EJTIR

Issue 17(2), 2017
pp. 279-303
ISSN: 1567-7141
tlo.tbm.tudelft.nl/ejtir

Specialisation, diversification, size and technical efficiency in ports: an empirical analysis using frontier techniques

Beatriz Tovar[1]
Infrastructure and Transport Research Group (EIT), Department of Applied Economics, University of Las Palmas de Gran Canaria, Spain.

Alan Wall[2]
Oviedo Efficiency Group, Department of Economics, University of Oviedo, Spain.

This paper explores the relationship between output specialisation, diversification, size and technical efficiency in ports. Using a sample of Spanish port authorities observed over the period 1993-2012, we calculate a normalised Herfindahl-Hirschman index of overall specialisation and indices of relative specialisation in the individual cargoes. An output distance frontier is estimated using non-parametric Data Envelopment Analysis techniques to calculate technical efficiency. These efficiency scores are then used to test that relationship with a bootstrapped truncated regression. We find that both overall and relative output specialisation have a positive influence on technical efficiency. Moreover, the positive effects of specialisation on technical efficiency is reinforced for larger ports, which are in a better position to take advantage of economies of scale and which can also attract different types of cargo (enjoying also economies of scope) and therefore protect themselves from adverse demand conditions. Our results underline the trade-off for smaller port authorities between efficiency gains from specialisation and their vulnerability to market conditions for their main output.

Keywords: *data envelopment analysis, economies of scale and scope, output distance frontier, output specialisation, ports, technical efficiency, truncated regression.*

1. Introduction

In recent decades, port activity has changed substantially due to a combination of technological change, cargo organisation and the manner in which ports are being managed. In particular, new cargo-handling technologies capable of dealing with new vessel design have been developed with the consequence that productivity has increased due to mechanisation and a reduction in the use of labour that has translated into shorter stays of the ships at the port. Vessels themselves differ substantially depending on the type of maritime cargo they are transporting, which has in turn led to greater specialisation in the types of infrastructure and equipment needed to deal with these cargoes (Tovar de la Fé *et al*, 2004). Moreover, technological change has affected port infrastructure, with some types of infrastructure being highly specialised whereas others permit a greater degree of flexibility[3].

[1] Campus Universitario de Tafira, Modulo D, Las Palmas de Gran Canaria, 35017 Spain. T: +34928451794, E: beatriz.tovar@ulpgc.es
[2] Avenida del Cristo s/n, 33071 Spain. T:+34985104872, E: awall@uniovi.es

[3] Liquid bulk cargo in a port, for example, is often associated with a nearby refinery and requires very specific infrastructure leading, at least in the short run, to captive demand. On the other hand, infrastructure designed to move containers permits different types of merchandise to be handled (González-Laxe and Novo-Corti, 2012).

Output specialisation and diversification can provide advantages to ports. On the one hand, ports will benefit from 'specialisation efficiencies' when increased specialisation among outputs leads to higher technical efficiency.[4] These efficiency gains from specialisation may arise from a variety of sources, including the advantages to be gained from the division of labour and management of resources, the use of specialist skills, learning by doing and product-specific scale economies, and the use of equipment and infrastructure specifically designed to be adapted to the needs involved in each type of port traffic. On the other hand, diversification can yield competitive advantages for ports through economies of scope,[5] where common assets (and hence costs) can be spread across different types of products, and greater market power through cross-subsidisation of products which may deter market entry and permit the ports to compete more effectively in prices[6]. Moreover, ports that are specialised in a given output may be vulnerable to changing market demand (level or mix of outputs). Meyler *et al.* (2011) proposed that port authorities need to implement a "strategy of port activity diversification" in order to improve port performance in adverse market conditions and pointed out that the ports reporting satisfactory operational results in the recession period of 2009 were those with diversified cargo. Port activity diversification, which can be measured by the weight of various traffic categories in overall seaport traffic (Huybrechts *et al.*, 2002), can represent a strategy to reduce risk during adverse market conditions. In this sense, Ducruet *el al.* (2010) argue that commodity specialisation represents a weakness for ports facing a difficult economic environment as they may suffer badly if their main commodity cargo is particularly affected by adverse demand conditions.

However, it is important to emphasize that specialisation efficiencies are compatible with greater diversification of outputs, especially in the case of large ports. Thus, large ports may be highly diversified, handling different types of traffic, while at the same time having specialised infrastructure and staff to manage all these outputs and thereby take advantage of specialisation efficiencies. To identify specialisation efficiencies, it is therefore important to control for the size of the ports, as a specialisation index may indicate that a port is diversified rather than specialised despite the fact that its output in certain types of traffic may be large relative to the sector as a whole and that it may be enjoying economics of scale in those outputs.

In light of the above, we will be trying to answer two related but different questions. The first is: *Controlling for their size, are port authorities that are more concentrated in outputs (i.e., higher overall specialisation) more efficient in production?* The second question focuses on individual outputs: *Controlling for their size, are port authorities that are relatively more specialised in certain outputs (relative specialisation) more efficient in production?* To provide answers to these questions, we will calculate the efficiency scores and then run two separate sets of regressions, one for each question. Given that we will be controlling for size in the regressions, other questions can be answered. For example, we can check whether the effect of specialisation (overall or relative) on efficiency varies according to port size. Alternatively, given the degree of (overall or relative) specialisation, are larger ports more efficient?

To explore the relationship between specialisation (economies of scale), diversification (economies of scope), size and the productive efficiency of port authorities empirically, we use a sample of 26 Spanish port authorities observed over the period 1993-2012. To measure productive performance, we use frontier techniques to calculate technical inefficiency for the port authorities.

---

[4] The term 'specialisation efficiencies' was initially introduced in the agricultural economics literature by Coelli and Fleming (2004).

[5] Potential advantages of joint production (economies of scope) in the port sector has been analysed in the Spain port system context in Jara Diaz *et al*. (2005), Tovar *et al*. (2007); Jara Diaz *et al*. (2008), Tovar and Wall (2012).

[6] On the negative side, greater diversification may increase coordination costs and lead to bureaucratic distortions (Riordan and Williamson, 1985) and managers may not be able to give adequate attention to different product lines (Grant *et al*, 1988). Of course, these coordination problems may be alleviated if different activities are grouped at different terminals.

Concretely, we estimate an output-oriented distance frontier using Data Envelopment Analysis (DEA), a non-parametric method widely used in the frontier literature. However, to the best of our knowledge no study has investigated the possibility that the degree of specialisation of ports can affect their technical efficiency, and our paper aims to fill this gap in the literature.

Our results show that the degree of specialisation leads to gains in productive efficiency. Larger port authorities which can be specialised in several outputs are more efficient, but smaller ports also gain from concentrating their output. From a policy standpoint, in a context where the size of smaller ports cannot be increased, this implies a trade-off for smaller ports between benefitting from economies of scale in a given output, on the one hand, and the amount of exposure to changing market demand conditions, on the other hand, as higher exposure would make them more vulnerable.

The paper is structured as follows. In Section 2 we provide a brief overview of the literature on the themes of efficiency measurement and specialisation in ports. Section 3 describes the methodological approach used to estimate the relationship between specialisation and efficiency. The data we use to measure efficiency and specialisation are presented in Section 4. In Section 5, we present the empirical specification and our results and Section 6 concludes and offers some ideas for future research.

2. Technical efficiency and specialisation in ports: A brief review

The study of efficiency in the port sector using frontier techniques has received a great deal of attention in recent years (Cullinane and Song, 2003; Barros, 2005; Cullinane *et al.*, 2006; Trujillo and Tovar, 2007; Chang and Tovar, 2014ab; Tovar and Wall, 2015, 2016; Chang and Tovar, 2016, 2017a).

In a frontier framework, efficiency can be measured with econometric tools, namely Stochastic Frontier Analysis (SFA), or the non-parametric, linear-programming-based method of Data Envelopment Analysis (DEA). Surveys of efficiency studies in ports using these frontier methods can be found in González and Trujillo (2009) and Cullinane (2010), while recent studies by Schøyen and Odeck (2013) and Nguyen *et al.* (2016) contain comprehensive literature reviews of studies using DEA to measure port efficiency. These papers highlight that the majority of port efficiency studies have used DEA, due most likely to its flexibility in handling multiple inputs and outputs and lack of assumptions about production technology. This comes at the cost that standard DEA does not account for randomness in the data. Recent bootstrapping methods have, however, allowed DEA to account for randomness in data (Simar and Wilson, 1998, 2000; Simar, 2007). In the port literature, bootstrapped DEA methods have been used, among others, by Barros *et al* (2010) and Nguyen *et al* (2016).

Apart from just measuring different types of inefficiency, several papers have tried to *explain* efficiency. Previous studies have analysed the influence of certain environmental variables on port efficiency, such as ownership structure and/or size (Liu, 1995; Cullinane *et al*. 2002; Tongzon and Heng; 2005; Bang *et al.*, 2012; Yuen *et al.*, 2013), regulatory changes (Nuñez-Sánchez and Coto-Millán 2012; Rodríguez-Álvarez and Tovar, 2012), or demand variability (Rodríguez-Álvarez *et al*., 2011; Tovar and Wall, 2014). Among these, Bang *et al*. (2012), Yuen *et al*. (2013) and Chang and Tovar (2017b) use two-stage approaches similar to that used in this paper, where DEA is used to calculate efficiency scores and these scores are then incorporated as dependent variables in posterior regressions on a series of explanatory variables. However, while the literature on technical efficiency in ports and its determinants is relatively large, as far as we are aware there are no studies which have addressed the issue of specialisation efficiencies, where increased specialisation is a potential determinant of technical efficiency.

The relationship between size and diversification in ports has been analysed by Ducruet *et al.* (2010), who find that bigger ports are characterised by greater diversification of port activity. However, and as we have commented above, large ports may handle different types of traffic and yet be able to take advantage of specialisation efficiencies because the quantities of some of those outputs are large enough to permit this. This was highlighted by Martinez-Budria *et al.* (1999) for the case of Spain. These authors divided the Spanish port authorities into three groups using a complexity criterion incorporating port size and the composition of the output vector and found that highly complex ports show higher comparative efficiency levels. Similar results have been recently obtained by Tovar and Rodríguez-Déniz (2015) who found that grouping by complexity, size and traffic mix may be adequate when it comes to benchmarking the cost-efficiency of Spanish port authorities. Using hierarchical clustering they found one of the clusters to be characterized by large scales of production and by large quantities of all types of cargo, with these port authorities being able to take advantage of both economies of scale and scope.

