Department of Economics
School of Social Sciences

Econometric Estimation of Scale and Scope Economies Within the Port Sector: A Review

Beatriz Tovar[1]
DAEA, Universidad de Las Palmas de Gran Canaria

Sergio Jara-Díaz
Departamento Ingeniería Civil Universidad de Chile[2]

Lourdes Trujillo[3]
CCRP, City University of London and DAEA, Universidad de Las Palmas de Gran Canaria

Department of Economics
Discussion Paper Series
No. 07/04

[1] Email: btovar@empresariales.ulpgc.es

[2] Email: jaradiaz@cec.uchile.cl

[3] Department of Economics, City University, Northampton Square, London, EC1V 0HB, UK. Email:ltrujillo@daea.ulpgc.es

ECONOMETRIC ESTIMATION OF SCALE AND SCOPE ECONOMIES WITHIN THE PORT SECTOR: A REVIEW.

Beatriz Tovar *

btovar@daea.ulpgc.es

Phone: +34 928 45 17 94 and Fax: +34 928 45 81 83

DAEA, Universidad de Las Palmas de Gran Canaria

Sergio Jara-Díaz

jaradiaz@cec.uchile.cl

Phone: (56-2) 678 4380 and Fax: (56-2) 689 4206

Departamento Ingeniería Civil

Universidad de Chile

Lourdes Trujillo

ltrujillo@daea.ulpgc.es

Phone: +34 928 45 18 09 and Fax: +34 928 45 81 83

DAEA, Universidad de Las Palmas de Gran Canaria and

CCRP, City University of London

Abstract

Seaports provide multiple services to ships, cargo, and passengers. Size of the port and type of service are two key elements when deciding whether competition is feasible and how to promote it, or conversely regulation is needed. Analyzing this requires a profound knowledge of the cost structure of the activity involved. This means not only knowing total costs for different volumes of aggregated traffic, but also the behavior of costs when part of the bundle is produced, i.e. when the mix changes. The optimal organization of the industry can be studied by means of cost and production functions. This paper offers a review of the relatively scarce literature about econometric ports' cost structure, and highlights the role of the multioutput approach as the correct one because allows the calculation of key cost indicators (economies of scale, scope, and so forth) to determine the optimal port industrial structure for a given forecast of demand (traffic mix and volume).

Key words: multiproduct, economies of scale and scope, port and cargo handling.

JEL Classification system: L9

1. Introduction

The estimation of key indicators representing the technical production properties of firms within an industry, such as economies of scale and scope, plays an essential role in the determination of the optimal industrial organization, i.e. that which induces the best assignment of resources. As known, technical properties can be analyzed directly through the study of the relations between inputs and outputs by means of production or transformation functions. They can be analyzed as well from the estimation of cost functions, taking advantage of the dual properties of these functions regarding technology. The estimation of production and cost functions represents a fundamental tool for the proper regulation of any industry, contributing with precise results to the regulatory process [1].

On the other hand, port activities are particularly important within an increasingly interacting world economy. Ports are a key component of the logistics chain and, therefore, their operation has a direct effect on relevant economic variables such as export competitiveness and final import prices, thus affecting economic development.

Economic activity within a port has multiple dimensions. They encompass from the administration provided by the Port Authority to pilotage, towage, supply of utilities such as water and power, cargo handling, catering, ship repair, and so on. Among those services, cargo handling requires special attention, as it means more than 80% of the bill of a vessel [2] that arrives to a port for loading and unloading. Both the provision of port facilities and cargo handling in ports are multioutput activities, as freight can arrive in many forms like containers, bulk, rolling stock, or non-containerised general cargo.

In most countries, the port industry is subject to some kind of control by the public sector, who has to plan future investments, has to regulate the operation of cargo handling firms and has to suggest optimal tariff structures, among other functions. In this context, the analysis of

the production and cost structure of the provision of port facilities and of the operation of firms is a helpful tool to help inducing policies for the optimal operation of port facilities.

There are three main reasons why it is useful to know the cost structure. First, estimating marginal costs as well as global and specific economies of scale is an essential tool for tariff cap regulation since it provides enough information in order to decide whether the application of marginal cost tariffs is feasible. Second, calculation of economies of scope provides information on whether it is advisable to have the firms diversified or whether they should be specialized instead, thus feeding with objective information a decision that might carry cost savings when the right choice is made. Lastly, the capacity to measure the degree of subadditivity (a measure of natural monopoly) provides a proper tool to decide the optimal intraport structure in terms of the adequate number of firms.

Furthermore, costs analysis also permits the evaluation of ports' return and productivity by calculating different indicators. Additionally, it allows comparisons of the productive efficiency among various firms and along time for a single firm. It is important to note, however, that this latter aspect has prompted a new branch of methodological research that has been applied to port analysis, in which cost minimization is not assumed, such that inefficient firm behaviour is allowed: the estimation of frontiers. The concept can be easily illustrated for a given amount of inputs or production factors. If one or more of the observed output components could be increased holding the other constant, then this point is inferior or inefficient. Analogously, an observed input combination is said to be inferior or inefficient in the production of a given output if one or more inputs could be diminished. Inefficiency translates into observed costs that are not minimum, which makes the association between expenses and production incorrect. The new technique is aimed at estimating the non-inferior sets, also called the frontiers.

Frontiers are essentially envelopes which can approximated from various methodologies from

the observed data on outputs or costs and the corresponding data on input quantities or input prices observed at the business unit level—either a port authority, a terminal, or any unit of interest to the analysis. The reveals those units which operate at efficient levels of production or costs and of those who do not and hence do not fulfill the assumption of maximization of production or minimization of costs. The first author who suggested the use of the borders of production for the efficiency analysis was Farell [3]. According to this author, the correct form to measure the efficiency was by means of the comparison of each observation (commonly companies) with the best observed practice.

Broadly, two different techniques are used to carry out efficiency studies. Both aim at estimating frontier functions but they differ in the way the frontier is estimated. The first uses non stochastic and non parametric mathematical programming methods. Within these optimization methods, the most popular was introduced by Charnes, et al. [4] and is known as *Data Envelopment Analysis* (DEA). For a detailed analysis of this mathematical method, the interested reader is referred to Charnes et al. [5].One of the main advantages of the approach is that it can yield results with a relatively modest set of data on the various business units to be analyzed. This is probably one reason why the technique has been used quite extensively in the economic analysis of ports efficiency, e.g. Roll and Hayuth [6], Tongzon [7], Wang et al [8], Estache et al. [9] and Cullinane et al. [10].

