A port choice model with logit models: a case study for the Spanish container trade

Abstract

Earlier research by the authors showed that transport costs, port specific dummy variables and proxies for quality of service aspects have a statistically significant impact on port choice. The objective of this paper is to improve the earlier models through the use of a nested logit model and further by including new variables. The results show that the nested model is an improvement, that the variable indicating the inland transport balance has a statistically significant impact and that the split of variables into distance dependant and independent components gives no improvement. The model outcomes can be used to assess the impact of interventions in ports, and thereby to serve as a basis for the economic and financial evaluation of port projects.

Keywords

Container port competition; port choice; nested logit model; disaggregated data; Spain.

1 Introduction

For interventions in container ports such as expansion projects and tariff setting it is important to understand the relation between port choice demand and the price of using a port. Port choice demand can be assessed by applying multinomial logit (logit) models to the bundles of container flows passing competing ports. During the last decade a nascent literature of more than a dozen publications came up on tests of logit models using revealed preference data. The authors took part in this development and this paper is an extension of their joint research with respect to Spanish container trades.

The port choice model tests published differ with respect to aspects such as the choice position (ports only, ports in combination with shipping lines or with inland modes of transport), type of data (aggregated or disaggregated), available statistical information (some countries publish data on the combined hinterland and foreland origin and destination of cargo flows and port of transfer) and statistical method (regression analysis or specific Maximum Likelihood estimation method). These differences often lead back to the geography of the ports and available statistical data. The methodology applied is a combination of all these factors and, of course, the size of the research budget.

Veldman et al. (Veldman et al., 2011), to be referred to as 'previous study', successfully tested logit models with respect to port choice for Spanish container trades. They used a conditional logit model to estimate the impact of inland transport costs, maritime transport costs, port specific dummy variables and proxies for quality of service aspects. The objective of this paper is to improve the model by (a) testing a two-phase choice function (first coastline - then port) by using a nested model, (b) including a new variable to indicate the inland container transport balance and (c) modifying the variables to reflect the impact of inland distance.

Section 2 describes tests of logit models with respect to port choice using revealed preference data. Section 3 gives a description of the structure of Spanish container trade by sea. Section 4 describes the attributes of the proposed model. Section 5 presents the results of alternative model specifications. Conclusions are drawn in Section 6.

2 Port choice models

2.1 Use of logit models for port choice

As far as we know statistical analyses of port choice using demand choice models, such as logit models, started about a decade ago with Malchow and Kanafani (Malchow and Kanafani, 2001). These analyses correspond with choice functions used with transportation planning models which started much earlier in the 1970s with the model split and traffic assignment stage of these models in particular, first for passenger transport and later on for transport of goods. The models are described in handbooks on transport modelling such as Hensher and Button (Hensher and Button, 2000) and Ortuzar and Willumsen (Ortuzar and Willumsen, 1990). Winston (Winston, 1981) expanded the applications to maritime transport by including it the modal split.

Publications on routing choice/ port choice are part of literature on port competition and competitiveness, and it is useful to give the application of choice models a place within this literature. The world-wide network of container liner services is becoming more fine-meshed and with this the number of routings offered per pair of hinterland - foreland regions. This increase is a reflection of the world-wide restructuring of global supply chains. The increase of routing options, each using a different port, leads to more competition between ports and thereby to a need for adequate tools of analysis, as noted by Robinson (Robinson, 2002). The world-wide network of container services with ports as transfer points shows similarities with the movements of passengers and goods at land. Therefore, it is logic that also similarities exist with demand choice situations with land transport and the models developed for land transport – the land transportation planning models referred to above – developed much earlier.

In their study of methodological issues in seaport research Woo *et al.* (Woo et al., 2011) counted 16 publications using logit models in the areas of port competition and performance, planning and development and management and strategy. Not all these include statistical tests with revealed preference data, which to our knowledge come at a total of more than a dozen. Paixao *et al.* (Paixao Casaca et al., 2010) identified 56 applications of port choice of which 11 using logit models.

At one end are publications on choice of ports located along a continuous stretch of coastline, such as for continental West Europe by Veldman and Buckmann (Veldman and Bückmann, 2003) and Veldman *et al.* (Veldman et al., 2005) and for Taiwan by Nir *et al.* (Nir et al., 2003). At the other end it concerns ports located on unconnected stretches of coast line such as in the United States by Malchow and Kanafani (Malchow and Kanafani, 2001) (Malchow and Kanafani, 2004), Blonigen and Wilson (Blonigen and Wilson, 2006) and Anderson *et al.* (Anderson et al., 2009) and in Russia by Veldman and Gopkalo (Veldman and Gopkalo, 2010). The information on choice ranges from disaggregate data on individual shipments such as by Malchow and Kanafani (Malchow and Kanafani, 2001) (Malchow and Kanafani, 2004), Tiwari *et al.* (Tiwari et al., 2003) to aggregate data related to flows of containers such as by Veldman and Buckmann (Veldman and Bückmann, 2003), Garcia-Alonso and Sanchez–Soriano (Garcia-Alonso and Sanchez-Soriano, 2009), Veldman *et al*. (Veldman et al., 2011) and Veldman and Gopkalo (Veldman and Gopkalo, 2010).