We will consider the relationship between technical efficiency and two different types of specialisation, namely *overall specialisation* and *relative specialisation*. Overall specialisation is defined as the extent to which the port authority is concentrated in one or few outputs, regardless of the individual outputs involved. We measure this using the Herfindahl-Hirschman Index (HHI), which is the most commonly-used concentration index. This index has been used in studies to measure the concentration in the containerised shipping liner industry (Sys, 2009), the European port system (Notteboom, 1997, 2010) and the Spanish port system (Mateo-Mantecón et al., 2015). Relative specialisation, on the other hand, is a measure of whether the port is more specialised in a given individual output than the port system as a whole. To calculate this we use a so-called Bird Index, which is one of the most common measures in the literature (Frémont and Soppé, 2007). This index has been used to measure relative specialisation in the Spanish system by González-Laxe and Novo-Corti (2012). Some recent studies have analysed the degree of specialisation of Spanish ports and concluded that they are becoming more and more specialised in terms of their traffic as well as in the services they offer (González-Laxe, 2012; Reina and Villena, 2013).

3. Estimating the relationship between specialisation and efficiency: methodological approach

In a multi-output activity such as the port sector, distance functions are a natural tool with which to measure technical efficiency. Input-oriented or output-oriented distance functions can be used, with the former measuring the extent to which inputs can be radially (proportionally) contracted while maintaining the present level of output whereas the latter measure the extent to which outputs can be radially expanded for a given input endowment. The orientation chosen will depend on the ability of ports to adjust inputs and/or outputs. As noted by Cheon *et al.* (2010), if ports have little control over adjusting inputs then they should maximize outputs given the input level. This output-oriented approach has been followed by, among others, Cullinane *et al.* (2004); Trujillo and Tovar (2007), Cheon *et al.* (2010) and Chang and Tovar (2014a,b). Moreover, González and Trujillo (2008) justify this orientation based on an analysis of the conditions under which port authorities in Spain develop their activities. They argue that when it comes to the provision of infrastructure services, port authorities have some control over the production level through the use of commercial policies and concessions. Port authorities actively market their services and facilities and use tariff discounts, within legal limits, to attract new traffic. González and Trujillo (2008) also point out that as long as port authorities decide on the type of firm that can operate at the different ports, they are effectively deciding on the ships and goods that will be handled. In our empirical setup, we will therefore use an output distance function to measure technical efficiency.

To estimate the effect of specialisation on technical efficiency, we use a two-stage bootstrap truncated regression procedure proposed by Simar and Wilson (2007) where efficiency scores are regressed on a series of variables which are expected to influence efficiency. In the port literature, this procedure has been used by Yuen *et al*. (2013).

More formally, to measure efficiency in a multi-output setting where there are N inputs and M outputs, define the production set as

$$\mathcal{P} = \{(x, y) \in R_{+}^{N+M} | x \; can \; produce \; y\} \quad (1)$$

Then, for a production unit located at $(x_0, y_0) \in R_{+}^{N+M}$ an output measure of technical efficiency can be defined as:

$$\delta_0 = sup\{\delta | (x_0, \delta y_0) \in \mathcal{P}, \delta > 0\} \quad (2)$$

which is the reciprocal of the Shephard (1970) output distance function. Note that when $(x_0, y_0) \in \mathcal{P}$, $\delta \geq 1$.

The estimated efficiency scores for each unit i, $\hat{\delta}_i$, are obtained by DEA estimators with an output orientation. To allow for greater flexibility and the possibility that port authorities are not operating at optimal scale, we use both variable returns to scale (VRS) and constant returns to scale (CRS) specifications of the technology.

Having estimated technical efficiency through DEA, in a second stage these estimated DEA efficiency scores $(\hat{\delta}_i)$ are regressed on a set of potential covariates (z_i):

$$\hat{\delta}_i = z_i \beta + \varepsilon_i \geq 1 \quad (3)$$

The bootstrap procedure of Simar and Wilson (2007) allows us to obtain unbiased beta coefficients and valid confidence intervals. This procedure[7] involves regressing the estimated technical efficiency scores on the covariates using a bootstrapped truncated regression.

To analyse the relation between efficiency and the degree of output specialisation and size of port authorities, we need to include some measure of specialisation and size among the covariates in the second stage regression. A commonly-used measure of specialisation is the Herfindahl-Hirschman Index (HHI) or its normalised equivalent. (Al-Marhubi, 2000). While this index is typically encountered in the context of measuring concentration of firms in markets, we will apply it to measure the concentration of output shares for each port authority. This indicates the degree of overall specialisation, i.e., the degree to which the port authority's output is concentrated in a given output or outputs. We use the normalised version of the HHI. Following Al-Marhubi (2000), the normalised HHI for port i is defined as:

$$HHI_i = \frac{\sum_{m=1}^{M} s_{mi}^2 - \frac{1}{M}}{1 - \frac{1}{M}} \quad (4)$$

where $s_{mi} = \frac{y_{mi}}{\sum_{m=1}^{M} y_{mi}}$ with y_{mi} representing the quantity of cargo m produced by port i and M representing the total number of cargo services provided by the port. The index is normalised to take values ranging from 0 to 1, where a value of 1 represents perfect specialisation (i.e., the port authority manages only one type of cargo) and values closer to 0 represent greater diversification in the sense of a more uniform distribution of cargoes.

The HHI gives a measure of overall specialisation but does not tell us whether a port authority is relatively more specialised in a given output or subset of outputs. The question then arises as to whether greater specialisation in a given output (liquids, containers, solids, general merchandise) matters. To capture this we calculate an index of relative specialisation (Bird Index) for each of the cargo outputs for port authority i ($RELSPECy_{mi}$) defined as:

---

[7] We use algorithm #1 in their paper.

$$RELSPECy_{mi} = \frac{y_{mi}/Y_i}{y_{mSYS}/Y_{SYS}} \quad (5)$$

where y_{mi} is the total traffic of cargo m in port authority i, Y_i is total traffic of port authority i $Y_i = \sum_m y_{mi}$, y_{mSYS} is the total traffic of cargo m in the system ($y_{mSYS} = \sum_i y_{mi}$), and Y_{SYS} is the total traffic of the system ($Y_{SYS} = \sum_m y_{mSYS}$). This index of specialisation or polarization indicates the degree of specialisation *in a given cargo* compared to the degree of specialisation in that cargo of the system as a whole. Clearly, values greater (less) than 1 indicate higher (lower) relative specialisation of the port authority in that output.

Furthermore, whereas values of the HHI closer to 0 represent greater diversification across cargoes, this does not preclude ports taking advantage of specialisation efficiencies. The same goes for relative specialisation: having a similar or even lower relative specialisation index in a given output compared to the system as whole does not preclude specialisation economies for ports that are large enough.

Larger ports may have specialised infrastructure and staff for several outputs without their output being concentrated in one or few of these, and also without being relatively more specialised that sector as whole, so we also control for the effect of size in our regressions. To account for port size, we use two different measures. The first is a measure of the overall relative size of the port ($RELSIZE_i$), defined as the ratio of total port cargo output to total system port cargo in a given year:

$$RELSIZE_i = \frac{Y_i}{Y_{SYS}} \quad (6)$$

To measure the size or importance of the port in a certain output, we use the share of overall system output for a given type of cargo corresponding to the individual port ($NATSHAREy_{mi}$):

$$NATSHAREy_{mi} = \frac{y_{mi}}{y_{mSYS}} \quad (7)$$

Note that the $RELSPEC$, $RELSIZE$ and $NATSHARE$ indices are related. In particular, from (5), (6) and (7) it can be shown that

$$NATSHAREy_{mi} = RELSPECy_{mi} * RELSIZE_i \quad (8)$$

4. Data

The data we use corresponds to the Spanish port system. In Spain, ports play a crucial economic role and are quite varied in terms of their size and specialisation, including small, medium-sized and large ports, ports that serve their hinterland and thus act as gateways, and others which serve as hubs. Our sample comprises 520 observations and is made up of a panel data set of 26 port authorities observed over the period 1993-2012.[8] The main sources of this information are the Spanish Public State Ports Body (EPPE), which publishes accounts and management reports, and the port authorities, which provide information in their annual reports and on their websites. The port authorities in the sample vary widely in terms of size and specialisation, with some managing ports whose activity involves cargo and passenger traffic, whereas others run ports whose main activity is cargo with almost no passenger transport. The fact that the ports under consideration are in the same country also has the advantage that the accounting data used are uniform and comparable. Moreover, these ports face the same regulations and operate in a very similar environment.

[8] The port authorities included are A Coruña, Alicante, Avilés, Bahía de Algeciras, Bahía de Cádiz, Baleares, Barcelona, Bilbao, Cartagena, Castellón, Ceuta, Huelva, Las Palmas, Málaga, Marín y Ría de Pontevedra, Melilla, Motril, Pasajes, Sta. Cruz de Tenerife, Santander, Sevilla, Tarragona, Valencia, Vigo and Vilagarcía.

In order to estimate the distance function technology for the ports in our sample, we need information on their outputs and inputs. Regarding outputs, port activity is multi-product and port infrastructure service provision may be viewed in terms of the goods handled and the passengers using the port. The amounts of different types of outputs for each of the sampled port authorities are known by type: tonnes of liquid bulk (y_1), tonnes of solid bulk (y_2), tonnes of containerised goods (y_3), tonnes of general non-container merchandise (y_4), and number of passengers (y_5).[9] The inputs used are number of labour units (x_1); intermediate consumption expenditures measured in euro (x_2); capital assets, including the port authority's capital assets, measured in euro (x_3); and the squared metres of deposit surface area (x_4).

As explanatory variables of port inefficiency, we use the specialisation/diversification and size indices described above, namely the normalised HHI and the relative specialisation indices. As passenger traffic is a minor activity for most ports, and in order to have common units of measurement for outputs, we focus on specialisation of cargo traffic only, so that that only outputs y_1-y_4 will be used to calculate the HHI and relative specialisation indices. While passenger transport is not included in the HHI, we control for port authorities having passenger traffic by including an indicator variable which captures the port authority has had passenger traffic in a given year (*DUMMYPASS*).

The average normalised HHI for each port authority over the sample period is presented in Figure 1. The average scores range from 0.0845 for Barcelona, which is the port authority with the most uniform distribution (the least specialised) of cargoes, to 0.701 for Gijón, which has the greatest overall specialisation (its output being highly concentrated in solid bulk).