The second technique is econometric modelling. Stochastic frontier modelling is the most common approach among these models. It relies on stochastic parametric regression-based methods. This technique starts with Aigner, et al. [11] and Meeusen and van de Broeck [12]. The main intuition in this approach is that the deviations from the frontier—i.e. best practice—is not totally under the control of the business unit analyzed because of policy, political, macroeconomic, institutional or other reasons. It has been applied to ports, among others, by Notteboom et al. [13], Estache et al. [14], Cullinane and Song [15], Rodríguez-

Álvarez et al. [16] and Cullinane et al. [17]. A full review of the literature on port efficiency can be seeing in Cullimane et al., [18], Wang et al., [19] or González and Trujillo [20].

This paper focuses on the contribution of the literature towards the econometric estimation of traditional production and cost functions in the port industry and provides a detailed and critical analysis of the relevant methodological aspects and empirical results. The paper is structured as follows: Section 2 provides an analysis of the theoretical aspects related to the econometric estimation of production and cost functions. Section 3 contains a detailed review of the economic literature dealing with the econometric estimation of production and cost functions in the port industry. Lastly, section 4 summarises the main conclusions.

2. Econometric aspects for the estimation of production and cost functions.

The first objective of the applied production analysis is the empirical measurement of economically relevant information, enabling the thorough description of economic agents' behaviour, either from production or cost functions. For smooth technologies this includes function value (for example, cost level or scalar production), function gradient (for example, derived demands or marginal costs) and the Hessian (for example, the elasticity matrix of derived demands or the derivatives of marginal costs). Therefore, when choosing the functional form to be used for the empirical estimation, the aim should be to choose a specification with enough parameters to enable the analysis of all these effects without imposing a priori restrictions.

The functional form used in the first empirical papers on production function estimations in the port sector was the Cobb-Douglas function, a functional form that has been widely used in the literature to evaluate scale effects, as the degree of scale economies can be easily computed parametrically as the inverse of the summation over the function exponents. The Cobb-Douglas functional form -although easy to estimate- presents important limitations.

This function belongs to the homogenous functions group and, therefore, it restrains the way in which scale effects and elasticities of substitution take place.

There are other functional forms, which do not present these limitations. Thus, the constant elasticity of substitution function (CES) is the natural extension of Cobb-Douglas function since it allows the elasticity of substitution to take values different from the unit. The following obvious step is to generate a function allowing the elasticity of substitution to change when the product or the proportion of productive factors used varies. The production function enabling these two generalizations is the translogarithmic function.

The estimation of a production function in the case of multiproduction is more complex, since the scalar representation $Y=f(X)$ has to be changed for more a general analytical treatment as $F(X,Y)=0$, which is usually approached assuming separability between inputs and outputs to enable representations as $g(Y)=f(X)$. In these cases, if data permits, it is very advantageous to estimate a cost function $C(w,Y)$ instead, since it relates the cost of obtaining a certain production with the product itself Y (which can be a vector) and the prices of the productive factors used, w. As known, duality theory enables the empirical study of production structure through cost function estimation; in other words, the properties of F can be studied from an estimated C function.

Most empirical studies apply cost functions that are twice differentiable with respect to productive factors prices. Under these conditions, an important property of the cost function known as Shephard's lemma can be applied, which allows the generation of factor demand functions as the derivative of the cost function with respect to each factor's price. The system's estimation, through cost function and the derived factor demand functions, generates more efficient parameter estimations than those obtained with only the cost function. This represents the most practical advantage of a cost function [12].

On the other hand, the functional form selected to represent a cost function needs to meet certain regularity conditions to ensure that it is a true cost function, i.e. a function consistent with the idea of achieving a certain production volume at the minimum expense, at given factor prices. It is widely known that the appropriate functional form representing a cost function must be non-negative, linearly homogenous, concave and non-decreasing in factor prices. Furthermore, a cost function must be non-decreasing in outputs when assuming free availability.

In addition to these generic conditions, a cost function must meet other requirements if it is to be used in the estimation of a multi-product process [13]. First, the function must provide reasonable cost figures for product vectors with some component at zero level, since not all the firms need to produce all industry products. The Cobb-Douglas function and the translog function, for example, violate this condition [14]. Secondly, the function must not prejudge the presence or absence of any cost property playing an important role in the analysis of the industry. On the contrary, the functional form must be consistent with the satisfaction or violation of those properties, so the empirical results obtained arise from the data and not as a consequence of the selected functional form. This property is called substantive flexibility of the functional form. Thirdly, the function must not require the estimation of an excessive number of parameters and, lastly, it must not impose restrictions on the value of the first and second partial derivatives. The Cobb-Douglas functional form applied to cost functions represents an example of the violation of some conditions since the form itself imposes that weak cost complementarity can not be detected, regardless of the fact that reality may be different. In fact, if there were cost complementary, an estimation using Cobb-Douglas function would produce biased marginal cost estimates. This is the reason why the Cobb-Douglas function is not suitable for multi-productive cost function estimations.

Therefore, based on the foregoing, it is clearly preferable to use functional forms which avoid restrictions imposed by the functional form itself -such as the so-called flexible functional forms- developed on the basis that they provide a good local approximation of a twice differentiable arbitrary function [15]. Moreover, this allows empirical contrast of additional restrictions, such as homogeneity, homotheticity, separability, constant returns to scale and constant elasticities of substitution, directly from the data instead of them being imposed a priori [16].

Caves et al. [17] indicates three problems that may turn the flexible functional forms used empirically less attractive, namely: the violation of regularity conditions in the production structure, the estimation of an excessive number of parameters and the impossibility to work with observations at zero production levels (in the case of a translog form). From this perspective, the most frequently flexible functional forms used in the port sector are analysed below: the quadratic function and the translogarithmic function.

The quadratic function, first proposed by Lau [18], is a Taylor expansion of second order and is consistent with the following equation:

$$C = \alpha_0 + \sum_i^m \alpha_i y_i + \sum_i^n \beta_i w_i + \frac{1}{2}\sum_i^m \sum_j^m \delta_{ij} y_i y_j + \frac{1}{2}\sum_i^n \sum_j^n \gamma_{ij} w_i w_j + \sum_i^m \sum_j^n \rho_{ij} y_i w_j \qquad (2)$$

where C is the total cost, $Y = (y_i)$ is the output vector, $W=(w_j)$ is the factor prices vector and, α_0, α_i, β_i, δ_{ij} γ_{ij}, and ρ_{ij} are parameters to be estimated. The variables are usually deviated around a meaningful point (point of approximation). The translog function is a quadratic form where variables have been expressed in logarithms.

Both flexible functional forms present advantages and drawbacks. Then, the selection between them depends on the policy issue and context at stake. One of the advantages of the quadratic function is that it is well defined for zero values so it allows consideration of those

cases in which some output vector element is nil, i.e. it enables the analysis of economies of scope and incremental costs [19]. On the contrary, the translog function does not allow zero values and then, it is not suitable for the study of economies of scope unless a proper output transformation is applied, such as a Box-Cox transformation.

