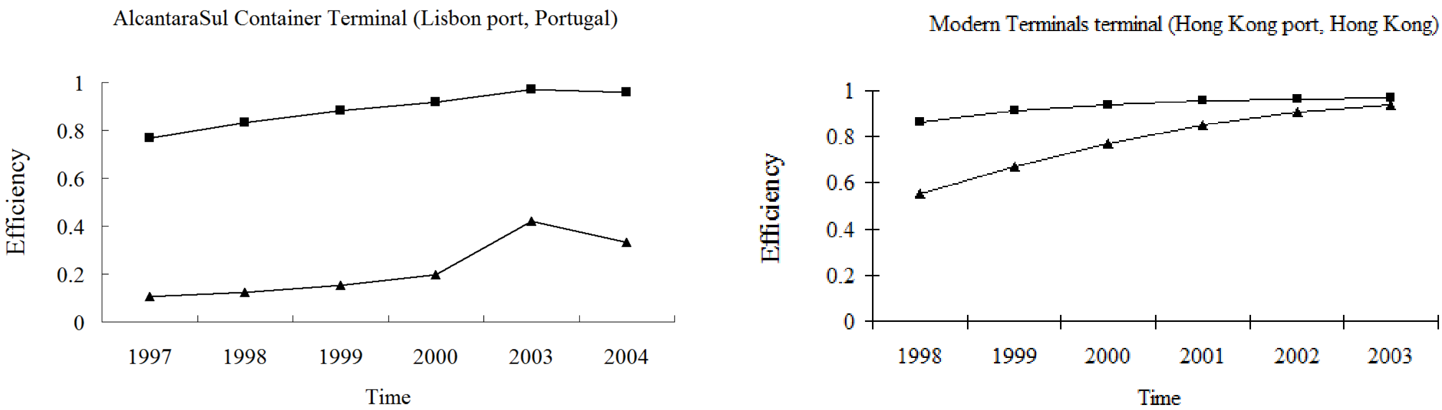
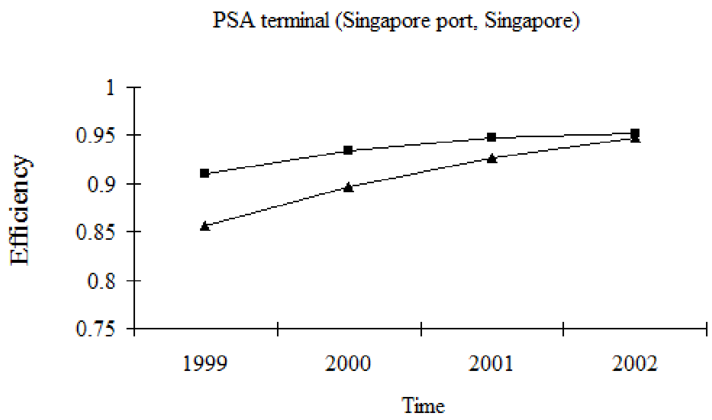
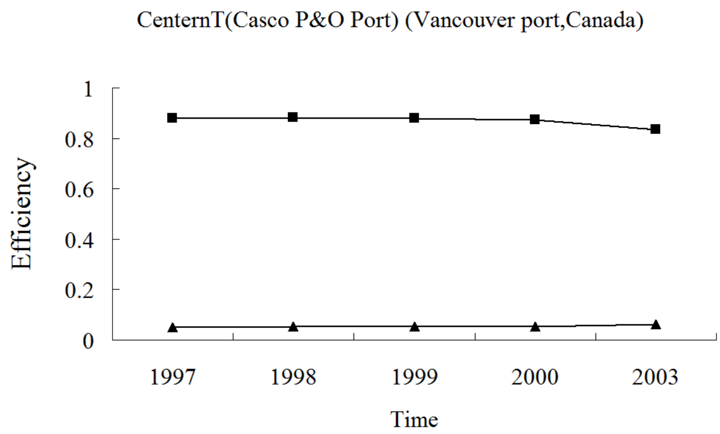


Figure 8: Examples of individual efficiency change of heterogeneous model and conventional model







Legend: square represent heterogeneous model, triangle represent conventional model