2.2 The port choice model

The analysis of the port choice from the perspective of discrete choice models considers that the chosen port is the one that maximizes the utility of the routing choice decision-maker, assuming that he or she is rational, i.e. with transitive and consistent preferences. However, the choices do not always satisfy these properties. This is why the utility is also considered as a random variable with two components: the deterministic one and the error term. The choice problem is then analyzed in probabilistic terms, and the type of model depends on the hypothesis about the distribution of the errors.

2.2.1 Conditional logit model

When the error terms are independently and identically distributed extreme value (Gumbel), and taking into account that the difference between two Gumbel distributions is a logistic distribution, the probability of choosing any alternative can be estimated by a logit model (McFadden, 1974). The independent variables can refer to the decision-makers' characteristics or to those of the alternatives, the attributes. When they are associated to the characteristics of the individual, the logit model is multinomial; when they describe the attributes, the model is conditional, as in our case.

The choice of seaport concerns the routings of Spanish imports or exports shipped between the gravity point of the Spanish province of import or export and the gravity point of the overseas trade partner. The logit model expresses the probability that an importer or exporter trading between one of the Spanish peninsular provinces *i* and one of the overseas trade partners *j*, chooses port *k* from a set of possible ports. Per combination of province and trade partner region the probability of choosing a routing via one of the ports, can be expressed as:

$$P_{ij}(i = k|k = 1..I) = \frac{e^{U_{ij}}}{\sum_{k=1}^{k=I} e^{kj}} \qquad (1)$$

Where P_{ijk} is the probability of choosing port *k* from all possible ports $p = 1..P$, for province $i = 1..I$ and trade partner $j = 1...J$; and U_{ijk} is the deterministic component of the *utility* for routing via port *k* of the trade flows between *i* and *j*. The probability P_{ijk} can be interpreted as the market share of a port *k* in the total of all ports serving the trade between province *i* and trade partner *j*, for either import or export. To test the model the probability P_{ijk} can be set equal to the observed market share of volume F_{ijk} of routing *k* in the trade between *i* and *j*.

The value which a trader attaches to routing *k* is measured in the utility, which can be expressed as a (linear) combination of all aspects impacting the choice between alternative ports. One of the models tested is:

$$U_{ijk} = \alpha_0^k + \alpha_1 CL_{ik} + \alpha_2 CM_{jk} + \alpha_3 Q_k \quad (2)$$

Where CL_{ik} is the inland transport cost between province *i* and port *k*; CM_{jk} is the maritime transport cost between trade partner *j* and port *k* and where Q_k is an indicator of quality of service aspects for *i*, *j* and port *k*.

The explanatory variables, referred to as attributes, are CL_{ik}, CM_{jk} and Q_k; α_0, α_1, α_2 and α_3 are the coefficients of the utility function. Alternative versions of equation 2 are discussed later on.

The relative position of one port against the other for trade pair i,j is expressed by the ratio of the probability that a trader chooses a routing via port k against the probability that he chooses a routing via port p. By subsequently substituting k and p in equation (1) and dividing the resulting probabilities, the ratio becomes:

$$P_{ijk}/P_{ijp} = e^{-U_{ijk}}/e^{-U_{ijp}} = e^{U_{ijp}-U_{ijk}} \quad (3)$$

The ratio of probabilities becomes a function of the differences of their attributes, which is a convenient form. If instead of differences a ratio form would apply (this applies to the situation where it is assumed that the utility function has a multiplicative instead of a linear form) the absolute level of the attribute values should have to be known and thereby of information on all shackles of the transport chain. Oum (Oum, 1989) states that with the ratio form the choice of base routing n affects the empirical results, including own and cross elasticities of demand.

By taking the logarithm of equation 3 the model becomes convenient for estimation with regression analysis:

$$Ln\left(P_{ijk}/P_{ijp}\right) = U_{ijp} - U_{ijk} = \alpha_0 + \alpha_1\left(LC_{ip} - LC_{ik}\right) + \alpha_2\left(MC_{jp} - MC_{jk}\right) + \alpha_3\left(Q_p - Q_k\right) (4)$$

Where $\alpha_0 = \alpha^k{}_0 - \alpha^p{}_0$

2.2.2 Nested logit models

In practice the hypothesis of independence from irrelevant alternatives (IIA) appears to very restrictive, as it imposes that the ratio of probabilities between any two alternatives is independent of the remaining alternatives and of their attributes. This hypothesis assumes no correlation between alternatives, which can be a problem when there is similarity between some of them. To solve this problem, the hypothesis of IIA can be relaxed applying nested models; that is, grouping alternatives in nests with some degree of similarity between them, in such a way that the hypothesis of IIA remains valid within each nest. So the process of choice is made in steps, following a tree structure: firstly one group is chosen among the set of groups, and then one of the alternatives in that group.

