AN AFFINE LONG-RUN PRODUCTION COST FUNCTION

Rickard E. Wall
Department of Management and Economics, Linköping University, SE 581 83 Linköping, Sweden.

ABSTRACT Economics-of-scale is an ubiquitous phenomena in real-world production processes. A common way to model economies-of-scale, e.g. in *Location Theory* models, is to assume that the production of a commodity entails fixed costs, F, and a constant marginal cost, c, so that the production costs of a plant are given by $F + cq$, where q = quantity produced. It is not obvious whether this function is a short- or a long-run cost function. Elementary economic theory postulates that a short-run cost function should have an explicit capacity limit, and that there should be no fixed costs in the long run. In order to derive long-run equilibrium results the model builder, however, may feel more comfortable using an unambiguously long-run cost function. Such a function, which still possesses the desired properties of the cost function given above, is derived in the present paper.

1. INTRODUCTION

A fundamental point-of-departure in many spatial economic models is the trade-off between using economies-of-scale in the production and increases in transport costs. This trade-off arises because the average transport distance of the goods increases as increased output typically cannot be disposed of without increasing the market area size of the plant. Indeed, “ .. the tradeoff between increased transport costs and decreased production costs is the heart of spatial economics”, Arnott (1987, p. 429) in *Palgrave,* and the roots of this trade-off can be traced back at least to Adam Smith’s *Wealth of Nations* (1976, p. 1776), in which the third chapter of the first Book is entitled “The Extent of the Market is Limited by the Division of Labour”. This idea has in modern times been further developed and couched in both analytical and diagrammatical terms, see e.g. Lewis (1945, p. 204), Beckmann (1968, p. 45) and Mohring and Williamson (1969, p. 225).

The presence of economies-of-scale characterizes many production processes in the real world, and it is therefore often a fruitful model assumption. As recognized by Eaton and Lipsey (1976, p. 77), an affine function - that is, a function $f(x) = ax$ (which, thus, passes through the origin) plus an intercept, see e.g. Hands (1991, p. 112) - is a convenient, and commonly-used, functional form for production costs of a plant in spatial economic model building:

$$TC = TC(q) = F + c \cdot q \tag{1}$$

where

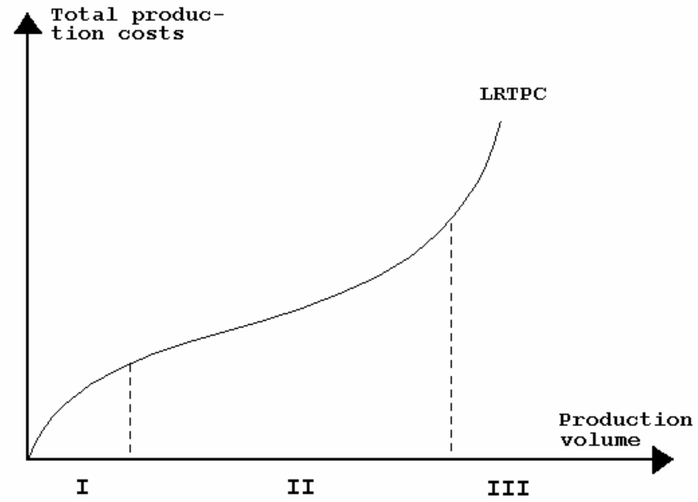

Figure 1. The Cubic Long-Run Total Production Cost Function.

TC = Total production costs
F = Fixed costs
c = marginal cost
q = quantity produced

It is not obvious whether (1) is a short- or a long-run cost function. Elementary economic theory postulates that a short-run cost function should have an explicit capacity limit, and that there are no fixed costs in the long run. In order to derive long-run equilibrium results, however, the model builder may feel more comfortable using an unambiguously long-run cost function. The purpose of this paper is to develop such a function, one that still possesses the required properties of cost function (1).

Section 2 reviews a few characteristics of the textbook cubic cost function, which are illustrated by some scattered empirical evidence shown in Section 3. What functional form is appropriate for the empirical data as well as for cost functions in general is the subject of Section 4. The main section of the paper is Section 5, in which an affine long-run production cost function is algebraically derived. Concluding remarks are given in Section 6.