Another advantage of the quadratic function is that it allows the direct determination of marginal costs considered at the approximation point α_i, and Hessian values δ_{ij}, which result essential for the subadditivity analysis [20]. On the other hand, there are two disadvantages generally mentioned about the quadratic function. The first is that it is not possible to make sure that linear homogeneity in prices is met. However, this condition can be imposed by normalizing cost function by one factor price. The second is that the cost function is very strict in regard with the specification of fixed costs whose effect should be captured by a single parameter (α_0 if variables are not deviated). The problem with this is that, in fact, fixed costs may vary depending on which subset of total products group is being produced. In order to solve this issue and give the functional form the capacity to capture these differences in fixed costs that may arise among firms producing different groups of products, dummy variables, F_i, are introduced. The value of these variables is represented by the unit when there is some production of product *i*, or zero otherwise. This leads to the following flexible fixed costs quadratic function [21]:

$$C = \alpha_0 + \sum_i^m \alpha_i y_i + \sum_i \alpha_i F_i + \sum_i^n \beta_i w_i + \frac{1}{2}\sum_i^m \sum_j^m \delta_{ij} y_i y_j + \frac{1}{2}\sum_i^n \sum_j^n \gamma_{ij} w_i w_j + \sum_i^m \sum_j^n \rho_{ij} y_i w_j \qquad (3)$$

In contrast, the translog function's main advantage is that it allows the analysis of the underlying production structure, such as homogeneity, separability, economies of scale, etc. through relatively simple tests of an appropriate group of estimated parameters. Its first order coefficients at the approximation point are the cost-product elasticities calculated at this approximation point (usually the mean) in a manner such that their addition represents an

estimation of the inverse of the degree of economies of scale.

The number of parameters to be estimated is larger in the quadratic function than in the translog [17] and this is so because the restrains imposed on the translog function to ensure it fulfils the conditions of homogeneity of degree one in factor prices, symmetry, etc., limit the number of free parameters to be estimated. Even though ordinary least squares could be used to estimate any one of these two functions, additional information can be used to improve the estimate efficiency (i.e. applying Shephard's lemma). This information can be factor demands, factor expenditures or factor shares. The maximum likelihood method can be used to estimate the unknown parameters, specifying that n+1 equations bear normal additive errors. Although cost functions can be estimated alone, it is clearly more efficient to estimate parameters from the n+1 equations system [22].

Lastly, it is worth mentioning that estimating these functions with the right hand variables deviated from an appropriate approximation point (generally the sample mean), represents usual practice in the empirical work and there are basically two reasons for this. First, this results in an immediate estimation of the gradient at the approximation point (marginal costs and factor demands). Second, multicollinearity between linear, square and cross terms is avoided because independent variables variations are magnified. In fact, if expansion is carried out around zero values, multicolinearity problems arise, although estimates are equivalent [20].

3. Production and cost functions in ports.

3.1 General aspects.

The literature on the econometric estimation of production and cost functions in ports is about 30 years old and is relatively scarce, as recently acknowledged by Turner et al. [23]. While it addresses many of the general issues raised in the theory, of particular interest is the set of

contributions focusing on economies of scale and, in some cases, economies of scope as defined by Baumol et al. [13]. Two different approaches can be distinguished within this group of papers. The first approach is represented by the studies using production functions, such as Chang [24], Reker et al. [25] and Tongzon [26]. The second approach encompasses the studies that estimate cost functions, either single-productive as in Kim and Sachish [27] and Martínez-Budría [28], or multi-productive as in Jara-Díaz et al. [29], Martínez-Budría et al. [30], Jara-Díaz et al. [31] and Jara-Díaz et al. [32]. In the following sections we present a synthesis of the empirical work, discussing output definition, independent variables, functional specification, and results on scale and scope. Tables *1* and *2* summarize the main features of these papers.

Although it is easy to define a seaport in broad terms as a transport interface, a detailed analysis of the economic activities taking place at seaport reveals a variety of different services that ships demand from a port, of which cargo loading and unloading is perhaps the more relevant but by no means the only one. Some of these services could be supplied under competitive conditions while other could present characteristics of natural monopoly. In the literature presented here, most of the papers deal with the provision of port infrastructure and, to a lesser extent, of cargo handling services, with the exception of Kim and Sachish [27], who try to study the port as a whole.

There are several reasons to justify the emphasis on infrastructure and cargo handling. First of all, both services often present conditions of natural monopoly (or oligopoly), which is why these services are of special interest to policy-makers and regulators. Secondly, in all the analyzed studies the service of infrastructure provision is provided by the Port Authorities, that are public organizations, which probably has facilitated data collection for this type of study. On the other hand, in the case of cargo handling the shortage of published articles

probably is due to the difficulty to collect the necessary data from private companies in spite of the interest in the study of this service.

Table 1. Production functions for the port sector

Author	Activity	Functional specification	Data	Variables (1)	Economies of Scale	Other Measurements
Chang (1978)	Infrastructure?	Cobb-Douglas	Time series Annual observations (21) (1953-1973)	$Y_1(X_1,X_2,e^{T/L})$	Constants	Average productivities Marginal Productivities
Reker el al. (1990)	Terminal-berth of Containers	Cobb-Douglas	Panel data Three terminals Monthly observations (70) (May 84-February-90)	$Y_2(X_6,X_4,X_5)$	Diminishing	Nothing
Tongzon (1993)	Terminal-berth of Containers	Cobb-Douglas	Panel data Three terminals Monthly observations (70) (May 84-February-90)	Y_3 (X_6,X_7,X_8)	Increasing	Efficiency for wharf

(1): Y_1 = Production measured as annual gross earning to the port in 1967 prices (wage payments to port workers not included)
Y_2 = Production measured as number of TEUS
Y_3 = Production measured as number of TEUS per berth hour
X_1 = Men years (excluding pork workers)
X_2 = Value of the clear assets of the port (prices of 1967)
$E^{T/L}$ =Proxy for the technological progress, (T/L = Tons for unit of labour)
X_3 = Net crane operation time
X_4 = Berth hours
X_5 = Labour
X_6 = Number of cranes per berth hour
X_7 = Labour per berth hours
X_8 = Number of TEUS carried by land per berth hour