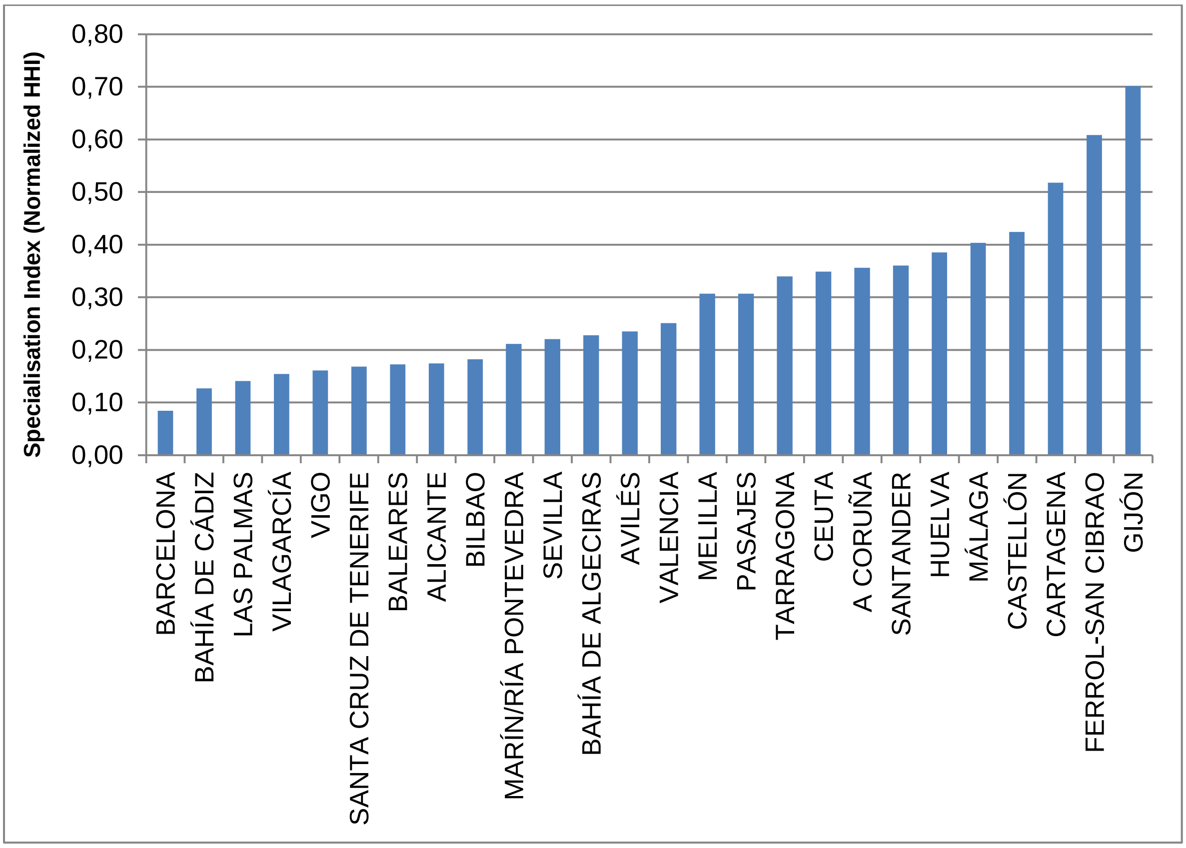

Figure 1. Average normalised HHI index by Port Authority: 1993-2012

[9] The cargo outputs could alternatively have been expressed in volume (TEU) terms. This had no effect on the results of our empirical analysis.

To gain an insight into the changes in the level of output specialisation over the sample period for the system as a whole, Figure 2 shows the evolution of the average output-weighted normalised HHI.[10] In the early years (1993-1999) there was a tendency toward more diversification. Since then, however, there has been a trend towards more specialisation of outputs within port authorities, and this process appears to have intensified since 2007.

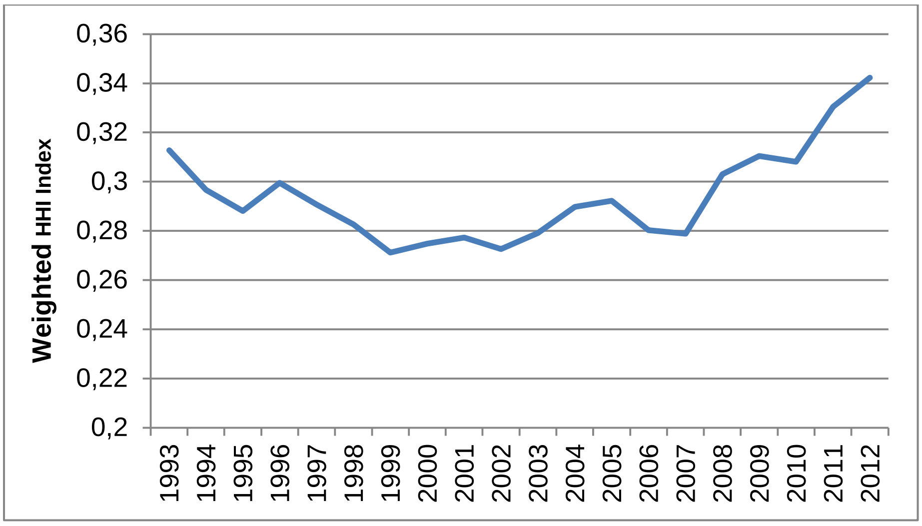

Figure 2. Evolution of overall output-weighted normalized HHI index: 1993-2012

The final issue that needs to be dealt with is size. The sample comprises port authorities of all sizes, and in particular contains some very large ports: These large ports, such as Barcelona, may not be very specialised in the sense that their cargos may account for more or less equal shares of their total output, yet because of their size they may have specialised infrastructure for each type of cargo[11]. While such ports may not appear to be specialised on the basis of an index such as the HHI, in reality they can be taking advantage of specialised infrastructure. To illustrate this, Figure 3 below show the shares of port authority output in total sector (i.e., national) output for each type of cargo, where it is clear that some port authorities have significant shares in several outputs.

Comparing the shares of the port authorities' individual cargoes with the HHI specialisation indices is instructive. Barcelona, for example, appears as the most diversified port authority as it has the smallest HHI score, yet it accounts for an average of 12% and 17% of Spanish general merchandise and container cargo respectively, and for over 5% of system output in the remaining two cargoes. Similarly, Bilbao has a HHI score in the second quartile and hence is relatively diversified according to this index, yet accounts for over 5% of system output in each of the four cargoes. As such, these port authorities are not specialised or concentrated in any particular cargo, but have specialised infrastructure in each and can thus be expected to gain from scale efficiencies. That is, the larger the port authority, the less concentrated it is overall but the greater the possibility of being able to take advantage of increased scale in one or more output. We control for the possible effect of size on efficiency by including the variables *NATSHARE* and *RELSIZE* defined above.

---

[10] That is, for each year t, the output-weighted HHI is $\sum_i\{(Y_i/Y_{SYS}) * HHI_i\}$, where Y_i is the total output of port i and Y_{SYS} is the total output of all ports in the system.

[11] Thus, Barcelona's port has a land surface area of 829ha, and has more than 20km of berths and docks. This space includes 35 specialised terminals, comprising 3 for ferries, 7 for cruises, 4 for containers, 2 for automobiles, 1 for fruit, 2 that are specialised in coffee, cacao and non-ferrous metals, 1 freezer facility, 9 for liquid bulk, and 6 for solid bulk.

Some descriptive statistics of the production variables (inputs and outputs) used to calculate the DEA efficiency scores and the covariates used to explain these efficiency scores are shown in Table 1.

Table 1. Descriptive statistics of variables

Variable	Description	Mean	Std. Dev.
Outputs and inputs used to calculate efficiency scores			
y_1	Liquid bulk cargo (tons)	5,129,363	6,481,987
y_2	Solid bulk cargo (tons)	3,200,887	3,378,799
y_3	Container cargo (tons)	3,540,684	8,429,694
y_4	General non-container cargo (tons)	1,709,522	1,912,089
y_5	Passengers (units)	858,984	1,477,287
x_1	Labour (units)	210	110
x_2	Supplies (€ deflated)	8,136,485	8,260,716
x_3	Capital assets (mill. € deflated)	364.375	323.745
x_4	Deposit surface area (m^2)	808,426	1,002,453
Variables used to explain efficiency scores			
HHI	Normalised Herfindahl-Hirschman Index	0.292	0.171
NATSHAREy_1	Output y_1 of port i/Output y_1 of all ports	0.038	0.048
NATSHAREy_2	Output y_2 of port i/Output y_2 of all ports	0.038	0.040
NATSHAREy_3	Output y_3 of port i/Output y_3 of all ports	0.038	0.078
NATSHAREy_4	Output y_4 of port i/Output y_4 of all ports	0.038	0.039
RELSPECy_1	Relative specialisation in output y_1	0.769	0.687
RELSPECy_2	Relative specialisation in output y_2	1.391	1.061
RELSPECy_3	Relative specialisation in output y_3	0.691	0.787
RELSPECy_4	Relative specialisation in output y_4	1.623	1.361
RELSIZE	Total cargo of port i/Total cargo of all ports	0.038	0.037
DUMMYPASS	Dummy if port has passenger traffic	0.894	0.308

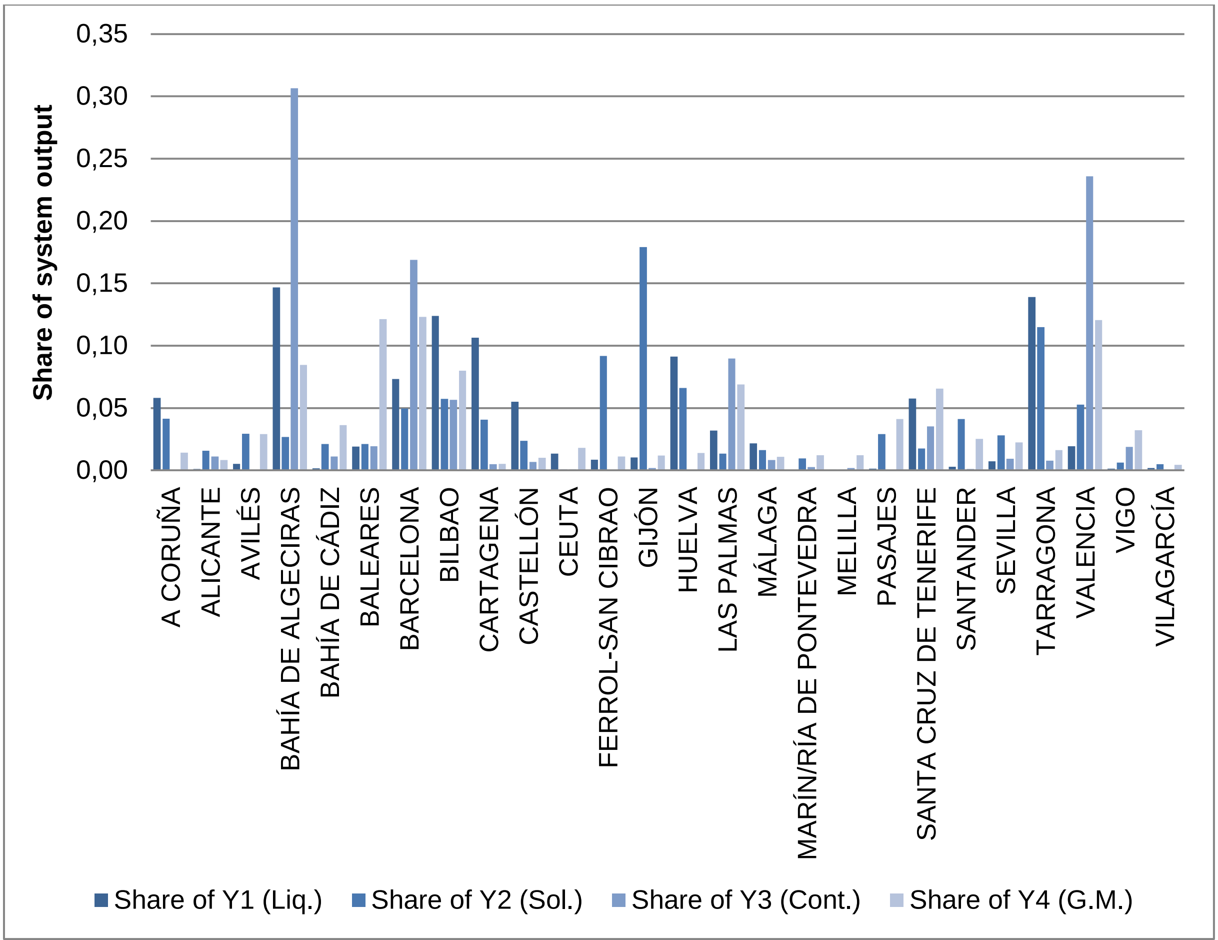

Figure 3. Shares of system output for each type of cargo by Port Authority: 1993-2012