2. THE TEXTBOOK CUBIC LONG-RUN PRODUCTION COST FUNCTION

The textbook cubic Long-Run Total Production Cost (LRTPC) function, shown in Figure 1, is very simple. It is non-linear, and postulates that there are no fixed costs in the long run.

For illustration purposes the production volume range in Figure 1 has been divided into three segments. In *Segment* I the LRTPC rises degressively, in *Segment* II the LRTPC rises by and large proportionally, and in *Segment* III the

LRTPC rises progressively with output. Thus, in *Segment III* diseconomies-of-scale is present. The graph in Figure 1 is a theoretical construction. Section 3 offers some empirical results.

3. SOME SCATTERED EMPIRICAL OBSERVATIONS

Production and cost data on plant level are often considered confidential information, and therefore virtually no private firm is willing to reveal it. A few cost functions have, nevertheless, been estimated in the present project. The selection of firms was not made at random, but was influenced by which firms were willing to offer the necessary data. Thus, the empirical evidence given in this section serves only to illustrate some actual, although not necessarily representative, cost and production volume relationships.

The research approach used in this project can be illustrated by the way in which the estimation of a production cost function in the Swedish single-plant bakery industry was obtained.

A dozen single-bakery firms in Sweden were identified using sources such as Web sites on the Internet, and the Federation of Swedish Industries. A letter of introduction was sent to each of these bakeries; this letter was followed shortly by a personal telephone call. The only question posed to bakery staff concerned the bakery's yearly production volume in tons. Most of the firms were willing to give out this single piece of information. The annual financial reports of the bakeries in question - from which production costs can be inferred - were ordered from the Swedish Patent and Registration Office (PRV). (Every Swedish company must submit its annual financial reports as public documents to PRV). In this way total production and total cost data for nine bakeries were obtained.

Production- and cost data were obtained in a quite similar manner for six hair salons, seven driving schools, eight oil depots, and 48 Primary Health Centers (PHC) – the latter were grouped into three categories – inner-city, suburban, and rural PHCs. All of these data sets, added with straight line fits, are plotted in seven figures; Figure 2 to Figure 8 (note, 1 AUD ≈ 5 SEK).

The observations in the figures, respectively, are quite few in number. Still, mere visual inspection of the figures indicates that affine functions approximate the cost and output observations of this particular sample reasonably well, which may serve as an argument for using cost functions like the one in (1). There is more to consider, however.

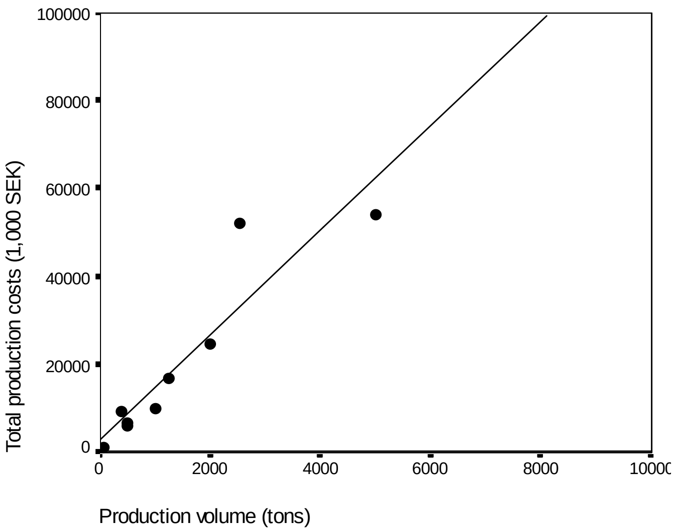

Figure 2. Production and Costs for Nine Swedish Bakeries in 1999.