With *P* alternatives, distributed among *N* nests, the probability of selecting alternative *k* belonging to nest *m* is (Train, 2009):

$$P_k = \frac{e^{U_k/\lambda m}\left(\sum_{l \in \mathrm{m}} e^{U_l/\lambda_m}\right)^{\lambda_m - 1}}{\sum_{n=1}^{N}\left(\sum_{l \in \mathrm{n}} e^{U_l/\lambda_n}\right)^{\lambda_n}} \quad (5)$$

Where *P* is the set of ports and *k* is the analysed one; *N* are the coastlines considered ($n = 1 \ldots N$), in this case two: the Atlantic and the Mediterranean; ($1-\lambda_l$) is a measure of the degree of correlation between the alternatives in nest *m*, i.e. $\lambda_l = 1$ implies no correlation; and $U_k = Z_k + W_m$, where Z_k are the independent variables of the model that describe alternative *k* (they vary over alternatives within nest *m*), and W_m the ones referring to the nest *m* (they differ over nests but not over alternatives within each nest).

Anderson *et al.* (Anderson et al., 2009) applied nested models and showed that the nested structure provided an improvement for the assessment of port choice with US imports. In this paper, we do the same for the Spanish container trades, as the analysed ports are located along unconnected coastlines in the Atlantic and in the Mediterranean. Our database is the Customs Statistics of the Ministry of Finance, which contains all the movements of the Spanish foreign trade. This database is the same as in the previous study, where we analysed the observed traffic distribution through ratios, interpreting them as probabilities. As the estimation of a nested model requires disaggregated data, we have to transform them. This is done by interpreting quantities chosen as survey data and by defining 1,000 tons of cargo as one unit of choice. This way, we consider flows amounting to 10,000 tons as 10 separate choices.

3 The Spanish container trades

The customs statistics contain information of import and export flows by province of origin and destination in peninsular Spain, the trading partner country, the province of the port of transfer in Spain[i], the mode of transport, the mode of shipment and the type, weight and value of cargo. According to customs statistics the volume of containerised seaborne trade generated by mainland Spain and destined for any of the countries included into one of the 8 sets of overseas trade partners was 24 million tons in 2007. With 13.3 million tons exports exceed imports, which amount to 10.7 million tons. Throughput volumes of Barcelona and Valencia dominate the market, while Algeciras located in Cadiz, Bilbao in Vizcaya and Vigo in Pontevedra lag much behind. Apart from said 5 ports it includes the ports of Cartagena in Murcia and Castellón in the province with the same name. See Figure 1. The volume of the 10 port provinces not included amounts to 3% of the total. The volumes of containerised imports and exports by port province are given in Table 1 (see their location by their number on the map -Figure 1-).

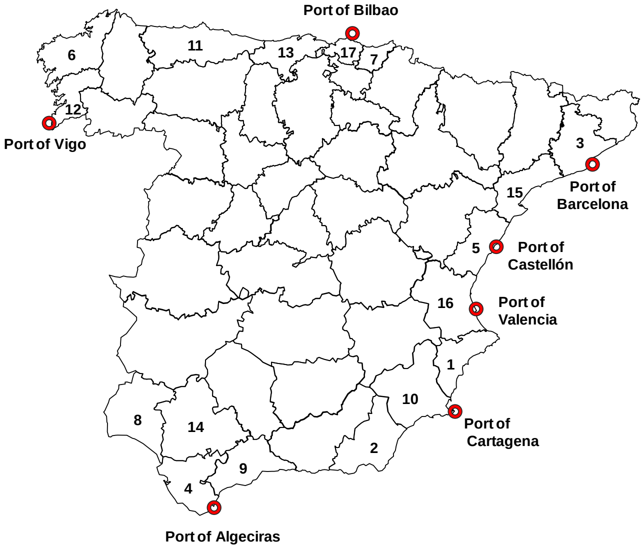
In this study the trade partner countries as given in the customs statistics are grouped into 8 more or less coherent coastal regions that correspond with liner shipping services. The volumes of imports and exports by maritime regions are given in Table 2. The total volume of containerised cargoes for which inland and overseas containerised cargo flows by one of the 7 ports is known, amounts to 11.1 million tons of exports and 10 million tons of imports.