Table 2. Cost functions for the port sector

Author	Activity	Functional specification	Data	Variables [1]	Scale economies evaluated in the approximation point	Other Measurements
			MONOPRODUCTIVE COST FUNCTIONS			
Kin y Sachish (1986)	Infrastructure and services	Translogarithmic	Time series Annual observations (19) (1966-1983)	$C(Y_1,W_1,W_5)$	Increasing	Minimal efficient scale Factor demand price elasticity Cross elasticities
Martínez-Budría (1996)	Infrastructure	Cobb-Douglas	Panel data 27 ports Annual observations (5) (1985-1989)	$C(Y_1,W_1,W_5,W_6,d_t,d_i,D_t,D_i)$	Increasing	Cost factor elasticities Individual specific effects of each port Second stage analysis
			MULTIPRODUCTIVE COST FUNCTIONS			
Jara-Díaz et al. (1997)	Infrastructure	Quadratic	Panel data 27 ports Annual observations (1985-1989)	$C(Y_2,Y_3,Y_5,Y_6,Y_7, W_1,W_5,W_6)$	Increasing	Marginal costs for each product Economies of scope (ED> 0)
Martínez-Budría et al. (1998)	Activity of the SEED (2)	Translogarithmic	Panel data 24 SEED Annual observations (1990-1996)	$C(Y_2,Y_5,W_1,W_6,T)$	Increasing	Marginal costs for each product Costs product elasticities Total factor productivity for a subsample of 14 SEED
Jara-Díaz et al. (2002)	Infrastructure	Quadratic	Panel data 26 ports Annual observations (1985-1995)	$C(Y_2,Y_3,Y_5,Y_6,Y_7,W_1,W_5,W_6)$	Increasing	Marginal costs for each product i Economies of scope (ED> 0)
Jara-Díaz et al. (2005)	Cargo handling	Quadratic	Panel data 3 Port Terminals Monthly observations (1990-1999)	$C(Y_2,Y_3,Y_4,W_2,W_3,W_4,W_5,W_6,W_7,D_3,T)$	Increasing	Marginal costs for each terminal Marginal costs for each product Economies of scope (ED> 0)

(1): C = Long term total annual cost
Y_1 = Tons of cargo
Y_2 = Tons of non-containerized general cargo
Y_3 = Tons of containerized general cargo
Y_4 = Tons of ro-ro cargo for port i in year t
Y_5 = Tons of liquid bulks
Y_6 = Tons of dry bulks
Y_7 = Index added of other activities that use part of the infrastructure
T = Temporal trend

W_1, = Labour price
W_2 = Non port worker personal price
W_3 = Ordinary port worker price
W_4= Special port worker price
W_5 = Capital price
W_6 = Intermediate input price
W_7 = Area input price
d_i= Individual specific effect to port i
d_t= Individual specific effect to year t
D_i= Firm-specific dummy variable
D_t= Port-specific dummy variable

3.2. Product definition

In almost all empirical applications, the authors have had to deal with the problem of defining the product. The reason is that, although most economic production activities are multi-productive, it was not until the seminal book by Baumol et al. [13] was published that a systematic theoretical body of concepts was made available for the analysis of multi-productive activities. They introduced definitions of new specific concepts related to multiproduct processes that could be contrasted empirically. In spite of their contribution, the initial empirical analyses of multi-productive activities continued to rely on aggregates to represent the product, or used attributes to capture the multi-productive nature that was being omitted. In some cases, the real output vector dimension made it impossible to estimate cost functions under a flexible functional form and, thus, forced some kind of aggregation, as in the case of transport firms.

Once these theoretical developments started to be taken into account, the definition of the product used for empirical estimations of production and cost functions in the port industry began revealing a great variety of approaches to address the importance of the issue. One of the factors driving the specific definition of the product used is what each author considers to be port activity. Before analyzing the product definition, it is thus important to point out that not all the authors consider the port from an integral point of view.

Reker et al. [25], Tongzon [26] and Jara-Díaz et al. [32] focus their papers in cargo handling service in port terminals, while Jara-Díaz et al. [29], Martínez-Budría [28] and Jara-Díaz et al. [31] exclusively analyze the provision of port infrastructure by port authorities. In another paper, Martínez-Budría el al. [30] studies the activity performed by *Sociedades Estatales de Estiba y Desestiba*, SEED (State-owned Loading and Unloading Companies -pool of port workers-) operating in Spain. In Kim and Sachish [27], port activity is considered from an

integral perspective encompassing not only the services provided by infrastructure but also the rest of port services. Lastly, in the case of Chang [24], there is no mention at all of what the author considers port services. Even though the exclusion of payments to port workers from gross benefit and the hours worked by them from labor factor seems to indicate that the author is modelling the services provided by infrastructure, this is not clearly and unambiguously established since, on the other hand, he also rejects the tons handled as a variable representing the product and chooses the gross benefit instead, on the grounds that official statistics generally include tons that passed through the port but were not necessarily handled by the port. With this context in mind, the diversity of product definition can be summarized as follows.

Reker et al. [25], and Tongzon [26], define product as the number of Twenty feet equivalent unity (TEU) and the number of TEU by berth hour, respectively. In the latter case, the author justifies this product measure on grounds of consistency with the Port Authority's objective of maximizing berth use.

Kim and Sachish [27] and Martínez-Budría [28] use annual cargo tons by port as product measure. In both studies, the authors acknowledge the multi-productive nature of the activity under analysis; however they estimate a single-productive cost function. In the case of Kim and Sachish [27], this is because they have a limited number of observations, although they remark that it would be advisable to disaggregate by type of cargo to avoid the aggregation bias. Furthermore, Martínez-Budría [28] chooses an aggregate of the activity because he assumes that the cost share of a ton of cargo is independent of the activity where it is handled. This is a restrictive assumption according to results obtained in further studies carried out by Jara-Díaz et al. [29] and Jara-Díaz et al. [31], where port activity of infrastructure service provision is represented by a five-component vector: the tons of general non-containerized cargo handled , of general containerized cargo, of liquid bulk, of

solid bulk, and CANON, which consists of an aggregated index of other activities using part of the infrastructure and which basically represents room granted under a concession or leased to private firms by port authorities.

Although Martínez-Budría et al. [30] claim that they model the activity of those SEED that reflect cargo handling services, they are really analyzing only the management of port workers that belong to these SEED, who are hired by the handling cargo companies. Their product vector has two components: tons of general cargo handled and tons of dry bulk.

Lastly, in the study by Jara-Díaz et al. [32] in cargo handling services, a three-product vector is considered for the analysis of cargo handling service, namely tons of container movements, roll-on/roll-off and non-containerized general cargo handled.

3.3. Independent variables

The independent variables used in the different papers vary according to the activity under study, or to the type of function to be estimated (production or cost function). Productive factors are the independent variables in a production function, while in the case of cost functions, product and factor prices are considered.

In those papers where a production function is estimated, independent variables basically represent labour and capital and, in some cases, an index of technical change. There are at least two ways to define the labour factor. On one hand, labour can be defined as the total number of workers. On the other hand, it can be defined as the total number of worked hours. When working hours differ among the different workers, it seems more suitable to use the second measure; otherwise, either can be used. In all three analyzed papers estimating production functions, i.e. Chang [24], Reker et al. [25, 33], and Tongzon [26], the labour factor is defined as the number of workers.