5. Empirical specification and results

Our main model uses output-oriented DEA scores calculated from a pooled model, using both constant returns to scale (CRS and variable returns to scale (VRS). The averages of the pooled DEA efficiency scores for the entire sample were 0.63 and 0.69 for the CRS and VRS models respectively. Along the time dimension, in the CRS model the average efficiency scores went from 0.58 in 1993 to 0.63 in 2012, with a minimum of 0.54 in 2009 (coinciding with the economic crisis)[12] and a maximum of 0.70 in 2006 and 2007. For the VRS model, the scores went form 0.66 in 1993 to 0.68 in 2012, with a minimum of 0.59 in 2009 and a maximum of 0.75 in 2006 and 2007. The evolution of the scores over time for both models is illustrated in Figure 4, and a comparison of average scores by port for the CRS and VRS models is presented in Figure 5. These results indicate that port authorities have the potential to produce considerably more output with existing resources were the demand to exist.

[12] These results in the context of an international economic crisis are comparable to results reported by Estache *et al*. (2004) when analyzing the total factor productivity and efficiency evolution of Mexican ports during the East Asian Crisis, and Wilmsmeier *et al*. (2013) and Tovar and Chang (2014a,b) regarding the effect of the financial crisis on container port productivity in several countries in Latin America and the Caribbean. These constituted negative exogenous shocks to demand.

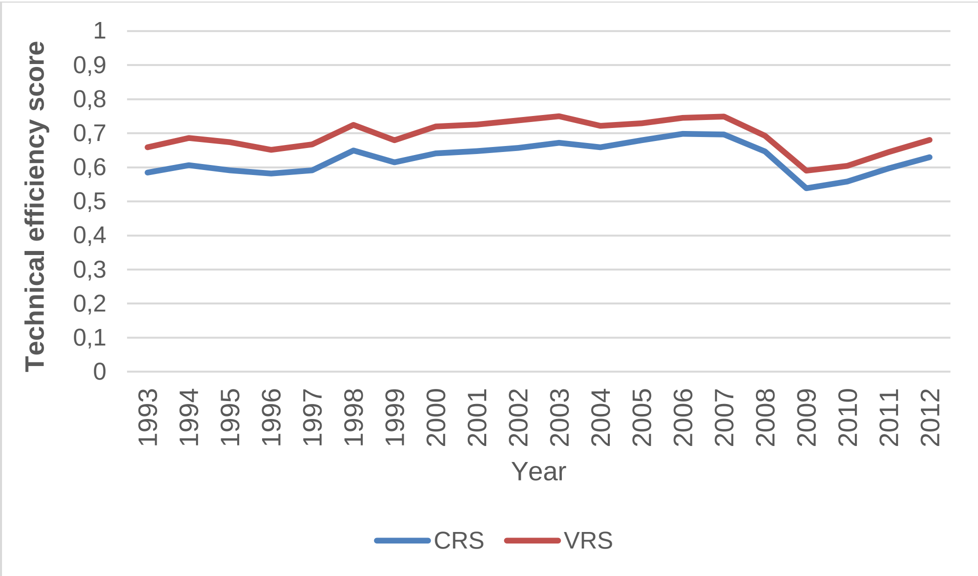

Figure 4. DEA efficiency scores for CRS and VRS models: 1993-2012

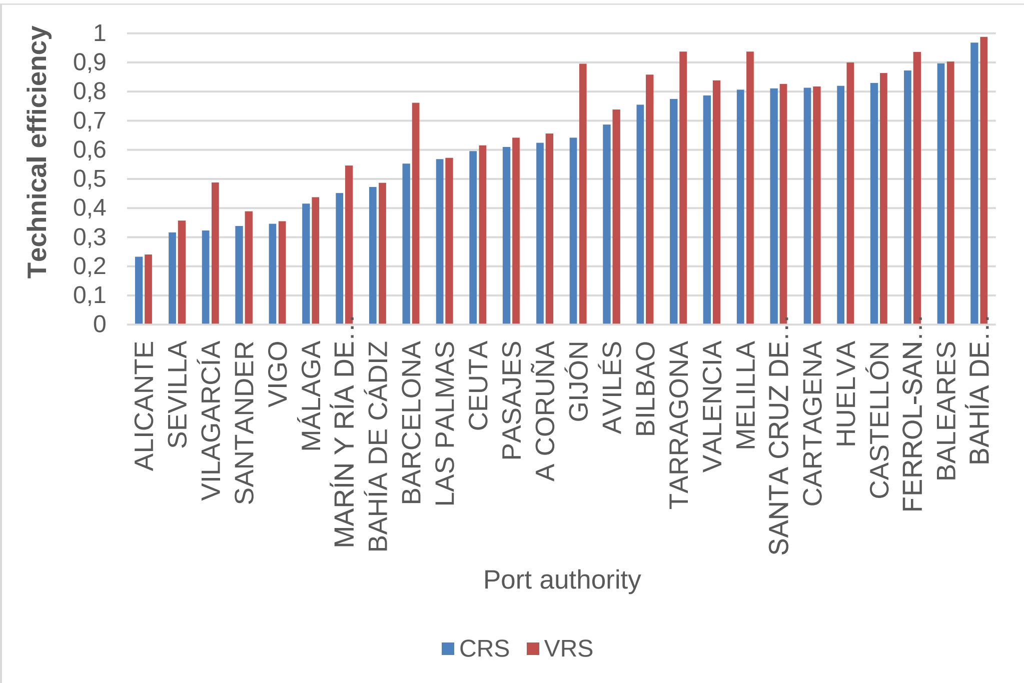

Figure 5. Average DEA efficiency scores for CRS and VRS by port authority

To visualize the relation between output specialisation and efficiency, Figures 6 and 7 present scatterplots of the average efficiency scores and the average overall specialisation indices for the port authorities for the CRS and VRS models, where the variables are expressed in deviation from their means. The figures seem to illustrate a positive relation between specialisation and efficiency, with many observations located in the south-west and north-east quadrants.

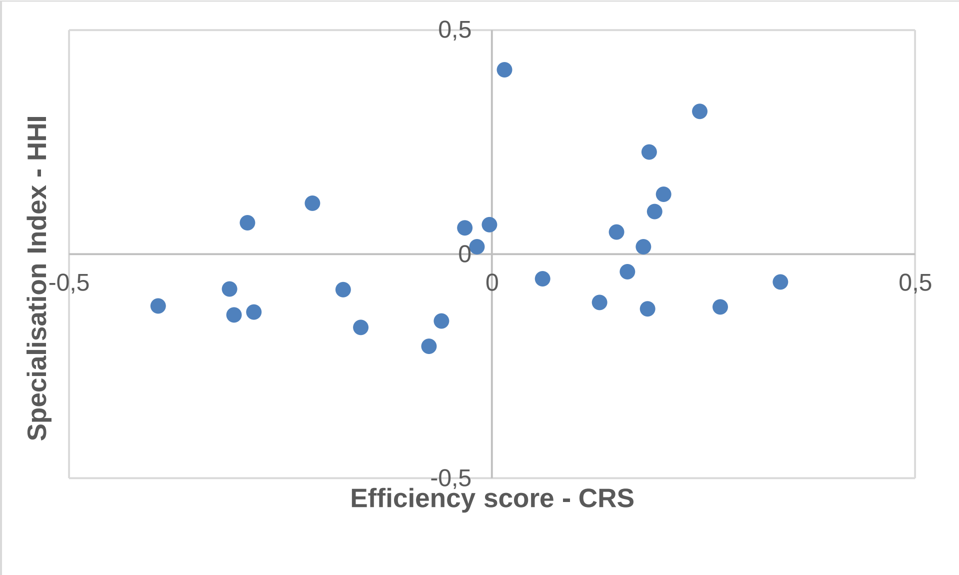

Figure 6. Scatterplot of Efficiency Scores-CRS vs. HHI

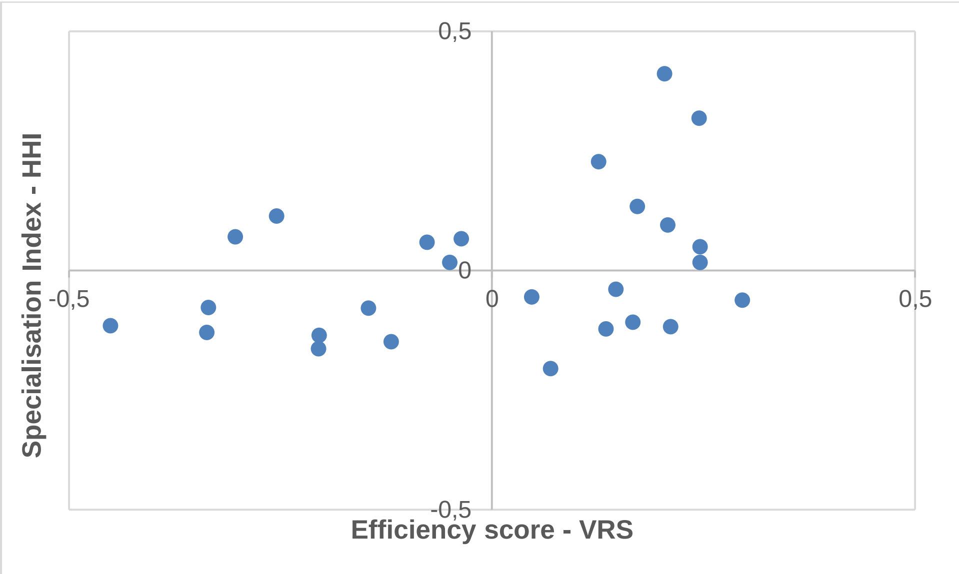

Figure 7. Scatterplot of Efficiency Scores-VRS vs. HHI

As an initial test of the relationship, we calculated the Spearman rank correlation coefficient between the efficiency scores and the HHI. The value of the coefficient was 0.343 and was significant well below the 1% level, confirming the positive correlation between efficiency and output specialisation.