Table 1. Containerised imports and exports by port province in 2007 in 1000 tons

Port province	Imports	Exports	Both	Imports	Exports	Both
1 Alicante	87.1	86.2	173.3	0.8%	0.6%	0.7%
2 Almería	0.3	2.2	2.5	0.0%	0.0%	0.0%
3 Barcelona	3,383.8	3,407.6	6,791.4	31.5%	25.7%	28.3%
4 Cádiz	646.1	631.4	1,277.5	6.0%	4.8%	5.3%
5 Castellón	34.8	737.8	772.6	0.3%	5.6%	3.2%
6 A Coruña	5.7	2.4	8.1	0.1%	0.0%	0.0%
7 Guipuzcoa	0.1	0.8	0.8	0.0%	0.0%	0.0%
8 Huelva	0.0	2.9	2.9	0.0%	0.0%	0.0%
9 Málaga	36.7	13.5	50.1	0.3%	0.1%	0.2%
10 Murcia	68.7	344.2	412.9	0.6%	2.6%	1.7%
11 Asturias	54.5	81.1	135.6	0.5%	0.6%	0.6%
12 Pontevedra	666.5	480.5	1,147.0	6.2%	3.6%	4.8%
13 Santander	7.5	1.6	9.1	0.1%	0.0%	0.0%
14 Sevilla	28.1	162.3	190.4	0.3%	1.2%	0.8%
15 Tarragona	36.0	82.7	118.7	0.3%	0.6%	0.5%
16 Valencia	4,483.4	5,988.3	10,471.7	41.8%	45.1%	43.6%
17 Vizcaya	1,189.7	1,255.8	2,445.6	11.1%	9.5%	10.2%
Total	**10,728.9**	**13,281.2**	**24,010.1**	**100.0%**	**100.0%**	**100.0%**

Source: derived from Spanish customs data

Figure 1. Location of ports and provinces



It should be noted that port throughput volumes based on customs differ from those collected by the port authorities. Escamilla *et al.* (Escamilla-Navarro et al., 2010) present a proposal to improve statistics by integrating both sources. To compare both sources for the Spanish container trades container throughput according to the Spanish Port Statistics (2007) has to be corrected by subtracting containers transhipped. These throughput volumes are assessed by taking the shares of transhipment cargo as published by Drewry (Drewry, 2009) and correcting the total throughput totals for 2007 for the ports of Algeciras, Barcelona and Valencia. It appears that for all ports together the volume of Table 2 is about one third of the volume according to the port statistics. This means that our analysis is based on a sample of about one third, assuming of course that port data represent the right volume.

Table 2. Containerised imports and exports by overseas trade partner region in 2007 in 1000 tons

		Exports	Imports	Total
1	West Mediterranean	1,207	105	1,312
2	East Mediterranean	1,386	833	2,219
3	West/South Africa	833	420	1,253
4	Arabian Sea Area	1,757	227	1,984
5	East Asia	2,415	5,437	7,852
6	North America	2,562	1,443	4,005
7	South America EC	680	1,001	1,681
8	South America WC	271	559	830
	Total	**11,111**	**10,024**	**21,135**

Source: derived from Spanish customs data

4 Attributes of the model

4.1 Variables taken from previous study

4.1.1 Inland transport cost

The transport of containers between the Spanish provinces and sea ports mainly takes place by road. For the biggest container ports of Valencia and Barcelona, where rail transport has the biggest chance to be used (CIM, 2009) the share of road is about 95%. For other ports the share most likely is even less. The costs of trucking depend on many variables such as distance, type of truck and availability of return cargo.

For this analysis we take the distance by road between the provincial gravity point and the seaport as basis and multiply it with the same price per ton-kilometre of € 0.085 (European Commission DG TREN, 2006) and is adapted for 2007 (Transport in Figures, 2007).

4.1.2 Maritime transport cost

Maritime transport concerns the transport between the gravity point of the overseas regions and seaports and consists of direct shipments and shipments via transhipment ports. The cost of maritime transport is reflected in liner shipping rates which are based on shipping costs and the related demand supply situation on the liner shipping market. The costs of a round trip connecting the port of origin or destination in Spain with the port of the overseas trade partner depend on the roundtrip distance, characteristics of the ship such as size and speed, roundtrip characteristics such as number of ports of call at each coast line, port productivity and the volume of containers carried. Important also are the ratios of inbound and outbound cargoes and the position of the route in the world container shipping network. For Spain this means for instance that trade links with the Far East are offered by shipping services passing Spain on their way between West Europe and the Far East and by services connecting the Mediterranean and the Far East.

The ports of Barcelona, Valencia and Algeciras are all well connected and have direct liner services with the main parts of the world. This concerns the east west trade routes connecting the major industrial centres of the world east of the Suez Canal and North America and also the north-south trade routes with Latin America and South and West Africa. Starting point of the assessment of liner freight rates for 2007 are the liner freight rates (Drewry, 2007). These freight rate benchmarks are for full container loads and include the base ocean rate, terminal handling charges at origin and destination and fuel and other surcharges. The freight rates are given for 20 ft and 40 ft containers. For the conversion into tons a TEU/box ratio of 1.6 and 14 tons per loaded TEU are applied, based on Spanish port data. Port costs are included in maritime costs and not taken explicitly.