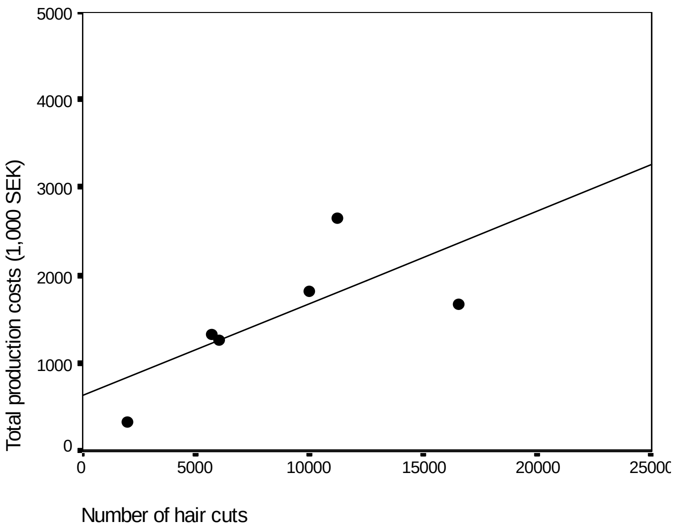

Figure 3. Production and Costs for Six Swedish Hair Salons in 1999.

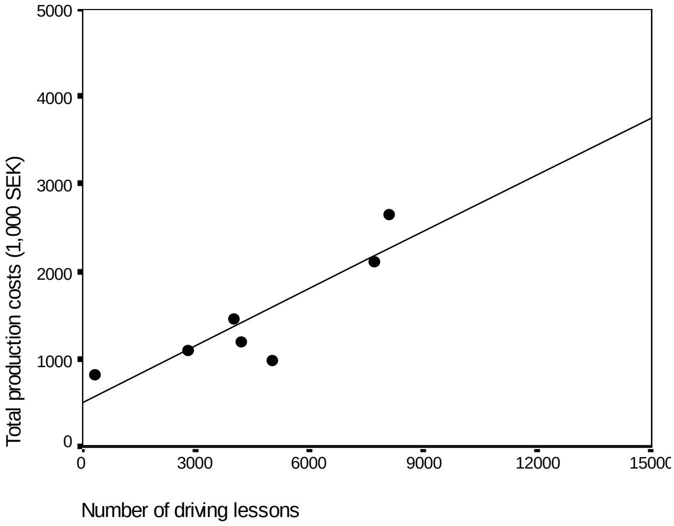

Figure 4. Production and Costs for Seven Swedish Driving Schools in 1999.

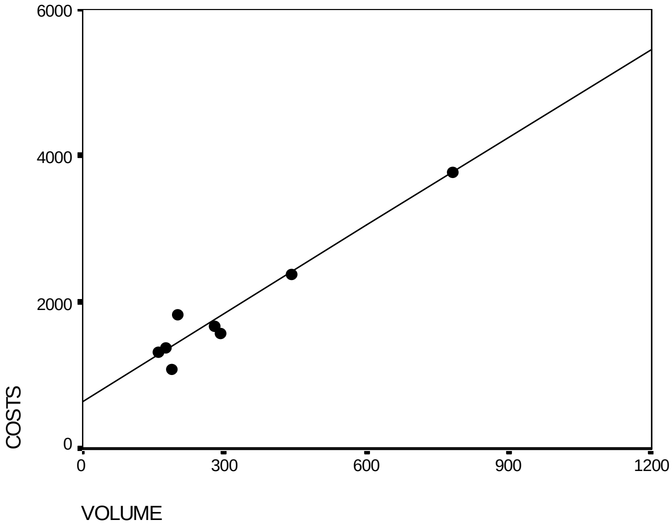

Figure 5. Production and Costs for Eight Swedish Petrol Depots in 1996.