To more formally test the impact of specialisation on efficiency, we carry out the bootstrapped truncated regression of the DEA efficiency scores following Simar and Wilson (2007). To capture the effects of overall specialisation and relative specialisation, we specify two separate regressions.

Beginning with overall specialisation, as measured by the concentration index (*HHI*), we specify the following regression model:

(9)

$$\hat{\delta}_i = \alpha_i + \beta HHI + \sum_{m=1}^{4} \gamma_m NATSHAREy_m + \sum_{m=1}^{4} \varphi HHI * NATSHAREy_m + \theta DUMMYPASS + \sum_{t=1994}^{2012} \beta_t D_t + \varepsilon_i \quad (9)$$

where we include time dummy variables to capture the evolution of efficiency over time. We carried out 1,000 bootstrap replicates of the truncated regression part of the procedure outlined in Algorithm #1 of Simar and Wilson (2007). The results of the bootstrapped regression are displayed in Tables 2 and 3 for the CRS and VRS models, and 95% confidence intervals are presented for the estimated coefficients. For the sake of space, we do not report the estimates of the individual effects or the time dummies.

Table 2. Efficiency and overall specialisation (HHI) - CRS model

Variable	Coefficient	95% Bootstrap Confidence Intervals	
		Low	*High*
HHI	1.303	-1.300	4.206
$NATSHAREy_1$	4.105	-30.369	41.234
$NATSHAREy_2$	-11.613	-43.417	12.852
$NATSHAREy_3$	-13.869	-43.651	18.392
$NATSHAREy_4$	-45.292*	-66.287	-21.027
$HHI*NATSHAREy_1$	-137.480*	-201.996	-60.124
$HHI*NATSHAREy_2$	-4.442*	-50.233	41.296
$HHI*NATSHAREy_3$	-17.767	-102.421	82.441
$HHI*NATSHAREy_4$	-16.291	-125.607	89.612
DUMMYPASS	0.104	-0.715	0.737
D_{1994}	-0.644	-1.297	0.154
D_{1995}	-0.415	-1.078	0.357
D_{1996}	-0.485	-1.190	0.261
D_{1997}	-0.719	-1.356	0.061
D_{1998}	-1.361*	-2.025	-0.501
D_{1999}	-1.260*	-1.944	-0.467
D_{2000}	-1.55*1	-2.202	-0.705
D_{2001}	-1.590*	-2.220	-0.644
D_{2002}	-1.731*	-2.391	-0.858
D_{2003}	-1.873*	-2.560	-0.958
D_{2004}	-1.70*2	-2.347	-0.805
D_{2005}	-1.867*	-2.484	-0.917
D_{2006}	-2.181*	-2.872	-1.179
D_{2007}	-2.209*	-2.917	-1.182
D_{2008}	-1.592*	-2.297	-0.750
D_{2009}	-0.201	-0.950	0.510
D_{2010}	0.089	-0.646	0.768
D_{2011}	-0.761	-1.503	0.007
D_{20012}	-1.198*	-1.955	-0.445
Observations	520		

* Significant at 5% level

Table 3. Efficiency and overall specialisation (HHI) - VRS model

Variable	Coefficient	95% Bootstrap Confidence Intervals	
		Low	*High*
HHI	-1.142	-3.467	1.515
NATSHAREy_1	9.326	-15.727	39.788
NATSHAREy_2	-41.436*	-69.942	-11.597
NATSHAREy_3	3.983	-26.168	37.673
NATSHAREy_4	-55.484*	-78.823	-31.672
*HHI*NATSHARE*y_1	-109.766*	-163.315	-51.976
*HHI*NATSHARE*y_2	52.236	-14.615	103.401
*HHI*NATSHARE*y_3	-61.583	-150.233	36.626
*HHI*NATSHARE*y_4	-6.098	-100.812	91.129
DUMMYPASS	0.683	-0.001	1.137
D_{1994}	-0.297	-0.881	0.365
D_{1995}	0.028	-0.555	0.651
D_{1996}	0.032	-0.554	0.639
D_{1997}	0.023	-0.520	0.679
D_{1998}	-0.561	-1.126	0.149
D_{1999}	-0.360	-0.932	0.291
D_{2000}	-0.468	-1.011	0.207
D_{2001}	-0.530	-1.067	0.180
D_{2002}	-0.653	-1.228	0.012
D_{2003}	-0.765*	-1.310	-0.081
D_{2004}	-0.830*	-1.357	-0.124
D_{2005}	-0.842*	-1.379	-0.126
D_{2006}	-1.083*	-1.684	-0.304
D_{2007}	-1.067*	-1.675	-0.319
D_{2008}	-0.595	-1.224	0.080
D_{2009}	0.776*	0.150	1.321
D_{2010}	1.014*	0.375	1.576
D_{2011}	0.211	-0.374	0.837
D_{20012}	-0.354	-0.997	0.251
Observations	520		

* Significant at 5% level

The coefficients on the HHI and NATSHARE variables will be used to calculate the marginal effects of these variables. Before turning to that, some comments on the time dummy variables are in order. The estimated coefficients show that efficiency was increasing over time until 2008. In 2008 we see a sharp drop in efficiency which continues for the following three years, reflecting the impact of the recent economic crisis.

As previously mentioned, to assess the impact of the overall specialisation on technical efficiency, we calculate the overall marginal effects of HHI based on the coefficient estimates. These are reported at the top of Table 4, where it can be seen that the HHI is negative and statistically significant for both the CRS and VRS models.[13] The negative sign on the marginal effect implies that HHI is associated with greater technical efficiency (as technical inefficiency is lower). This result for overall specialisation measured by the HHI concentration index provides an affirmative answer to our first question: *controlling for their size, ports that are more concentrated in outputs are more productively efficient.*

[13] Table 4 contains results for a series of different DEA models. We will discuss these at the end of the section.

Table 4. Overall marginal effects of specialisation indices

Model	Variable	*CRS*		*VRS*	
		Mean	*S.E.*	*Mean*	*S.E.*
Pooled	HHI	-3.761*	1.275	-3.703*	1.034
Yearly	HHI	0.378	0.798	-2.241*	0.605
Window (5-year)	HHI	0.480	0.818	-1.260	0.680
Sequential	HHI	0.480	0.818	-2.180	1.159
Pooled-IV	HHI	-4.742*	1.698	-5.540*	1.721
Pooled: Input-Oriented	HHI	-	-	-1.783*	0.306
Pooled	$RELSPEC_{y1}$	-2.729*	0.471	-1.511*	0.467
	$RELSPEC_{y2}$	-0.152*	0.382	-0.447	0.401
	$RELSPEC_{y3}$	-0.851	0.333	-0.379	0.372
	$RELSPEC_{y4}$	-1.707*	0.286	-1.488*	0.305
	RELSIZE	-134.746*	21.057	-167.332*	23.614
Yearly	$RELSPEC_{y1}$	-2.033*	0.332	-0.775*	0.291
	$RELSPEC_{y2}$	-0.753*	0.216	-0.676*	0.198
	$RELSPEC_{y3}$	-0.763*	0.216	-0.168	0.213
	$RELSPEC_{y4}$	-1.610*	0.187	-1.624*	0.158
	RELSIZE	-124.880*	16.057	-212.979*	18.948
Window (5-year)	$RELSPEC_{y1}$	-2.417*	0.301	-1.016*	0.273
	$RELSPEC_{y2}$	-0.796*	0.199	-0.715*	0.190
	$RELSPEC_{y3}$	-1.030*	0.210	-0.302	0.197
	$RELSPEC_{y4}$	-1.624*	0.177	-1.518*	0.166
	RELSIZE	-146.226*	15.178	-213.114*	20.382
Sequential	$RELSPEC_{y1}$	-2.417*	0.301	-0.600	0.313
	$RELSPEC_{y2}$	-0.796*	0.199	-0.458*	0.219
	$RELSPEC_{y3}$	-1.030*	0.210	-0.608*	0.293
	$RELSPEC_{y4}$	-1.624*	0.177	-1.194*	0.199
	RELSIZE	-146.227*	15.178	-225.003*	27.640
Pooled-IV	$RELSPEC_{y1}$	-2.327*	0.474	-0.974*	0.438
	$RELSPEC_{y2}$	0.465	0.364	0.071	0.354
	$RELSPEC_{y3}$	-0.853*	0.357	-0.182*	0.356
	$RELSPEC_{y4}$	-1.502*	0.268	-1.066*	0.248
	RELSIZE	-135.650*	21.601	-147.118*	23.782
Pooled: Input-Oriented	$RELSPEC_{y1}$	-	-	-0.460*	0.117
	$RELSPEC_{y2}$	-	-	0.345*	0.096
	$RELSPEC_{y3}$	-	-	0.047	0.097
	$RELSPEC_{y4}$	-	-	-0.503*	0.074
	RELSIZE	-	-	-31.864*	4.692

* Significant at 5% level

The interaction terms between the *HHI* and the *NATSHARE* variables shed further light on the relationship between overall specialisation and size. Care must be taken when interpreting the coefficients of these interaction terms in Tables 2 and 3 given the non-linear nature of the truncated regression (Ai and Norton, 2003; Karaca-Mandic *et al*. 2012). Thus, in Table 5 we present the derivatives of the interaction terms of *HHI* with the different *NATSHARE* variables, where the cross-derivatives of *HHI* are evaluated at the 25th, 50th, 75th and 95th percentiles of the *NATSHARE* variables.

Table 5. Marginal effects of overall specialization interaction terms

		CRS		VRS	
Variable	Percentile	Marginal	SE	Marginal	SE
$NATSHARE_{Y1}$	25	-0.247	1.413	-1.662	1.223
	50	-1.446	1.323	-2.624*	1.165
	75	-8.507*	2.354	-8.286*	2.001
	95	-18.632*	5.162	-16.405*	4.255
$NATSHARE_{Y2}$	25	-4.571*	1.458	-6.109*	1.226
	50	-4.620*	1.460	-5.526*	1.210
	75	-4.709*	1.635	-4.457*	1.527
	95	-5.020	3.122	-0.764	3.831
$NATSHARE_{Y3}$	25	-4.222	2.191	-3.090	1.919
	50	-4.290*	2.048	-3.382	1.775
	75	-4.548*	1.633	-4.494	1.381
	95	-7.474	8.406	-17.071*	8.531
$NATSHARE_{Y4}$	25	-4.165*	1.583	-4.779*	1.489
	50	-4.337*	1.398	-4.849*	1.290
	75	-5.063*	2.308	-5.145*	1.904
	95	-6.343	5.808	-5.667	4.901

* Significant at 5% level.