Information of the liner shipping services as at July 2007 (Drewry, 2009) shows that Algeciras, Barcelona and Valencia are called at by all the major east-west and north-south shipping lines, while the other ports are called at by one or two services or not at all. In these cases transhipment is needed: for Cartagena and Castellon via ports in the West Mediterranean and for Bilbao and Vigo via ports in Northwest Europe. This means that full feeder costs have to be added or, if such ports are called at by one or two services, part of the feeder costs. The additional costs of feeder transport and the cost of short sea shipping, as applies for the intra-Med trades, are based on roundtrip costs calculations.

4.1.3 Other cost and quality of service aspects

There is an extensive discussion as to what factors impact the choice of routing[ii], such as transport costs, transit time, frequency of service and reliability of service. User surveys dating back to the 1980s show that quality of service aspects are important (see, for example, (Peters, 1989) and (Collison, 1984). This research is refined by the use of the analytical hierarchy process method to analyse survey data (such as by (Lirn et al., 2004) and (Song and Yeo, 2004).

Zhang (Zhang, 2008) mentions that a larger hinterland of a port allows for:

i) a larger size of ships being attracted thus realising economies of ship size (Jansson and Shneerson, 1982);

ii) higher frequencies of service resulting in Mohring effects (Scherer, 1980) and (UNITE, 2003);

iii) stronger roles as load centres;

iv) better availability of third party logistic service providers and

v) more value added clusters (Langen, 2004).

Taken together these effects can be referred to as hub port effects.

Above-mentioned user surveys include competition aspects at a detailed sometimes operational level as it applies to specific ports, operators and carriers. They are useful for operators to strengthen their market share by improving their competitive edge. Our analysis is at a more abstract level and disregards the performance of individual operators. Instead it concentrates on cost parameters derived from the average performance of port and liner shipping operators, such as the port's place in the liner shipping network and economies of ship size.

The basic costs considered are those featuring in transportation planning models on modal split and route choice and include transport costs, transit time, frequency of services and all quality of service aspects such as service reliability impacting port choice as experienced by users and producers of shipping services and which are included in the utility function. Transit time is not included as it is practically proportional to transport costs, while precise information is lacking. Service reliability is not included as it averages out between operators.

Important, however, are the so called Mohring effects (Scherer, 1980). Users of transport facilities impact the situation of other users. In the negative case, e.g. when the activity of one user causes extra costs for others, we talk of congestion costs. In the positive case, when users' activities improve the welfare situation of other users we talk about the 'Mohring effect'. The project UNIfication of accounts and marginal costs for Transport Efficiency (UNITE, 2003), is a project under the Fifth Framework package by the European Commission, where quantitative assessments are made of Mohring effects through case studies concerning passenger and freight transport. Logit models were used in particular with respect to freight transport.

Mohring effects are related to frequency of service or headway[iii]. At the level of a route the frequency can be calculated as the average of number of sailings per unit of time. At the level of a port, being linked by more than one route and being served by more than one shipping line, the frequency can be equated to the average over all routes and services. This average comes close to Mohring effects measuring quality of service aspects related to the level of trade.

In this paper, the Mohring variable is expressed by 1 minus the inverse of container throughput of the ports in 500,000 tons, where throughput concerns both import, export and transhipment containers, as is shown in Table 3.

Table 3. Mohring variable

	Port	Mohring index
1	Algeciras	0,9882
2	Barcelona	0,9803
3	Bilbao	0,9155
4	Cartagena	0,0476
5	Castellon	0,6232
6	Valencia	0,9846
7	Vigo	0,7943

Source: (Puertos del Estado, 2007) and (Drewry, 2009).

4.2 Other variables introduced in this paper

4.2.1 Choice of coast line

The Spanish ports analysed are located along different coast lines. It is possible that the ports located on the same coast line share some attributes, so that the IIA hypothesis would not hold anymore. Therefore, we propose to use a nested model with two nests, one for each of the two main Spanish peninsular coast lines: Atlantic and Mediterranean. The ports included in the Atlantic nest are Vigo and Bilbao, while the ones of the Mediterranean nest are Barcelona, Valencia, Castellón, Cartagena and Algeciras.

4.2.2 Inland transport cargo balance

In the previous study inland trucking costs were taken as a linear function of the distance between province and seaport disregarding the situation of the cargo balance, which has an impact on trucking costs. The impact of an additional variable reflecting the cargo balance situation, is tested. For non-containerised cargoes Walters (Walters, 1960) derived a formula by assessing the impact of the cargo balance on trucking costs per front and back-haul part of the trip, using a probabilistic model. Talley (Talley, 1986) refined this formula.

The role of the cargo balance situation on maritime costs/freight rates is studied by Singer *et al.* (Singer et al., 2007), who developed a model based on Nash bargaining. They deducted that the ratio of (rates paid from A to B)/(rates paid from B to A) depends linearly on the ratio (flows from B to A)/(flow from A to B). They empirically verified this ratio with data from the world three major maritime trade routes.