Regarding the variables used to represent the capital factor, a large variety is found. In Chang [24] capital is measured as the value of port net assets. In Reker et al. [25], they use berth hours, while in Tongzon [26] the selected variable is the number of cranes per berth hour. Furthermore, these papers include other variables used to reflect different aspects such as technical change [24] or the effect on production caused by factors other than capital and labor, such as land connections [26].

When estimating cost functions, it is necessary to have information on production (reviewed above) and on the price of each production factor involved in the process. In Kim and Sachish [27], Martínez-Budría [28], Jara-Díaz et al. [29] and Jara-Díaz et al. [31] consider the ratio between personnel expenses and the number of workers [34] to approximate labour price. On the other hand, Martínez-Budría et al. [30] applies the ratio between personnel cost and the number of worked hours. Lastly, in Jara-Díaz et al. [32], both approximations are used and a further distinction is made between port and non-port workers. The price for non-port workers is approximated by the ratio between total expenses for this type of labor and the number of workers, while in the case of port workers, the number of worked hours is used as denominator.

For capital price, there are slight differences across studies. In Kim and Sachish [27] capital input involves three types of assets: equipment (cranes), other depreciated equipment (service equipment and facilities) and other non-depreciated assets and materials. Capital input price is calculated through the Christensen-Jorgenson [35] User Cost. In Martínez-Budría [28] and Jara-Díaz et al. [29, 31], capital price is approximated by the ratio between the amortization of the period and the number of linear meters of berths with a depth over four meters, adding a rate of return on net fixed assets to the amortization of the period. In Jara-Díaz et al. [32], the approximation used is the ratio of the cost of capital to the active capital of the period. The cost of capital results from the addition of the accounting amortization for the period plus the

return on the active capital of the period and the shares of stock of the SEED. Lastly, in Martínez-Budría et al. [30], capital input is not incorporated since, according to the authors, it is considered to be a residual category in this type of activity.

Finally, Martínez-Budría [28], Jara-Díaz et al. [29, 31], Martínez-Budría et al. [30] and Jara-Díaz et al. [32] use in their papers –in addition to labour and capital- an additional factor denominated intermediate input which consists of a variable capturing other activity-related cost allocations and whose price arises from the ratio between all cost allocations other than personnel costs and depreciations, and the total activity. This total activity is represented by the total number of tons handled by the port in Martínez-Budría [28] and Jara-Díaz et al. [29], and it is represented by the annual revenues in Jara-Díaz et al. [31]. Lastly, in Jara-Díaz et al. [32] electricity price is used as the indicator of intermediate input price.

3.4. Objectives and functional specification

The study carried out by Chang in 1978 represents the first reference in the literature where a port production function was estimated. This author estimates a production function to analyze the productivity and convenience of expanding the capacity of the port of Mobile (Alabama-USA). Formally, he estimates a Cobb-Douglas production function.

Also to analyze productivity and offering an alternative indicator of partial factor productivity measures, Reker et al. [25] estimate a production function for three container terminals situated in the Port of Melbourne. This is the first study estimating a production function for port terminals and modeling cargo handling service. Reker et al. [25], following De Neufville and Tsunokawa [36], consider that it is better to estimate a production function than resorting to the usual approach of estimating a cost function. The main reason for this lies with the difficulty to obtain reliable data on productive factor prices. In order to take advantage of the large number of individual performance measures which had been

previously calculated, Reker et al. [25] consider these measures as productive factors (independent variables) of the production function. Although they acknowledge that the selected productive factors are not completely independent, they assume that the degree of dependency can be omitted. They also estimated a Cobb-Douglas function.

Tongzon [26], as Reker et al. [25], estimates the production function of container handling, although in this case the author's objective is to examine whether the new tariff policy at the Port of Melbourne improved port efficiency. At the same time, he assess the contribution of the different factors involved in port efficiency. This study also considers a Cobb-Douglas function.

As seen, the Cobb-Douglas function has been widely used in the literature on production functions to evaluate scale effects, since the degree of economies of scale can be easily calculated but it presents important limitations. For this reason, as seen in section 2, it is clearly preferable to use functional forms which avoid restrictions imposed by the functional form itself -such as the so-called flexible functional forms. Flexible functions are of common use to estimate cost functions.

A flexible functional form commonly used in the empirical analyses through cost functions is the translogarithmic function. Indeed, this is the approach followed by Kim and Sachish [27] in their study of the port of Ashdod. This paper represents the first reference in the literature of port cost function estimation. In this paper, the authors focus on three objectives: first, they analyzed the port production structure placing special emphasis on the productive factors substitution pattern and the determination of the existence or absence of economies of scale. Secondly, they attempt to determine the nature and the impact of technological change, i.e. not only the technological change ratio is analyzed but also whether possible biases in this change could alter the productive factors shares. Lastly, they explore the interrelation

between the port internal economies of scale and the external technological change, in the determination of total factor productivity. For this purpose, total factor productivity is decomposed into two parts: one related to economies of scale and the other induced by technical change.

In order to estimate the technical change and the technology of port operations –taken as the services provided by infrastructure and cargo handling-, Kim and Sachish [27] estimate, a system of equations consisting of the translog total cost function, the factor cost share equation (Shephard's lemma) and the appropriate parametric restrictions such that the cost function meets the conditions of symmetry and homogeneity in factor prices.

Some years later, Martínez-Budría [28] analyzed the provision of port infrastructure services managed by the port committees that evolved into the present port authorities and, at the same time, analyzed the differences between them. He assumed that the technology used by all port authorities is the same and that it can be analyzed through a model with an error structure including a fixed time effect and a specific individual effect. The fixed time effect is common to all firms although it varies along periods and it reflects the technical change during the observation period. On the other hand, the specific individual effect enables the analysis of the reasons for cost differences. He estimated the cost function with a panel of data of observations from 27 Spanish Ports of general interest. The functional form specification selected by the author is a Cobb-Douglas function, which also has important limitations when used as a cost function, as explained earlier.

Using the same original data plus six more years of observations, a multi-product version of this model was developed by Jara-Díaz et al. [31], which has a previous version [29] using the same data period as in Martínez-Budría [28]. The aim of Jara-Díaz et al. [29, 31] was to determine the specific marginal costs of each product, the multiproduct degree of economies

of scale and the degree of economies of scope for the services of the Spanish port infrastructure. The model estimated differs from the one used by Martínez-Budría [28] basically in three aspects. Firstly, Jara-Díaz et al. [29, 31] change the functional form specification and decide to apply a flexible functional form (quadratic). Secondly, although Martínez-Budría [28] acknowledged the multi-productive nature of the activity in his paper, he used a product-aggregated measure, while in Jara-Díaz et al. [29, 31], product is defined as a five-component vector. Lastly, panel data estimation techniques used in both papers differ. Thus, in Martínez-Budría [28], function estimation is performed through least squares with dummy variables (fixed effect model), while in Jara-Díaz et al. [29,31], a system of equations made up by a long run quadratic cost function and a set of factor-derived demand functions is estimated, using Zellner's iterative technique.