As can be seen from Table 5, the cross-derivatives are negative for all percentiles for each of the four *NATSHARE* variables for both the CRS and VRS specifications, with the majority being significant. This implies that for each of the outputs, the larger the size of the output that a port authority has, the greater the overall specialisation efficiencies.

We now turn to relative specialisation. For this, we re-specify the model in (9) as follows:

$$\hat{\delta}_i = \alpha_i + \sum_{m=1}^{4} \beta_m RELSPECy_m + \gamma RELSIZE + \sum_{m=1}^{4} \varphi RELSPECy_m * RELSIZE + \theta DUMMYPASS + \sum_{t=1994}^{2012} \beta_t D_t + \varepsilon_i \quad (10)$$

Given the relationship between the variables *RELSPEC*, *RELSIZE* and *NATSHARE* in (8), it can be seen that the basic difference between this model (10) and the previous model (9) is that we have substituted the overall specialisation index (*HHI*) with the four indices of relative specialisation (Bird indices). As in the previous model, size is controlled for and the model is estimated with individual effects and time dummy variables. The results from the bootstrapped truncated regressions are presented in Tables 6 and 7 for the CRS and VRS models.

Based on the estimations in Tables 6 and 7, the overall marginal effects of the relative specialisation variables are reported in Table 4. All values of the marginal effects are negative in both the CRS and VRS models. The *RELSPEC* indices are negative and significant for liquids (y_1) and general merchandise (y_4) in both models, indicating that port authorities that are relatively specialised in these cargoes tend to be more efficient. Of the remaining values, the index for solids (y_2) is significant in the CRS model. We also report the marginal effects of the size variable (*RELSIZE*). These are negative and significant in both the CRS and VRS specifications, showing that larger ports also gain from specialisation economies and are more efficient.

Table 6. Efficiency and relative specialisation (RELSPEC) - CRS model

Variable	Coefficient	95% Bootstrap Confidence Intervals	
		Low	*High*
$RELSPECy_1$	0.129	-0.996	1.047
$RELSPECy_2$	1.261*	0.684	1.753
$RELSPECy_3$	0.445	-0.309	1.079
$RELSPECy_4$	-0.434*	-0.766	-0.106
RELSIZE	59.220	-14.102	101.309
$RELSPECy_1$*RELSIZE*	-84.011*	-109.950	-38.959
$RELSPECy_2$*RELSIZE*	-40.310*	-60.829	-14.436
$RELSPECy_3$*RELSIZE*	-38.015*	-56.163	-9.486
$RELSPECy_4$*RELSIZE*	-37.789*	-49.645	-15.219
DUMMYPASS	0.072	-0.596	0.587
D_{1994}	-0.506*	-1.031	0.164
D_{1995}	-0.223	-0.785	0.380
D_{1996}	-0.405	-0.971	0.211
D_{1997}	-0.668	-1.181	0.002
D_{1998}	-1.203*	-1.748	-0.464
D_{1999}	-0.983*	-1.560	-0.275
D_{2000}	-1.204*	-1.737	-0.479
D_{2001}	-1.442*	-1.952	-0.662
D_{2002}	-1.582*	-2.143	-0.811
D_{2003}	-1.782*	-2.299	-0.973
D_{2004}	-1.692*	-2.211	-0.900
D_{2005}	-1.833*	-2.309	-1.001
D_{2006}	-2.133*	-2.700	-1.234
D_{2007}	-2.163*	-2.756	-1.304
D_{2008}	-1.696*	-2.273	-0.959
D_{2009}	-0.644*	-1.205	-0.008
D_{2010}	-0.555*	-1.102	0.139
D_{2011}	-1.183*	-1.754	-0.479
D_{20012}	-1.494*	-2.035	-0.820
Observations	520		

* Significant at 5% level

We also assess the interaction terms ($RELSPECy_i$*RELSIZE*), which correspond to the national share of the output (*NATSHARE*). The disaggregation of the *NATSHARE* variable into its two components as represented by the interaction of *RELSPEC* and *RELSIZE* is insightful in that it shows explicitly that the effect on efficiency of being relatively more specialized in a given output. Following a similar procedure to that followed for the absolute specialisation interaction terms, Table 8 reports the marginal effects where the cross derivatives of each *RELSPEC* variable are evaluated at the 25th, 50th, 75th and 95th percentiles of the size variable (*RELSIZE*). Some differences exist between the CRS and VRS results. The interaction terms are almost all negative and significant, with exceptions being *RELSIZE**$RELSPEC_{y2}$ (solids) for both models and *RELSIZE**$RELSPEC_{y3}$ (containers) for the VRS model. The negative values show that the larger the port, the greater the positive effect of relative specialisation in a given output on efficiency. From the results in Table 8, the effect of size on relative specialisation appears to be particularly strong for liquids and general merchandise regardless of the model used.

Table 7. Efficiency and relative specialisation (RELSPEC) - VRS model

Variable	Coefficient	95% Bootstrap Confidence Intervals	
		Low	*High*
$RELSPECy_1$	0.833	-0.266	1.673
$RELSPECy_2$	0.598*	0.172	1.136
$RELSPECy_3$	0.516	-0.202	1.112
$RELSPECy_4$	-0.272	-0.636	0.024
RELSIZE	-4.827*	-74.248	61.425
$RELSPECy_1$*RELSIZE*	-69.363*	-106.147	-24.863
$RELSPECy_2$*RELSIZE*	-28.969*	-58.876	-7.355
$RELSPECy_3$*RELSIZE*	-27.778	-50.585	7.447
$RELSPECy_4$*RELSIZE*	-37.709*	-51.721	-11.125
DUMMYPASS	0.625	-0.033	1.046
D_{1994}	-0.267	-0.804	0.349
D_{1995}	0.162	-0.378	0.718
D_{1996}	0.010	-0.557	0.573
D_{1997}	-0.019	-0.545	0.580
D_{1998}	-0.476	-1.023	0.153
D_{1999}	-0.224	-0.828	0.363
D_{2000}	-0.355	-0.902	0.258
D_{2001}	-0.431	-0.965	0.203
D_{2002}	-0.547	-1.156	0.074
D_{2003}	-0.716*	-1.260	-0.072
D_{2004}	-0.770*	-1.318	-0.118
D_{2005}	-0.794*	-1.310	-0.103
D_{2006}	-1.066*	-1.661	-0.368
D_{2007}	-1.086*	-1.699	-0.386
D_{2008}	-0.670*	-1.288	-0.054
D_{2009}	0.505	-0.081	0.994
D_{2010}	0.604*	0.014	1.119
D_{2011}	-0.102	-0.695	0.458
D_{20012}	-0.648*	-1.213	-0.085
Observations	520		

* Significant at 5% level

Table 8. Marginal effects of relative specialization interaction terms

		CRS		VRS	
Variable	Percentile RELSIZE	Marginal Effect	SE	Marginal Effect	SE
$RELSPEC_{Y1}$	25	-0.771	0.433	0.064	0.389
	50	-2.136*	0.393	-1.104*	0.385
	75	-4.008*	0.732	-2.707*	0.797
	95	-8.976*	1.987	-6.960*	2.163
RELSPECY2	25	0.819*	0.232	0.284	0.211
	50	0.142	0.311	-0.259	0.330
	75	-0.786	0.561	-1.003	0.635
	95	-3.249*	1.338	-2.978	1.529
RELSPECY3	25	0.038	0.288	0.218	0.256
	50	-0.582*	0.286	-0.225	0.306
	75	-1.432*	0.484	-0.832	0.612
	95	-3.687*	1.227	-2.443	1.575
RELSPECY4	25	-0.821*	0.147	-0.665*	0.130
	50	-1.439*	0.225	-1.275*	0.244
	75	-2.286*	0.434	-2.113*	0.499
	95	-4.534*	1.047	-4.336*	1.221

* Significant at 5% level

Taken together, the results from the model exploring the relationship between relative specialisation and efficiency (10) provide an affirmative response to our second question: *controlling for size, ports that are relatively specialised in a given output compared to the sector as a whole are more efficient*. This effect, for the pooled DEA model, seems particularly strong when the ports are relatively specialised in liquid bulk and general merchandise.

The results reported so far regarding the relationship between technical efficiency and specialisation were all based on DEA technical efficiency calculations from a pooled model, i.e., all observations were used to construct the frontier, with no distinction made between port authorities or years. To test the robustness of our results, it would be interesting to see how they change when different DEA methods are used to construct the efficiency scores. As noted by Cullinane and Wang (2010) and Cullinane, Ji and Wang (2005), outcomes can vary considerably depending on the DEA model used. To investigate this, Table 9 reports the DEA efficiency scores from a series of different models. At the other extreme from the *Pooled* model is the *Yearly* model, where a different frontier is estimated for each year. Intermediate cases include the 5-year *Window* DEA model, where DEA scores are calculated using moving 5-year windows, and a *Sequential* model where previous periods' technologies are always assumed feasible for firms and efficiency scores are calculated with reference to present and past observations only (see Fried *et al*, 2008, for a general reference). Both constant returns to scale (CRS) and variables returns to scale (VRS) models are reported. We also report the DEA efficiency scores from the Pooled model where inefficiency is measured using an input-oriented frontier model (*Pooled-IO*) to check whether the orientation used to calculate the efficiency scores alters the relationship between specialisation and technical efficiency. As the Pooled input-oriented and output-oriented efficiency scores are identical under the assumption of CRS, only the VRS efficiency scores are reported for the input-oriented model.