Blonigen and Wilson (Blonigen and Wilson, 2006) tested the impact of cargo imbalance on the US import and export on maritime freight rates and took the logarithm of the difference between imports and exports as a measure. The imbalance indicators appeared not to be statistically significant. Clark *et al.* (Clark et al., 2004) tested the impact for US import and export trades by taking the difference between both divided by the sum of both, which appears to be statistically significant in half of the models they tested. Wilmsmeier and Martinez Zarzosa (Wilmsmeier and Martínez-Zarzosa, 2010) analysed maritime transport costs of intra-Latin American trades with regression analysis and used trade imbalance, defined in a way comparable to Clark *et al.* (Clark et al., 2004), which variable appeared to be statistically significant in some of their models.

The costs of empty return trips can be reduced by combining them with other loaded trips. For network effects we can refer to the literature of minimising transport of empty containers, which generally concentrates on the reduction of total system costs rather than on price setting for containers per origin-destination pair. For a literature review on the matter see, for example, (Yur and Esmer, 2011).

For Spain there is no published information on trucking costs per combination of port and inland origin or destination of containers. We therefore apply a variable specific for the directional cargo imbalance, which we define in three different ways:

$$BH1 = \left|\frac{X-I}{X+I}\right| \qquad (6)$$

$$BH2 = \begin{cases} \left|\frac{X}{X+I}\right| for\ imports \\ \left|\frac{I}{X+I}\right|\ for\ exports \end{cases} \qquad (7)$$

$$BH2 = \begin{cases} \left|\frac{X}{X+I}\right| for\ imports \\ \left|\frac{I}{X+I}\right|\ for\ exports \end{cases} \qquad (8)$$

The symbols *X* and *I* represent export and import flows respectively.

4.2.3 Variables split into a distance dependent and independent part

Veldman and Buckmann (Veldman and Bückmann, 2003) and Veldman *et al.* (Veldman et al., 2005) tested port choice models for Northwest European ports and split their variables in a distance dependent and independent part. This appeared to give a considerable improvement in terms of model fit. We also tested models with split variables for the non-nested model versions, which appear to give no improvement.

5 Model testing

5.1 Comparison nested versus non-nested version

In the previous study we used regression analysis to estimate the coefficients of the logit model according to equation (4) and obtained R-square values of 37.7% and 35.2% for exports and imports respectively, as given in Table 4 and 5 under the heading *Previous study*. The best results, in terms of R-square values of 73.5% and 73.1% respectively, were obtained for the model with port specific dummy variables, instead of the Mohring variable Mo (see the Table 5 of the previous study).

This time we transform the aggregate data into disaggregate data, as described in Section 2.2.2, and estimated the model coefficients of the base model with the statistical package "R" (R Core Team, 2010). See *Base model* under the heading *Nested models* in Table 4 and 5. To check the validity of the applicability of a nested structure, we compared it against the non-nested alternative with the Likelihood Ratio test (see equation 9), resulting in a strong difference between both: 448.06 for exports and 21.45 for imports, relative to a 95 per cent critical value of 5.99. We can conclude that the nested structure fits our data better.

$$LR\ test = 2\ (lnL_{non-nested} - lnL_{nested}) \quad (9)$$

5.2 Nested models tested

Given the superiority of the nested structure for the base model we tested the same base model adding a variable reflecting the Backhaul effect (BH1) and further this model by replacing the Mohring effect variable (Mo) with port specific dummy variables.

For the *Backhaul variable* the specification according to equation 6, i.e. with variable BH1, gives clearly the best result. Both for exports and for imports, the improvement when adding the backhaul variable is negligible in terms of the "hit rate". For exports, the Log-Likelihood statistic shows an important reduction, which can be interpreted as that the model with that new variable fits the data equally well, but is more robust (see Table 4).

In the case of imports, Log-Likelihood improvement is not important. The results using dummy variables are slightly better than those of the version with the Mohring variable in terms of Log-Likelihood. However, the explanatory power of the model is less as the dummy variables represent unknown factors contributing to the appeal of a port (see Table 5).

The inclusive values resulting from the nested structure (iv.Atl. and iv.Med.) indicate that, for exports, the Mediterranean coast seems more appealing than the Atlantic one, whereas for imports both look similar but with a slight preference for the Atlantic coast

Table 4. Test results of non-nested and nested models for exports

	Previous study		**Nested models**					
	Base model		**Base model**		**Base and Backhaul**		**Port dummies**	
	Coeff	**t**	**Coeff**	**t**	**Coeff**	**t**	**Coeff**	**t**
iv. Atl.			0.176	13.6	0.378	3.5	0.340	3.0
iv. Med.			0.470	39.0	0.524	36.0	0.560	31.5
LC	-0.133	-21.8	-0.067	-43.8	-0.068	-41.3	-0.069	-37.2
MC	-0.116	-6.3	-0.060	-20.2	-0.070	-22.0	-0.068	-16.2
Mo	15.400	29.3	1.107	19.3	1.659	20.7		
BH1					0.845	11.1	1.022	11.0
D.Alg							-0.528	-6.5
D.Barc							-0.313	-8.3