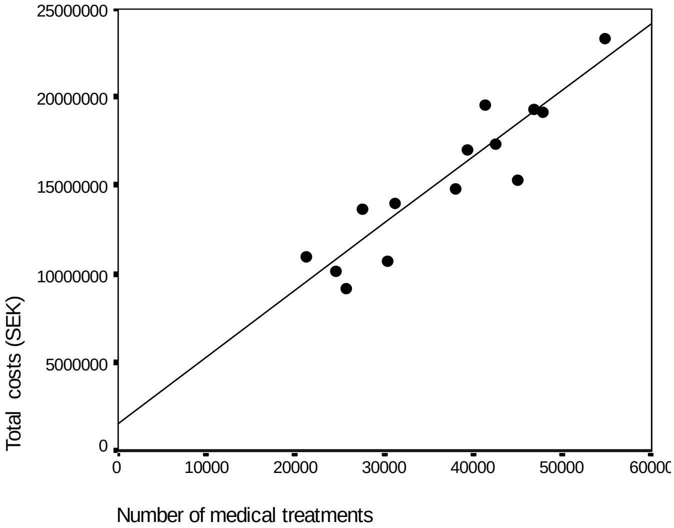

Figure 6. Production and Costs for 14 Swedish Inner-city Primary Health Centres in 1995

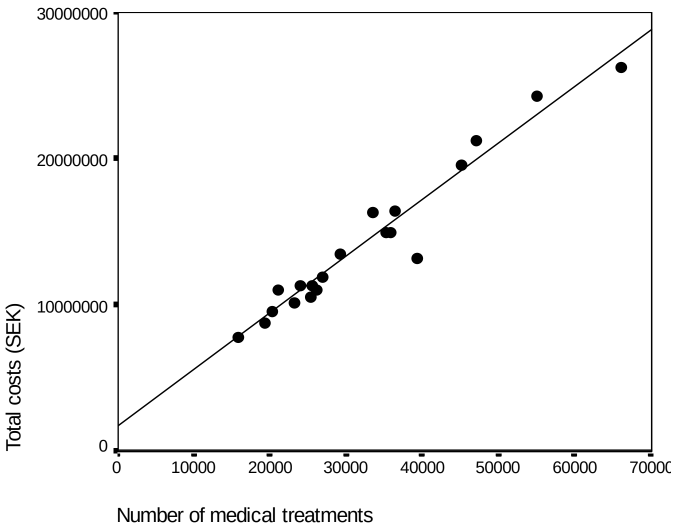

Figure 7. Production and Costs for 20 Swedish Suburban Primary Health Centres in 1995

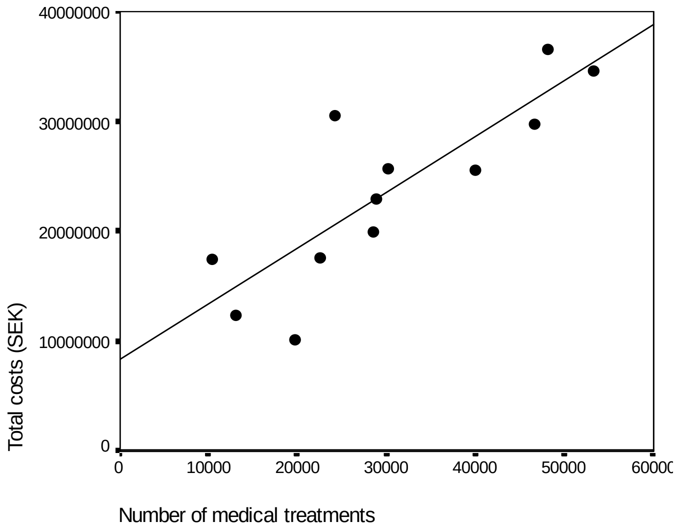

Figure 8. Production and Costs for 12 Swedish Rural Primary Health Centres in 1995.

4. WHICH FUNCTIONAL FORM IS APPROPRIATE?

Two observations can be made in Figures 2-8. First, there is an upper limit for the output volume of a production facility. Indeed, with the exception of the large petrol depot in the upper right corner of Figure 5, output volumes of the largest production facility generally is no more than three or four times larger than the smallest. Second, very small output volumes are rare.

Firms seldom operate plants in *Segment* III, as it generally pays in such cases to split the production into two separate plants. Therefore, the researcher will seldom find observations in the very large output volume *Segment* III. Furthermore, as pointed out by e.g. Miller (1977, pp. 79-80), once a plant is built it is seldom operated at very small production volumes *relative* to the set-up costs involved. This is because firms running such plants most often would be out-competed by firms operating plants at higher production volumes and, thus, at lower average costs. Therefore, very small output volume observations rarely are found in empirical cross section data. What the researcher often does find are *Segment* II observations.