Table 9. DEA efficiency scores

	CRS Models				VRS Models				
Port	**Pooled**	**Yearly**	**Window**	**Sequential**	**Pooled**	**Yearly**	**Window**	**Sequential**	**Pooled-**
A Coruña	0.62	0.86	0.86	0.86	0.66	0.89	0.89	0.71	0.66
Alicante	0.23	0.42	0.41	0.41	0.24	0.44	0.44	0.27	0.49
Avilés	0.69	0.93	0.93	0.93	0.74	0.99	0.99	0.84	0.82
B.	0.97	1.00	1.00	1.00	0.99	1.00	1.00	1.00	0.98
B. Cádiz	0.47	0.64	0.64	0.64	0.49	0.66	0.66	0.55	0.53
Baleares	0.90	1.00	1.00	1.00	0.90	1.00	1.00	0.98	0.91
Barcelona	0.55	0.68	0.68	0.68	0.76	0.98	0.98	0.97	0.61
Bilbao	0.75	0.88	0.88	0.88	0.86	0.99	0.99	0.96	0.76
Cartagena	0.81	0.97	0.97	0.97	0.82	0.97	0.97	0.91	0.84
Castellón	0.83	0.97	0.97	0.97	0.86	1.00	1.00	0.88	0.89
Ceuta	0.59	1.00	0.80	0.80	0.61	1.00	0.98	0.64	0.72
Ferrol-S.	0.87	1.00	1.00	1.00	0.94	1.00	1.00	0.95	0.92
Gijón	0.64	0.77	0.77	0.77	0.90	1.00	1.00	0.95	0.85
Huelva	0.82	0.99	0.99	0.99	0.90	1.00	1.00	0.93	0.87
Las Palmas	0.57	0.74	0.74	0.74	0.57	0.77	0.77	0.71	0.61
Málaga	0.41	0.76	0.70	0.70	0.44	0.79	0.72	0.46	0.55
Marín-Ría-	0.45	0.68	0.68	0.68	0.55	0.95	0.95	0.60	0.81
Melilla	0.81	0.97	0.97	0.97	0.94	1.00	1.00	0.99	0.99
Pasajes	0.61	0.83	0.83	0.83	0.64	0.87	0.87	0.75	0.65
S.C.	0.81	0.99	0.96	0.96	0.83	1.00	0.97	0.90	0.84
Santander	0.34	0.47	0.47	0.47	0.39	0.53	0.53	0.47	0.42
Sevilla	0.32	0.46	0.46	0.46	0.36	0.51	0.51	0.42	0.45
Tarragona	0.77	0.95	0.95	0.95	0.94	1.00	1.00	0.97	0.90
Valencia	0.79	0.93	0.93	0.93	0.84	0.99	0.99	0.99	0.80
Vigo	0.35	0.54	0.51	0.51	0.35	0.57	0.56	0.44	0.43
Vilagarcía	0.32	0.50	0.50	0.50	0.49	1.00	1.00	0.63	0.91
Mean	*0.63*	*0.80*	*0.79*	*0.79*	*0.69*	*0.88*	*0.88*	*0.76*	*0.74*

As we would expect, the DEA scores from the *Yearly*, *Window* and *Sequential* models are higher than those from the Pooled models. This is a consequence of the fact that these models use subsets of the data when calculating the scores, so that firms have fewer observations to be compared with and are more likely to be efficient. The correlation coefficients among the DEA scores from the models are presented in Table 10 for the CRS and VRS specifications. As can be seen, the scores are highly correlated across models, especially for the CRS specifications.

Table 10. Correlations between DEA efficiency scores from different models

CRS MODELS

	POOLED	YEARLY	WINDOW	SEQUENTIAL
POOLED	1.00	-	-	-
YEARLY	0.94	1.00	-	-
WINDOW	0.97	0.98	1.00	-
SEQUENTIAL	0.97	0.98	1.00	1.00

VRS MODELS

	POOLED	YEARLY	WINDOW	SEQUENTIAL	POOLED-IO
POOLED	1.00	-	-	-	
YEARLY	0.85	1.00	-	-	
WINDOW	0.86	1.00	1.00		
SEQUENTIAL	0.98	0.86	0.87	1.00	
POOLED-IO	0.85	0.88	0.89	0.81	1.00

The overall marginal effects of the overall specialisation (*HHI*), relative specialisation (*RELSPEC*) and relative size (*RELSIZE*) variables for these alternative DEA models for both the CRS and VRS specifications are presented on Table 4. From these results it is clear that The positive relationship between specialisation efficiencies and overall specialisation (*HHI*) for the pooled model manifested by a negative sign of the derivative (which shows that increases in HHI are associated with decreases in technical inefficiency) is repeated for the alternative DEA models under a VRS specification. The derivative is negative in all cases, and significant at the 5% level for the panel model (it is significant at the 10% level for the window and sequential models). In the CRS specifications, on the other hand, the signs of the derivative are positive, though not significantly significant.

Regarding relative specialisation, the results from the pooled model hold for all the alternative DEA models under both CRS and VRS specifications. In the CRS specifications, all derivatives for the *Yearly*, *Window* and *Sequential* models are negative and significant, as was the case for the pooled model. In the VRS specifications, all derivatives are again negative except for input-oriented poled model, with the vast majority significant. In the *Pooled-IO* (input-oriented) model, the relative specialisation derivatives for liquids and general merchandise (y_1, y_4) are negative and significant, as with the output-oriented model. However, whereas in the output-oriented model there was no statistically significant relationship between relative specialisation in solids (y_2) and technical efficiency, this relationship was found to be negative and significant for the input-oriented model. Finally, the size variable (*RELSIZE*) is negative and significant for all models in both CRS and VRS specifications, as in the (output-oriented) pooled model.

Overall these results show that the relationship between relative specialisation in liquids and general merchandise and technical efficiency is strongly robust and holds across different DEA specifications. The relationship between relative specialisation in solids and containerised merchandise and technical efficiency is also positive and statistically significant in virtually all of the models presented, an exception being the negative relationship between relative specialisation in solids and efficiency in the input-oriented pooled DEA model. The relationship between overall specialisation and technical efficiency holds across different DEA models, including the input-oriented pooled model, for the VRS specifications. For the CRS specifications, the relationship is positive and statistically significant for the pooled output- and input-oriented models, but is not significant for the other models.

As a final robustness test, we allow for the possibility that the specialisation indices in the second-stage regression are endogenous. To check this, we use the one-period lagged values of the specialisation indices as instruments and carry out Hausman specification tests, and endogeneity was rejected. The marginal effects of the specialisation indices from the bootstrapped regression using the lagged indices as instruments, which we label *Pooled-IV*, are reported in Table 4 and it can be seen that they are very similar to those obtained under the *Pooled* model.

6. Conclusions

We have contributed to the literature on port efficiency by analysing the relationship between output specialisation/diversification, size and technical efficiency. Using the frontier technique of Data Envelopment Analysis (DEA), we have estimated an output distance frontier to measure technical efficiency for a set of 26 Spanish Port Authorities observed over the period 1993-2012. We calculate indices of overall output specialisation for these port authorities using the normalised Herfindahl-Hirschman Index, as well as indices of relative specialisation for each type of cargo. We then estimate the relation between DEA efficiency scores, size and our measures of output specialisation using bootstrapped truncated regressions.

We find that output specialisation has a positive effect on port authority technical efficiency, with the DEA efficiency scores being positively related to the overall specialisation index. Port authorities with greater relative specialisation in two of the cargoes – liquid bulk and general merchandise – were also found to be more efficient, whereas the relation between relative specialisation in containerised goods and solid bulk and efficiency was less clear and depended on the type of DEA model used. Moreover, we find that the effect of specialisation is greater for larger port authorities. That is, for any given of level of overall specialisation measured by the concentration index, larger port authorities, which may be specialised in several types of cargo, are found be more technically efficient. This can be interpreted as evidence that these large ports, though diversified, also take advantage of specialisation efficiencies. Thus, larger port authorities, which tend to be somewhat less specialised overall (as measured by the HHI) are more likely to be specialised in several different cargoes and are more efficient as a result due to the fact that they are in a position to benefit from economies of scope and scale simultaneously. The same argument applies for relative specialisation. Ports relatively specialised, compared to the system as a whole, in liquid bulk and general merchandise tend to be more efficient regardless of size. However, as the ports grow in size, the positive effect of relative specialisation on efficiency intensifies.

Our work shows that both overall and relative specialisation increase technical efficiency, with the effect of specialisation being reinforced as ports grow in size because larger port tends to be specialised in several outputs which let them take advantages form economies of scale and scope. Smaller ports have a relative disadvantage compared to larger ports, as they cannot specialise in multiple outputs. However, while the benefits from specialisation increase with size, our results suggest that smaller ports can also benefit from specialisation.

From a policy perspective, our results suggest that expanding the size of smaller port authorities to allow them to take advantage of economies of scale in different goods is to be recommended. Increasing the size of ports allows them to take advantage of economies of scale in a range of individual outputs, which means they can better compensate for adverse changes in demand for a particular cargo. This is highlighted in our result that larger ports are closer to their frontier regardless of the level of specialisation. However, if this is not possible due to, say, budgetary constraints, smaller port authorities can make better use of resources by concentrating their output. Our work shows that specialised ports take better advantage of their resources, implying that there is a trade-off between the dangers of being exposed to the vicissitudes of the market for the output in question and the benefits of taking advantage of specialisation in that output.

Finally, some words are in order about where future research on the issues raised in this paper could go. Firstly, we see no reason why the results of our study of the Spanish port system should not be applicable to other countries. It would therefore be interesting to see to what extent our results are corroborated by studies using data from other countries or continents. Secondly, it would be interesting to whether our results might be affected by using stochastic frontier analysis instead of DEA. Several stochastic frontier models for panel data are available in the literature, with some recent models permitting the simultaneous incorporation of individual heterogeneity and time-invariant and time-varying inefficiencies. An overview of these models is provided by Kumbhakar *et al*. (2014) and studies of the relationship between specialisation, size and efficiency using these parametric models would provide a natural comparison for DEA-based studies such as this one.

Acknowledgements

The authors would like to thank the three anonymous reviewers for their insightful comments on the paper. Addressing the points raised has led to the improvement of this work. Financial assistance for this work was provided by Grant ECO2015-68345-R (MINECO/FEDER) from the Spanish Ministry of Economy and Competitiveness and the European Regional Development Fund (FEDER), and Grant FC-15-GRUPIN14-048 from the Regional Government of the Principality of Asturias in conjunction with FEDER.

References

Ai, C. and Norton, E.C. (2003). Interaction terms in logit and probit models. *Economic Letters*, 80, 123-129.

Al-Marhubi, F. (2000). Export diversification and growth: an empirical investigation. *Applied Economics Letters*, 7, 559-562.

Bang, H.-S., Kang, H.-W., Martin, J. and Woo, S.-H. (2012). The impact of operational and strategic management on liner shipping efficiency: a two-stage DEA approach. *Maritime Policy & Management*, 39(7), 653–672.

Barros, C. (2005). Decomposing growth in Portuguese seaports: A frontier cost approach. *Maritime Economics & Logistics*, 7, 297–315.

Barros, C. P., Assaf, A. and Ibiwoye, A. (2010). Bootstrapped technical efficiency of African seaports. In P. Coto-Millán, M. A. Pesquera, and J. Castanedo (eds) *Essays on Port Economics,* Heidelberg.

Chang, V. and Tovar, B (2014a). Efficiency and productivity changes for Peruvian and Chilean ports terminals: A parametric distance functions approach. *Transport Policy,* 31, 83- 94.

Chang, V. and Tovar, B (2014b). Drivers explaining the inefficiency of Peruvian and Chilean ports terminals. *Transportation Research Part E,* 67, 190-203.