D.Bilb							-0.223	-2.2
D.Cart							-1.521	-17.1
D.Cast							-0.944	-19.7
D.Vigo							-0.819	-3.5
R-square	0.377		n.a.		n.a.		n.a.	
Log-Likelihood			-4335.5		-4192.4		-4084.2	
Hit rate			85.3%		85.4%		85.4%	

Table 5. Test results of non-nested and nested models for imports

	Previous study		**Nested models**					
	Base model		**Base model**		**Base and Backhaul**		**Port dummies**	
	Coeff	**t**	**Coeff**	**t**	**Coeff**	**t**	**Coeff**	**t**
iv. Atl.			0.926	16.3	0.927	16.0	0.915	11.1
iv. Med.			0.845	30.4	0.844	30.1	0.865	26.7
LC	-0.161	-22.5	-0.077	-33.6	-0.078	-33.5	-0.079	-29.3
MC	-0.109	-5.6	-0.051	-13.4	-0.051	-13.4	0.001	0.1
Mo	15.700	25.0	4.459	23.6	4.603	22.5		
BH1					0.171	1.1	0.827	5.0
D.Alg							-0.449	-4.4
D.Barc							-0.469	-8.2
D.Bilb							-1.730	-5.2
D.Cart							-3.836	-15.6
D.Cast							-4.434	-17.0
D.Vigo							-2.438	-6.1
R-square	0.352		n.a.		n.a.		n.a.	
Log-Likelihood			-3200.9		-3199.5		-2944.7	
Hit rate			91%		91%		91.4%	

6 Conclusions

Earlier research on port competition with Spanish container trades with non-nested logit models showed that inland transport costs, maritime transport costs, port specific dummy variables and proxies for quality of service aspects have a statistically significant impact. The objective of this paper is to improve the model by testing a two-phase choice function (first coastline - then port) by using a nested model, by including a new variable to indicate the inland container transport balance and by modified variables reflecting the impact if distance.

The test of two-phase choice function (first coastline - then port) by using a nested model appeared to lead to an improvement, resulting in more accurate model coefficients. The introduction of a variable reflecting the inland container transport balance showed mixed results: a slight improvement for container export flows and no improvement with imports.

Statistical tests on port choice with respect to North Sea container ports showed that the adoption of variables split into a distant dependent and a distance independent part lead to a considerable improvement. We also tested models with split variables for the conditional model versions, which clearly appeared to give no improvement. We therefore decided to do no further tests.

The resulting model outcomes can be used to predict container port market shares and help to support the assessment of economic benefits and the effect of port revenues resulting from port interventions.

Future research should concentrate on a disaggregation of input data such as by type and value of container cargo and a refinement of information on logit model attributes such as inland transport costs and freight rates. To this end interviews are needed with shippers and carriers.

Acknowledgements

The authors thank the support provided by the Government of Spain, FEDER under projects MTM2008-06778-C02-01 and MTM2011-23205.

References

Anderson, C.M., Opaluch, J.J., Grigalunas, T.A., 2009. The demand for import services at US container ports. Maritime Economics and Logistics 11, 156-185.

Blonigen, B.A., Wilson, W.W., 2006. International Trade, Transportation Networks and Port Choice.

CIM, 2009. From road to rail. Containerisation International Monthly, December.

Clark, X., Dollar, D., Micco, A., 2004. Port efficiency, maritime transport costs, and bilateral trade. Journal of Development Economics 75, 417-450.

Collison, F.M., 1984. Market Segments for Marine Liner Service. Transportation Journal 24, 40-54.

Drewry, S.C., 2007. Container Freight Rate Insight, Bi-Monthly Pricing Benchmarks on the Container Market (No. Six different issues over the year). Drewry Shipping Consultants.

Drewry, S.C., 2009. Container Forecaster Annual – 2007/08. Drewry Shipping Consultants.

Escamilla-Navarro, L., García-Menéndez, L., Pérez-García, E., 2010. Integration of foreign trade and maritime transport statistics in Spain. Maritime Policy & Management 37, 347-375.

European Commission DG TREN, 2006. Analysis of the impact of oil prices on the socio-economic situation in the transport sector. Final Report. ECORYS Transport (Netherlands) and Consultrans (Spain).

Garcia-Alonso, L., Sanchez-Soriano, J., 2009. Port selection from a hinterland perspective. Maritime Economics and Logistics 11, 260-269.

Hensher, D.A., Button, K.J., 2000. Handbook of transport modelling / edited by David A. Hensher, Kenneth J. Button. Elsevier Science B.V., New York.

Jansson, J., Shneerson, D., 1982. Port Economics. The MIT Press. Cambridge, Massachusetts.

Langen, P.D., 2004. Governance in Seaport Clusters. Maritime Economics & Logistics 6, 141-156.

Lirn, T.C., Thanopoulou, H.A., Beynon, M.J., Beresford, A., 2004. An Application of AHP on Transhipment Port Selection: A Global Perspective. Maritime Economics and Logistics 6, 70-91.