Using a similar structure as in Jara-Díaz et al. [29,31], Martínez-Budría et al. [30] analyze the results of the reform of loading/unloading operations in Spanish ports. The functional specification applied to carry out the estimation is the generalized translog function which is estimated together with factor-derived demand equations. Finally, Jara-Díaz et al. [32] represents the first paper estimating a long-run total cost function for cargo handling service in multi-purpose port terminals moving mostly containers. They used data on cost and production collected directly from the files of three concessionaire firms in the port of La Luz and Las Palmas located in Gran Canaria, Spain. The paper reports marginal costs and economies of scale and scope. Using the iterative technique modified by Zellner, the authors estimate a system of equations made up by the total cost function and the factor share equations. A detailed account of data collection and models can be found in Tovar,B., [37].

3.5. Estimation results: Economies of scale and/or scope

As explained above, the published studies are heterogeneous in relation to the activity analyzed, the functional specification applied and the objectives pursued. This explains why the measures carried out by the different authors are also heterogeneous in general, but have a common denominator, which is the estimation of the degree of economies of scale. This section contains an analysis of results regarding scale economies and economies of scope. So, emphasis is made on the calculation of quantities that are meaningful when dealing with regulation. These are pieces of information that, although necessary and important, are not readily available to the regulators but have to be calculated from functions estimated using data specifically collected.

When comparing the empirical estimations of economies of scale of the different papers referred to in this survey, it is important to bear in mind that the activity under analysis differs across the studies. All the papers focusing on the study of the services provided by infrastructure, i.e. Chang [24], Martínez-Budría [28] and Jara-Díaz et al. [29, 31], conclude that there exist increasing returns to scale, although it is necessary to make further comments on this.

In this study for the port of Mobile, Chang [24] finds increasing returns to scale, although the hypothesis of constant returns cannot be rejected when the calculated confidence intervals are taken into account. On the other hand, the degree of economies of scale estimated by Martínez-Budría [28] using a single-productive approach is 3.47, much larger than the results arising from the application of a multi-product approach using similar data, which yielded 1.43 [29] and 1.69 [31] at the sample mean (the degree of economies of scale by port was also calculated in both of these latter studies). According to Jara-Díaz et al., the reason for this difference is the presence of economies of scope, which cannot be revealed using the

aggregate description of the product but have been clearly detected in the last two studies. As shown by Jara-Díaz [38], the use of aggregates to represent products implies not only loss of information but also may cause misinterpretation of the coefficients estimated in the empirical analyses, a problem that the policy makers and regulators need to be aware of. For example, in a single output approach the existence of economies of scale is enough to conclude that the activity is a natural monopoly, but in a multiproduct context even the simultaneous existence of economies of scale and scope is not enough to establish it. Therefore, if the analyst chooses a single-product context when the correct one is a multioutput one, then the hidden presence of economies of scope could create a 'false' natural monopoly.

Indeed, in Jara-Díaz et al. [29, 31] the authors calculate the degree of economies of scope at the sample mean for three subsets of the output vector and their complements. First, they studied whether separating liquid bulks from the rest resulted in any benefit in costs. Secondly, they analyzed the convenience of separating general cargo from liquid and solid bulks. Finally, they considered general cargo specialization. In both papers, the results obtained indicated the presence of economies of scope showing that, if all outputs have to be produced, then it is not advisable to specialize port infrastructure by type of product because of some 40% $\pm$ 5% savings in joint production.

The literature that analyzes container terminals-berths using a production function [25, 26], arrives at contradictory conclusions. On one hand, although Reker et al. [25] do not mention this, their parameter estimates suggest the existence of decreasing returns but they neither report standard errors of parameters nor t-stats, which prevents the determination of confidence intervals in order to verify the hypothesis of decreasing returns. On the other hand, the parameters estimated by Tongzon [26] suggest that his production function is subject to increasing returns, although this contradicts an earlier conclusion drawn by the

author in the sense that constant returns are present. Again, confidence intervals are not reported and neither hypothesis can be tested.

The study carried out by Jara-Díaz et al. [32] analyzes cargo handling activity in port multi-purpose terminals through cost function estimations. The multi-productive approach developed in this paper allows estimations of marginal costs by product, scale economies (global and specific) and economies of scope, which are quite useful in order to determine the type of terminals (multipurpose or specialized) that should be allowed at the specific port analyzed for the given levels of demand (traffic mix and volume). The degree of economies of scale evaluated at the mean yielded a 1.64 estimate, thus indicating the existence of increasing returns of scale. Economies of scope were analyzed in order to evaluate if there was any benefits in costs resulting from jointly handling containers, general cargo and Ro-Ro cargo. All relevant partitions of output vector were analyzed. In every case, the conclusion drawn is that joint production at the given levels [39] was convenient (as opposed to specialization). Note that this should be interpreted in a very strict sense, i.e. ***if*** the three products are to be produced at the given levels because of exogenous reasons, it would be better done (less costly) with one firm than with two or three firms. Furthermore, the authors analyze firm specific scale economies, concluding that the two neighbour and smaller firms could operate jointly, increasing scale advantages while keeping competition with the single larger operator.

Kim and Sachish [27] analyze port activity from an integral point of view. Estimation of scale economies was 1.3, thus, leading to the conclusion that port services production in the port of Ashdod is subject to increasing returns to scale at the approximation point. The authors explain that Ashdod is an artificial port located between two important breakwaters. This situation not only hinders its physical expansion but also supports the existence of increasing returns to scale.

Finally, Martínez-Budría et al. [30] only analyze management of port workers depending on the SEED. Based on their analysis, the authors conclude that the degree of economies of scale evaluated at the mean results in 1.126, thus evidencing, the existence of increasing returns to scale.

4. Conclusions

This paper focuses on the contributions in the literature to the study of traditional production and cost functions in the port industry and provides a detailed and critical analysis of the relevant aspects of cost and production functions within the sector. The estimation of key concepts in the firms' cost structure, such as marginal costs, economies of scale and scope, plays an essential role in the determination of the optimal industrial structure and therefore, represents a fundamental tool for efficient regulation.

The first important conclusion from this review is the finding of a quite limited literature on production and cost structure of port activities, particularly in connection with multi-productive cost functions of port terminals. Probably, this reflects deficiencies in the data rather than a lack of interest in the subject.

Within the port area, a great diversity of activities are performed: infrastructure services, generally provided by port authorities, cargo handling services, in most ports provided by private firms, and other services such as mooring, towage, etc. Each of these activities shows well-differentiated features and their own technology, such that they can be considered as different services (although related). From this perspective, it is difficult to speak about port activities in a broad sense. At this stage, it is clearly evident that some papers lack a precise definition of the service under analysis.