It is furthermore interesting to note that the data in Figures 2-8 can be said to lie, by and large, on straight lines. In the figures, graphs are fitted to the data and extrapolated downward to zero output volumes and upward to varied output volumes. It is perfectly possible, however, to imagine the presence of observations in *Segments* I and III of Figures 5-8 in such a pattern that these will approximately fit into the basic cubic LRTPC function. With the exception of left-most observation, the same holds true for the data sets in Figure 3 and 4, respectively. The data structure is less obvious in Figure 2.

Thus, it is not obvious from the *Segment* II observations of Figures 2-8 what kind of cost function – an affine, a cubic, or some other – is relevant over the output range 0 to ∞. Furthermore, it is typically not possible to fit cubic long-run cost functions with the observations at hand. In the next section, however, an affine cost function is derived, that may be regarded as long- run, as the fixed costs can be disposed of.

5. THE AFFINE LONG-RUN TOTAL PRODUCTION COST FUNCTION DERIVED

In this section the task is to develop a long-run cost function that can be squared into *Segment* II of the cubic function in Figure 1 without explicitly violating the shapes of *Segments* I and III. The function should also be reconcilable with the functional form of (1), in order to conform with established spatial economic model building.

It seems unsatisfactory to assume a constant-elastic cost function. Sooner or later constant returns-to-scale is likely to set in – it would have some odd consequences in a spatial context to assume that the average cost is falling forever. Within the range of actual observations in Figures 2-8 it is found that a linear function with a positive abscissa fits the data quite well. This does not, of course, mean that the shape of a LRTPC function like (1) above applies in the whole output range - from zero onwards. It is here assumed that this shape applies only within the output range where q is larger than some threshold quantity (of, say, q_0). Such a truncated, linear LRTPC function can be derived from very reasonable underlying characteristics of capacity costs and short-run operating costs. As will be demonstrated presently, the chosen cost function is also defined in the initial output range, in which $q < q_0$. For reasons that will soon be clear, this output range will rarely be relevant, however. In order to derive a long-run cost function that possesses the desired properties, the following should be assumed:

(i) The capacity cost per unit of capacity is falling with the increase in capacity, a fact that is empirically well documented, see e.g. Haldi and Whitcomb (1967, p. 383).

(ii) The operating unit cost is falling with the increase in capacity, which is also well documented – big plants are typically labour-saving, see e.g. Pratten, 1971, pp. 4-5.

Convenient forms of the total capacity cost and the total operating cost that take these characteristics into account are:

$$\text{Total capacity cost} = c_1 + a_1 K \tag{2}$$

$$\text{Total operating cost} = \left(\frac{c_2}{K} + a_2\right) q \tag{3}$$

where

K = capacity ≥ q

q = volume of output

and a_1, a_2, c_2 are constants > 0, $c_1 \geq 0$

As the capacity gets very large, the capacity unit cost will approach the constant a_1, and the operating unit cost will approach the constant a_2. The constant c_1 in (2) may appear to be a half-measure - just a reduced counterpart of the fixed costs in (1) - and ($c_2/K + a_2$) may appear to be an unnecessarily involved expression for the marginal cost. However, these extensions do have some bearing on the real world that should be taken into account. Take, for example, a patent on an efficient production technology for a particular product. The patent may be sold by its holder, and therefore represents an opportunity cost. This motivates the inclusion of the capacity-independent constant $c_1 \geq 0$ in (2). As for the operating costs, it has been long observed that marginal production costs tend to decline with increases in plant size - that is, capacity. This motivates the inclusion of c_2/K in (3).