Malchow, M., Kanafani, A., 2001. A disaggregate analysis of factors influencing port selection. Maritime Policy and Management 28, 265-277.

Malchow, M., Kanafani, A., 2004. A disaggregate analysis of port selection. Transportation Research. Part E 40, 317-337.

McFadden, D., 1974. The Measurement of Urban Travel Demand. Journal of Public Economics 3.

Nir, A.S., Lin, K., Liang, G.S., 2003. Port choice behaviour from the perspective of the shipper. Maritime Policy and Management 30, 165-173.

Ortuzar, J. de D., Willumsen, L.G., 1990. Modelling transport. John Wiley and Sons, England.

Oum, T.H., 1989. Alternative Demand Models and Their Elasticity Estimates. Journal of Transport Economics and Policy 23, 163-187.

Paixao Casaca, A., Carvalho, S., Oliveira, M., 2010. Port choice in the European short sea shipping market: the view of point of port authorities, in: IAME 2010 Annual Conference Proceedings.

Peters, H.J., 1989. SeaTrade, Logistics and Transport (No. 8185). The World Bank.

Puertos del Estado, 2007. Anuario estadístico. Madrid.

R Core Team, 2010. R: A Language and Environment for Statistical Computing.

Robinson, R., 2002. Ports as elements in value-driven chain systems: the new paradigm. Maritime Policy and Management 29, 241-255.

Scherer, F.M., 1980. Industrial Market Structure and Economic Performance. Rand McNally College, Chicago.

Singer, M., Pardo, I., Sánchez, R.J., 2007. Is applicable Nash bargaining rates to the evaluation of freight rates?, in: IAME 2007: Challenges and Trends in Shipping: Markets, Investments and Policies. Athens.

Song, D.W., Yeo, K.T., 2004. A Competitive Analysis of Chinese Container Ports Using the Analytic Hierarchy Process. Maritime Economics and Logistics 6, 34-52.

Talley, W.K., 1986. A rule for allocating joint truck-carrier costs. Transportation Research Part B: Methodological 20, 49-57.

Tiwari, P., Itoh, H., Doi, M., 2003. Shipper´s port and carrier selection behaviour in China: a discrete choice analysis. Maritime Economics and Logistics 5, 23-39.

Train, K.E., 2009. Discrete choice methods with simulation, 2nd ed. Cambridge University Press.

Transport in Figures, 2007. In Dutch: Transport in cijfers, editie 2007.

UNITE, 2003. UNIfication of accounts and marginal costs for Transport Efficiency. Project funded by the EC under the Fifth Framework Transport RTD (1998-2002).

Veldman, S., Bückmann, E., 2003. A model on container port competition: an application for the West European container hub-ports. Maritime Economics and Logistics 5, 3-22.

Veldman, S., Bückmann, E.H., Saitua, R.N., 2005. River Depth and Container Port Market Shares: The Impact of Deepening the Scheldt River on the West European Container Hub-Port Market Shares. Maritime Economics & Logistics 7, 336-355.

Veldman, S., Garcia-Alonso, L., Vallejo-Pinto, J.Á., 2011. Determinants of container port choice in Spain. Maritime Policy & Management 38, 509-522.

Veldman, S., Gopkalo, O., 2010. A Model on Port Basin Choice with Russian Container Trade, in: International Forum on Shipping, Ports and Airports (IFSPA) 2010: Integrated Transportation Logistics: From Low Cost to High Responsibility. Chengdu (Sichuan, China), pp. 409-420.

Walters, A.A., 1960. The Allocation of Joint Costs with Demands as Probability Distributions. The American Economic Review 50, pp. 419-432.

Wilmsmeier, G., Martínez-Zarzosa, I., 2010. Determinants of maritime transport costs – a panel data analysis for Latin American trade. Transportation Planning and Technology 33, 105-121.

Winston, C., 1981. A Multinomial Probit Prediction of the Demand for Domestic Ocean Container Service. Journal of Transport Economics and Policy 15, 243-252.

Woo, S.-H., Pettit, S.J., Kwak, D.-W., Beresford, A.K.C., 2011. Seaport research: A structured literature review on methodological issues since the 1980s. Transportation Research Part A: Policy and Practice 45, 667-685.

Yur, T., Esmer, S., 2011. A review of the studies on empty container repositioning problem, in: ECONSHIP 2011. University of the Aegean, Department of Shipping, Trade and Transport, Chios.

Zhang, A., 2008. The impact of hinterland access conditions on rivalry between ports. Discussion paper 2008-8. Joint Transport Research Centre, OECD and ITF.

---

[i] We assume that the province where the customs process is realised, is where the port of transfer is located.

[ii] Routing is defined here as the door-to-door shipment of imports and exports including land transport, maritime transport and transfer in ports.

[iii] Headway is also referred to as inter-arrival time. Frequency of calls per year equals 365/headway.