Chang, V. and Tovar, B (2016). Metafrontier Analysis on Productivity for West Coast of South Pacific Terminals. Transportation Research Part E. 67, 190-203. Paper presented at the *International Association of Maritime Economist*, *IAME 2016*, August 2016, Hamburg. Germany.

Chang, V. and Tovar, B (2017a). Heterogeneity unobserved and efficiency: A latent class model for west coast of South Pacific terminals. *Journal of Transport Economics and Policy (In press)*.

Chang, V. and Tovar, B (2017b). The Determinants of efficiency for the West Coast of South Pacific Terminals: A Second-Stage Dea Approach in a Metafrontier Framework. Departamento de Economia Aplicada Working Paper. Spain.

Cheon S., Dowall, D. and Song D-W. (2010). Evaluating impacts of institutional reforms on port efficiency changes: Ownership, corporate structure, and total factor productivity changes of world container ports. *Transportation Research Part E*, 46, 546–561.

Cheon, S., Dowall, D., and Song, D.-W., (2010). Evaluating impacts of institutional reforms on port efficiency changes: ownership, corporate structure, and total factor productivity changes of world container ports. *Transportation Research Part E: Logistics and Transportation Review*, 46 (4), 546–561.

Coelli, T. and Fleming, E. (2004). Diversification economies and specialisation efficiencies in a mixed food and coffee smallholder farming system in Papua New Guinea. *Agricultural Economics*, 31, 229-239.

Cullinane, K. (2010). Revisiting the Productivity and Efficiency of Ports and Terminals: Methods and Applications. In Grammenos, C.T., Nomikos, N.K. and Alizadeh, A.H. (eds.). *The Handbook of Maritime Economics and Business*, Lloyd's List, London.

Cullinane, K.P.B. and Song, D-W. (2003). A stochastic frontier model of the productive efficiency of Korean container terminals. *Applied Economics*, 35, 251–67.

Cullinane, K.P.B. and Wang T. (2010). The efficiency analysis of container port production using DEA panel data approaches, *OR Spectrum*, 32(3), 717–738.

Cullinane, K.P.B., Ji, P. and Wang, T. (2005) The relationship between privatization and DEA estimates of efficiency in the container port industry, *Journal of Economics and Business*, 57, 433-462.

Cullinane, K.P.B., Song, D-W., Ji, P. and Wang T. (2004). An application of DEA windows analysis to container port production efficiency. *Review of Network Economics,* 3(2), 184-206.

Cullinane, K., Song, D.W. and Gray, R. (2002). A stochastic frontier model of the efficiency of mayor container terminals in Asia: Assessing the influence of administrative and ownership structures. *Transportation Research Part A*, 36, 743-762.

Cullinane, K., Wang, T.F., Song, D.W. and Ji, P. (2006). The technical efficiency of container ports: Comparing data development analysis and stochastic frontier analysis. *Transportation Research Part A*, 40, 354-374.

Ducruet, C., Van der Beek, D.J. and Koster, H.R.A. (2010). Commodity variety and seaport performance. *Regional Studies*, 44(9), 1221-1240.

Estache, A., Tovar, B., & Trujillo, L. (2004). Sources of efficiency gains in port reform: A DEA decomposition for a Malmquist TFP index for Mexico. *Utility Polic*y, 12(4), 221–230.

Fried, H.O., Lovell, C.A.K. and Schmidt, S.S. (2008). *The Measurement of Productive Efficiency and Productivity Growth*. OUP, New York.

Frémont, A. and Soppé, M. (2007). Northern European Range: Shipping Line Concentration and Port Hierarchy. In J. Wang, T. Notteboom, D. Oliver, B, Slack, (eds) *Ports, cities, and global supply chains*. Aldershot.

Grant, R. M., Jammine, A. P. and Thomas, H. (1988). Diversity, Diversification, and Profitability among British Manufacturing Companies. *Academy of Management Journal*, 31,771-801.

González-Laxe, F. (2012). La nueva gestión portuaria española. *Boletín Económico del ICE*, nº 3024.

González-Laxe, F. and Novo Corti, I. (2012). Spanish port Competitiveness: Port system response to Economic crisis. Working Paper. Universidad de A Coruña.

González, M.M. and Trujillo, L. (2008). Reforms and infrastructure efficiency in Spain's container ports. *Transportation Research Part A*, 42, 243-257.

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

Jara Diaz, S., Tovar De La Fe, B. and Trujillo, L. (2005). Multioutput analysis of cargo handling firms: An application to a Spanish port, Transportation, 32(3), 275-291.

Jara Diaz, S., Tovar, B. and Trujillo, L. (2008). On the proper modelling of multioutput port cargo handling costs. *Applied Economics*, 40, 1699-1705.

Huybrechts, M., Meersman, H., van de Voorde, E., Van Hooydonk, E., Verbeke, A. and Winkelmans, W. (2002). *Port competitiveness: an economic and legal analysis of the factor determining the competitiveness of seaports*. Editions De Boeck Ltd, Antwerp.

Karaca-Mandic, P., Norton, E.C. and Dowd, B. (2012). Interaction Terms in Nonlinear Models, *Health Services Research*, 47(1), 255-274.

Kumbhakar, S.C., Lien, G., and Hardaker, J.B. (2014). Technical efficiency in competing panel data models: A study of Norwegian grain farming. *Journal of Productivity Analysis,* 41, 321–337.

Liu, Z. (1995). The comparative performance of public and private enterprise: The case of British ports. *Journal of Transport Economics and Policy*, 29, 263-274.

Martinez-Budria, E., Diaz-Armas, R., Navarro-Ibanez, M., & Ravelo-Mesa, T. (1999). A study of the efficiency of Spanish port authorities using data envelopment analysis. *International Journal of Transport Economics*, XXVI (2), 237–253.

Mateo-Mantecón, I., Rodríguez-Vázquez, E., Casares-Hontañón, P. and Coto Millán, P. (2015). Indicadores económicos del sistema portuario español. Paper presented at the *Spanish regional Economcis Association meeting*, Rovira-Virgili University, Reus, Spain.

Meyler, L., Moiseenko, S. and Volkogon, V. (2011). Diversification of the seaport activity as an effective post-crisis strategy. Paper presented at the *European Conference on Shipping Intermodalism & Ports (ECONSHIP)*, Chios, Greece.

Nguyen, H.-O., Nguyen, H.-V., Chang, Y.-T. and Chin, A.T.H. and Tongzon, J. (2016). Measuring port efficiency using bootstrapped DEA: the case of Vietnamese ports. *Maritime Policy & Management*, 43(5), 644-659.

Notteboom, T. E. (1997). Concentration and load centre development in the European container port system. *Journal of Transport Geography*, 5(2), 99–115.

Notteboom, T. E. (2010). Concentration and the formation of multi-port gateway regions in the European container port system: an update. *Journal of Transport Geography*, 18(4), 567–583.

Nuñez-Sánchez, R., Coto-Millán, P. (2012). The impact of public reforms on the productivity of the Spanish ports: A parametric distance function approach, *Transport Policy*, 24, 653-666.

Reina, R. and Villena, F. (2013). Concentration of goods traffic in Spanish ports during the period 2000-2009. *Regional and Sectoral Economic Studies*, 13, 59-72.

Riordan, M. H. and Williamson, O. E. (1985). Asset specificity and economic organization, *International Journal of Industrial Organization*, 3, 365-378.

Rodríguez-Álvarez, A. and Tovar, B. (2012). Have Spanish port reforms during the last two decades been successful? A cost frontier approach. *Transport Policy*, 24, 73-82.

Rodríguez-Álvarez, A., Tovar, B. and Wall, A. (2011). The effect of demand uncertainty on port terminals costs. *Journal of Transport Economics and Policy*, 45, 303-328.

Schøyen, H. and Odeck. J. (2013). The technical efficiency of Norwegian container ports: A comparison to some Nordic and UK container ports using Data Envelopment Analysis (DEA). *Maritime Economics & Logistics*, 15(2), 197–221.

Shephard, R.W. (1970). *Theory of Cost and Production Functions*. Princeton University Press, Princeton.

Simar, L. (2007). How to improve the performances of DEA/FDH estimators in the presence of noise? *Journal of Productivity Analysis*, 28(3), 183–201.

Simar, L. and Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, 44(1), 49–61.

Simar, L., and Wilson, P. W. (2000). A general methodology for bootstrapping in non-parametric frontier models. *Journal of Applied Statistics* 27 (6): 779–802.

Simar, L. and Wilson, P.W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, 136, 31–64.

Tongzon, J. and Heng, W. (2005). Port privatization, efficiency and competitiveness: Some empirical evidence from container ports (terminals). *Transportation Research*, 39, 405-424.

Sys, C. (2009). Is the container liner shipping industry an oligopoly? *Transport Policy*, 16, 259–270.

Tovar de la Fé, B., Trujillo, L. and Jara-Díaz, S. (2004). Organization and regulation of the port industry: Europe and Spain. In P. Coto-Millan (ed) *Essays on Microeconomics and Industrial Organisation*. Second Edition. Physica-Verlag. A Springer-Verlag Company. Germany.

Tovar, B. Jara Diaz, S. and Trujillo, L. (2007). Econometric estimation of scale and scope economies within the port sector: a review. *Maritime Policy and Management*, 34, 203-223.

Tovar, B. and Wall, A. (2012). Economies of scale and scope in service firms with demand uncertainty: An application to a Spanish port. *Maritime Economics & Logistics*, 14, 362–385.

Tovar, B. and Wall, A. (2014). The impact of demand uncertainty on port infrastructure costs: Useful information for regulators? *Transport Policy*, 33, 176-183.

Tovar, B. and Wall, A. (2015). Can Ports increase traffic while reducing inputs? Technical efficiency of Spanish Port Authorities using a directional distance function approach, *Transport Research Part A*, 71, 128-144.

Tovar, B. and Rodríguez-Déniz H. (2015). Classifying Ports for Efficiency Benchmarking: A Review and a Frontier-based Clustering Approach. *Transport Reviews: A Transnational Transdisciplinary Journal*, 35(3), 378-400.

Tovar, B. and Wall, A. (2016). Dynamic cost efficiency in port infrastructure using a directional distance function: accounting for the adjustment of quasi-fixed inputs over time. *Transportation Science.* Forthcoming.

Trujillo, L. and Tovar, B. (2007). The European port industry: An analysis of its economic efficiency. *Maritime Economics & Logistics*, 9, 148-171.

Wilmsmeier, G. Tovar, B and Sanchez, R. (2013). The evolution of container terminal port productivity and efficiency under changing economic environments. *Research in Transportation Business and Management,* 8, 50–66.

Yuen, A. C.-L., Zhang, A. and Cheung, W. (2013). Foreign Participation and Competition: A Way to Improve the Container Port Efficiency in China? *Transportation Research Part A: Policy and Practice*, 49, 220–231.