Combining the capacity cost funtion and the short-run operating cost function, the capacity variable K can be eliminated, and the Long-Run Total Production Cost can be expressed as a function of just volume of output (q) by minimizing the total cost for every level of output, observing that $q \leq K$. Form the Lagrangian expression L:

$$L = c_1 + a_1K + (\frac{c_2}{K} + a_2)q - \lambda(K-q) \tag{4}$$

and find the Kuhn-Tucker conditions for minimum total costs:

$$\frac{\partial L}{\partial K} = a_1 - (\frac{c_2}{K^2})q - \lambda \geq 0 \tag{5}$$

$$K\frac{\partial L}{\partial K} = 0 \tag{6}$$

$$\frac{\partial L}{\partial \lambda} = K - \mathrm{q} \geq 0 \tag{7}$$

$$\lambda\frac{\partial L}{\partial \lambda} = 0 \tag{8}$$

The capacity constraint can be binding or not binding. Where it is binding ($q = K > 0$) λ takes a positive value, and from (5) it can be concluded that $q > c_2/a_1$. If $K > q$, it follows from (7) and (8) that $\lambda = 0$. The capacity can now be found from (5):

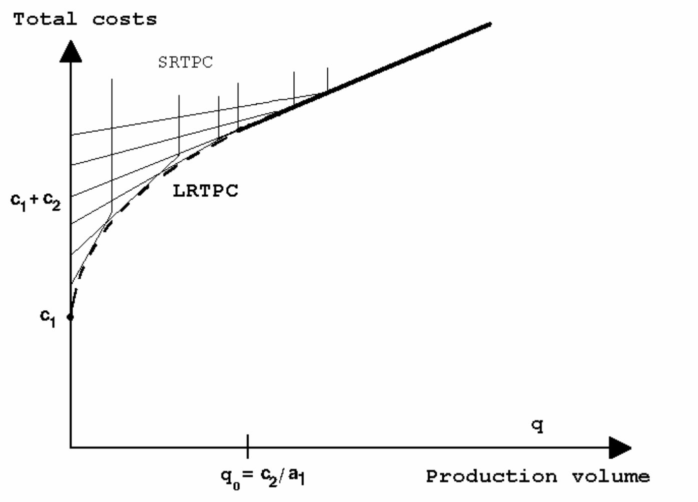

Figure 9. Short- and Long-Run Total Production Cost Functions.

$$K = \sqrt{\frac{c_2 q}{a_1}}\text{, in the range where } q < c_2/a_1 \tag{9}$$

By substitution of (9) and $K = q$, respectively, in (2) + (3), it is evident that the LRTPC-function, consists of two parts:

$$\text{* LRTPC} = c_1 + 2\sqrt{a_1 c_2 q} + a_2 q\text{, for } q < \frac{c_2}{a_1} \tag{10a}$$

$$\text{* LRTPC} = c_1 + c_2 + (a_1 + a_2)q\text{, for } q > \frac{c_2}{a_1} \tag{10b}$$

The cost functions (10a) and (10b) are illustrated in Figure 9 in which a family of Short-Run Total Production Cost (SRTPC) curves are tangent to the (dashed part of the) long-run curve up to the $q_0 = c_2/a_1$ production volume. The dashed part of the LRTPC function in Figure 9 is the graphical counterpart to the function (10a), whereas the heavy straight line for production volumes larger than c_2/a_1 is the counterpart to (10b).

Note that the SRTPC-curves in Figure 9 stop at the capacity limits. This occurs initially to the right of the point of tangency of the SRTPC- and LRTPC-curves, but the discrepancy diminishes as the SRTPC-curves flatten. Eventually the tangency point coincides with the production capacity limit – at $q_0 = c_2/a_1$. In the range where $q < c_2/a_1$, installing capacity that is larger than necessary for the planned volume of output is optimal, see e.g. Cheenery (1953, pp. 320-321). This can be justified by the operating cost savings realised by a larger capacity.