In spite of the limited number of studies available, the literature analyzed shows that not only infrastructure but also cargo handling services present increasing returns to scale. The few

multi-productive studies also suggest the presence of economies of scope between cargo types for both services. Note that this prevents marginal cost pricing if the infrastructure investment is to be recovered through pricing, and a similar thing occurs with cargo handling, where second best prices could be an appropriate alternative. Demand elasticities should be calculated and considered if this scheme is desired.

Ultimately what this survey reveals is that applied research still has a long way to go in the sector. The proper approach is multi-productive, and the consequences of not taken into account this important fact are quite relevant for the adequate regulation in port activities. Single output studies could suggest erroneous efficient structures. Most probably, the lack of studies on the sector's cost behavior is due to the great difficulty of collecting information. This is unfortunate since the sector continues to have some important components of its business which have monopolistic features. Rents are thus being created and unless costs levels and structures are properly assessed, the levels and distribution of these rents are unlikely to be known. If governments are serious about their commitment to improve the competitiveness of their countries, they will have to ensure that port costs and hence rents are minimized and this can only be done if costs are measured and assessed properly. One of the regulator's tasks should be to help obtaining relevant information from adequate sources in order to undertake properly these types of studies. One way to advance in this direction is to make it mandatory for regulated firms to release such information in order to take full advantage of the methodological analysis that could be applied to port activities.

Acknowledgements

This research was partially funded with grants from Fondecyt, Chile, project 1050643, the Millennium Nucleus 'Complex Engineering Systems' and Gobierno Autónomo de Canarias,

project PI2000/047. Prof. Jara-Díaz would like to thank the hospitality at the Universidad de Las Palmas de Gran Canaria while working on this research.

References and notes

1. See appendix for a synthesis of the theoretical concepts related with production and cost function.
2. The bill of a vessel or scale account is the total amount that the ships have to pay to use all the different services that has been used during the time the ship has spent in port.
3. Farrel M. (1957): *The Measurement of Productive Efficiency*, Journal of the Royal Statistical Society, Series A, Part III, Vol. 120.
4. Charnes, A., Cooper, W. and Rhodes, E., 1978, *Measuring the Efficiency of Decision Making Units*. European Journal of Operational Research, 2 (6), 429-444.
5. Charnes, A., Cooper, W. Lewin, W. and Seiford, L., 2000, *Data Envelopment Analysis: Theory, methodology and applications*, Kluwer Academic Publishers.
6. Roll, Y. and Hayuth, Y., 1993: Port performance comparison applying data envelopment analysis (DEA). *Maritime Policy and Management.* 20, 2, 153-161.
7. Tongzon, J.L., 2001, Efficiency measurement of selected Australian and other international ports using Data Envelopment Analysis, *Transportation Research, Part A* 35, 113-128
8. Wang, T, Song, D-W and Cullinane, K.P.B., 2003, Container port production efficiency: a comparative study of DEA and FDH approaches. *Journal of the Eastern Asian Society for Transportation Studies*, 5, 698-713.

9. Estache, A., Tovar, B. and Trujillo, L., 2004, Sources of efficiency gains in port reform: a DEA decomposition of a Malmquist TFP index for Mexico. *Utility Policy* 12, 221-230

10. Cullinane, K., Song, D-W., Wang, T-F., 2005, The application of mathematical programming approaches to estimating container port production efficiency. *Journal of Productivity Analysis*, 24, 73-92.

11. Aigner, D., Lovell, C. and Schmidt, P., 1977, Formulation and estimation of stochastic frontier production function models, Journal of Econometrics, 6.

12. Meeusen, W. And Van den Broeck, J., 1977, Efficiency estimation from Cobb-Douglas production functions with composed error, International Economic Review, 18 (2).

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

14. Estache, A., González, M.M. and Trujillo, L., 2002, Efficiency Gains from Port Reform and the Potential for Yardstick Competition: Lessons from Mexico. *World Development*. 30(4), 545-560

15. Cullinane, K., Song, D-W., 2003, A stochastic frontier model of the productive efficiency of Korean container terminals. *Applied Economics*, 35, 251-267

16. Rodríguez-Álvarez, A. Tovar, B. and Trujillo, L., 2004, Technical and allocative efficiency for cargo handling firms in Spanish ports. *North American Productivity Workshop 2004.* Toronto. Canada**.**

17. Cullinane, K., Wang, T-F., Song, D-W., and Ji, P., 2006, The technical efficiency of

container ports: comparing data envelopment analysis and stochastic frontier analysis. *Transportation Research Part A.*, 40,354-374.

18. Cullinane, K., 2002, The Productivity and Efficiency of Ports and Terminals: Methods and Applications in Grammenos, Costas TH. (ed), The Handbook of Maritime Economics and Business, Informa Proffesional, London.

19. Wang, T. F., Cullinane, K. and Song D.W., 2005, Container Port Production and Economic Efficiency. Palgrave MacMillan. New York

20. González, M.M. and Trujillo, L., 2006, Efficiency Measurement in Port Industry: A Survey of the Empirical Evidence. Departamento de Análisis Económico Aplicado. ULPGC (mimeo)

21. McFadden, D., 1978, Production economics a dual approach to theory dual applications. Amsterdam North-Holland Publishing Company

22. Baumol, W., Panzar, J. and Willig, R.,1982, Contestable markets and the theory of industry structure. Harcourt, Bruce and Jovandich, Inc., New York.

23. In this case, applying Box-Cox transformations can solve the problem, although this highly complicates the interpretation of parameters.

24. Diewert, W. E., 1974, Applications of duality theory. In M.D. and D.A. Kendnick (ed.) *Frontiers of Quantitative Economics II*, 106-171. Amsterdam North-Holland

25. Dodgson, J.S., 1985, A survey of recent developments in the measurement of rail total factor productivity. In Button, K.J. and Pitfield, D. (Eds.) *International Railway Economics*, 13-48) Gower, Aldershot.