As explained previously, it is not necessary - either in the theoretical modelling or in the empirical studies - to take into consideration what shape the cost function may take where q < c_2/a_1. Instead, the assumption is now made that the only “relevant range” is the one to the right of c_2/a_1. That is, where:

$$q > q_0 = \frac{c_2}{a_1} \tag{11}$$

Thus, only the second part of LRTPC (denoted TPC from now on) will be considered. It is apparent that, given the units of measurement of output, the economies-of-plant size will be more pronounced the higher the ratio of:

$$\frac{c_1 + c_2}{a_1 + a_2} \tag{12}$$

will be. This ratio is given a designation of its own, namely *b*. If, in addition, (a_1+a_2) is put together and denoted *a*, (12) may be written:

$$b = \frac{c_1 + c_2}{a} \tag{13}$$

and, consequently:

$$ab = (c_1 + c_2) \tag{14}$$

Now (10b) can be written:

$$\text{TPC} = a\text{b} + aq,\ q > q_0 \tag{15}$$

and the (long-run) Average Production Cost (APC) is written:

$$\text{APC} = a + \frac{ab}{q} \tag{16}$$

which is a cost function of the kind sought-after in this paper.

6. CONCLUDING REMARKS

Function (15) is similar to cost functions frequently used in spatial economics - that is, equation (1). It is worth stressing, however, that (10b) and (15) are only defined for $q > q_0$, and that *ab* should not be literally interpreted as the fixed costs. The “long-run fixed costs” are $c_1 \geq 0$, which are less than *ab*, and possibly zero. Thus, the cost function in (15) is a cost function that is different from cost functions like that of (1). Nevertheless, the cost function developed in Section 5 is not only a new cost function, but may also serve as justification for a careful use of cost functions of the kind: TPC = $F + c \cdot q$ in spatial economics long-run equilibrium modelling.

Furthermore, (16) shows that as output volume goes to infinity, average production cost approaches *a*, but can never be below *a*, no matter how large the output volume is. Therefore, *a*, the marginal cost, can serve as a convenient proxy for the unit-value. This an advantage of the present cost function compared to e.g. homogeneous production functions with a scale-elasticity greater than unity, in which there is no limit to the fall in APC as the output volume increases (in such a case the concept of unit-value becomes quite ambiguous). In Wall (2001) it is argued that the standard *Location Theory* model - which typically includes:

- fixed costs or some indicator of the degree economies-of-scale in the production
- the transport costs per unit distance of the product in question
- the density of demand

as independent variables in explaining density of plants or market area size – benefits from the inclusion of the unit-value of the product as a fourth explanatory variable. It is shown in Wall (2001) how the present cost function (15) can be used in developing such a model.

REFERENCES

Arnott R. (1987) Spatial economics. In: J. Eatwell, M. Milgate and P. Newman. (eds), *The New Palgrave Dictionary.* MacMillan: London.

Beckmann M.J. (1968) *Location Theory.* Random House: New York.

Cheenery H.B. (1953) Process and production functions from engineering data. In W. Leontief (ed.), *Studies in the Structure of the American Economy*. Oxford University Press: New York.

Eaton B.C. and Lipsey R.G. (1976) The non-uniqueness of equilibrium in the Löschian location model. *American Economic Review*, 56, pp. 77-93.

Haldi J. and Whitcomb D. (1967) Economies of scale in industrial plants. *Journal of Political Economy*, 75, pp. 373-385.

Hands D.W. (1991), *Introductory Mathematical Economics.* D.C. Health and Co., Lexington, MA.

Lewis W.A. (1945) Competition in retail trade. *Economica*, 12, pp, 202-234.

Mohring H. and Williamson H. (1969), Scale and “industrial reorganisation” economies of transport improvements. *Journal of Transport Economics and Policy*, 3, pp. 251-271.

Miller E.M. (1977) Selection bias in the estimation of cost and production functions. *Journal of Economics and Business*, 30, pp. 79-81.

Pratten C.F. (1971) *Economies of Scale in Manufacturing Industry*, Occasional Paper, No. 28, Department of Applied Economics, University of Cambridge Press: Cambridge.

Smith A. (1976]) *An Inquiry into the Nature and Causes of the Wealth of Nations*, Claredon Press: Oxford, p. (1776

Wall R.E. (2001) *The Importance of Transport Costs for Spatial Structures and Competition in Goods and Service Industries.* Linköping Studies in Management and Economics, Dissertation No. 50. Linköpings Universitet, Sweden.