26. Caves, W., Christensen, L. R., and Tretheway, M. W.,1980 Flexible cost functions for multiproduct firms. *The Review of Economics and Statistics* 62, 477-481

27. Lau, L.J., 1974, Applications of duality theory. Comments. In M.D. and D.A. Kendnick, II (ed) *Frontiers of Quantitative Economics II* ,176-199. Amsterdam North-Holland

28. Röller, L.H., 1990, Proper quadratic cost functions with an application to the bell system. *The Review of Economics and Statistics* 72, 202-210

29. Jara-Díaz, S., 1983, Freight transportation multioutput analysis. *Transportation Research* 17A, 6, 429-438

30. Mayo, J.W., 1984, Multiproduct monopoly, regulation, and firm costs. *Southern Economic Journal* 51, 208-218

31. From n equations of factors derived demand only n – 1 are linearly independent.

32. Turner, H., Windle, R. And Dresner, M., 2004, North American containerport productivity 1984-1997. *Transportation Research Part E* 40, 339-356

33. Chang, S.,1978, Production function and capacity utilization of the port of Mobile. *Maritime Policy and Managemen.* 5, 297-305

34. Reker, R.A., Connell, D. and Ross, D.I., 1990, The development of a production function for a container terminal in the port of Melbourne. *Papers of the Australiasian Transport Research Forum* 15, 209-218

35. Tongzon, J.L., 1993, The Port of Melbourne Authority's pricing policy its efficiency and distribution implications. *Maritime Policy and Management* 20, 197-203

36. Kim, M. and Sachish, A., 1986, The structure of production, technical change and productivity in a Port. *International Journal of Industrial Economics* 35, 209-223

37. Martínez-Budría, E., 1996, Un estudio econométrico de los costes del sistema portuario español. *Revista Asturiana de Economía* 6, 135-149

38. Jara-Díaz, S., Martínez-Budría, E., Cortes, C. and Vargas, A., 1997, Marginal costs and scale economies in spanish ports. *25th European Transport Forum*, Proceedings Seminar L, PTRC, London, 137-147

39. Martínez-Budría, E., González-Marrero, R. and Díaz, J., 1998, Análisis económico de las Sociedades Estatales de Estiba y Desestiba en España. Working paper 97/98-1. Universidad de La Laguna. Tenerife.

40. Jara-Díaz, S., Martínez-Budría, E., Cortes, C. and Basso, L., 2002, A multioutput cost function for the services of spanish ports' infrastructure. *Transportation* 29, 419-437

41. Jara-Díaz, S. Tovar, B., and Trujillo, L., 2005, Marginal costs, scale and scope for cargo handling firms in Spain. *Transportation*. 32, 275-291

42. This work shows several weaknesses. First, there is no explicit definition of independent variables, which are obscurely presented. For instance, they include crane operation time variable in the estimation, and later on they acknowledge that the elimination of that variable enables a better adjustment of data. Finally, independence among explanatory variables is assumed, although the authors themselves believe that they are not completely independent.

43. Kim and Sachish, 1986, use 'actual' number of workers, which is determined through the application of an index. In this way, they aggregate the hours worked by workers performing different tasks weighted by the importance that each task bears in total labour costs.

44. Christensen, R. and Jorgenson, D. W, 1969, The measurement of U.S. real capital input, 1929-1967. *The Review of Income and Wealth* 5, 293-320

45. De Neufville, R. and Tsunokawa, K.,1981 Productivity and returns to scale of container ports. *Maritime Policy and Management* 8, 2, 121-129

46. Tovar, B., 2002, Análisis multiproductivo de los costes de manipulación de mercancías en terminales portuarias. El Puerto de La Luz y de Las Palmas. (A multioutput cost analysis for cargo handling services in port terminals. La Luz y de Las Palmas' Port) Ph D. Departamento de Análisis Económico Aplicado. Universidad de Las Palmas de Gran Canaria. España.

47. Jara-Díaz, S., 1982, The estimation of transport cost functions a methodological review. *Transport Reviews.* 2, 257-278.

48. This is a logical result if the size of these multipourpose terminals is taken in account.

49. Panzar, J. and Willig, R., 1977, Economies of scale in multioutput production. *Quaterly Journal of Economics***,** 91, 481-494.

Appendix: Production and Cost Concepts

Production of goods and services can be defined as a process that permits the transformation of some production factors into some products. The production function is an attempt at representing analytically the technical range of possibilities open to producers. Let $X = (x_{1, \ldots,} x_n)$ be a vector of non-negative production factors and $Y = (y_1 \ldots, y_m)$ a vector of non-negative products. Let T be the set of all combinations (X,Y) that are technically feasible. Under some regularity conditions efficient production can be described through the general transformation function

$$F(X,Y) = 0. \qquad (1)$$

If production is described by a single homogeneous product y, then the usual concept of production function is recovered

$$y = f(X). \tag{2}$$

From equations (1) or (2) technical properties like input substitution or scale effects can be analyzed by means of the function and its first and second derivatives. When the basic economic unit, the firm, faces competitive markets (input and output prices are exogenous) the observed relations between inputs and outputs are purely technical and the properties can be empirically studied. Under certain regularity conditions this technical properties can be also studied from the cost function using its duality properties regarding production.

In general, if F in equation (1) is strictly convex with respect to X, there exist a unique cost function that is dual to F that, for given input prices and level of production, can be written as

$$C = C(W, Y) \tag{3}$$

The cost function $C(W,Y)$ represents the minimum expenditure necessary to generate the products contained in vector Y, at factor prices W. The latter has been eliminated in the expressions below in order to simplify the mathematical formulae, which follows Baumol, Panzar and Willig [13]. First, the local variation in costs after the increase of one product keeping all other products constant is the marginal cost of product i, calculated as

$$\frac{\partial C}{\partial y_i} = Cm_i \tag{4}$$

On the other hand, the degree of global economies of scale S is a technical property of the productive process which is defined in the transformation or production functions. However, dual relations allow the calculation of S directly from the cost function [40] as

$$S = \frac{C(Y)}{Y \ \nabla_y C(Y)} \tag{5}$$

The degree of global economies of scale represents the maximum growth rate that the product vector can reach when productive factors increase by the same proportion. Therefore, the

presence of increasing returns of scale (S>1) implies that a proportional growth of all products induces a less than proportional growth of costs, i.e. a production expansion exhibits advantages from the point of view of costs. If prices are set equal to marginal costs under increasing returns, the firm will have losses.

Two products are said to exhibit cost complementarity when the marginal cost of one of them diminishes as the other product increases. Formally, this means that

$$C_{ij}(Y) = \frac{\partial^2 C(Y)}{\partial \ y_i y_j} \leq 0, \qquad (6)$$

and represents some form of advantage in joint production, with the inequality holding strictly over a set of non-zero measure .

The concept of economies of scope is useful to analyse whether it is advisable or not to have the firm diversified or specialised. Thus, economies of scope measure the relative cost increase that would result from the division of the production of *Y* into two different production lines *T* and *N-T*. Formally, if an orthogonal partition of product vector *N* into two subsets *T* and *N-T* is carried out, the degree of economies of scope SC_T of subset of products *T* with relation to its complementary subset *N-T* is defined as

$$SC_T(Y) = \frac{1}{C(Y)}\left[C(Y_T) + C(Y_{N-T}) - C(Y)\right], \qquad (9)$$

in such a way that the partition of the production set will increase, decrease or not alter total costs depending on whether $SC_T\ (Y)$ is larger than, smaller than or equal to zero, respectively. Thus, if $SC_T(Y)$ >0 economies of scope are said to exist and it is cheaper to produce vector *Y* jointly than to produce vectors Y_T and Y_{N-T} separately. In other words, it is not advisable to specialise but to diversify production. It is easy to see that *SC* should be in the interval (-1, 1).