COMPETITIVE AND SUSTAINABLE GROWTH
(GROWTH)
PROGRAMME

UNITE

UNIfication of accounts and

marginal costs for Transport Efficiency

Annex A6

Deliverable 10: Infrastructure Cost Case Studies

Swedish Seaport Case Study:
Price-relevant marginal cost of Swedish seaport services

12 March 2002

Authors:

Jan Owen Jansson and Robert Ericsson

Contract: 1999-AM.11157
Project Coordinator: ITS, University of Leeds

Funded by the European Commission
5th Framework – Transport RTD

UNITE Partner Organisations
ITS/UNIVLEEDS (UK), DIW (De), NEI (Nl), CES/KUL (Be), TIS.PT (Pt), IWW/UNIKARL (De), VTI (Se), IER/USTUTT (De), CERAS/ENPC (Fr), HERRY (Au), EIET/ULPGC (Es), ISIS (It), STRATEC (Be), SYSTEMA (Gr), JP-TRANSPLAN (Fi), VATT (Fi), ECOPLAN (Ch), INFRAS (Ch), EKONO (Fi), EKI (Se)

UNITE

1999-AM.11157

UNIfication of accounts and marginal costs for Transport Efficiency

Annex A6: Swedish Seaport Case Study: Price Relevant marginal cost of Swedish seaport services

This document should be referenced as:
Jan Owen Jansson and Robert Ericsson Annex A6: Swedish Seaport Case Study: Price Relevant marginal cost of Swedish seaport services – UNITE (UNIfication of accounts and marginal costs for Transport Efficiency) Working Funded by 5th Framework RTD Programme. ITS, University of Leeds, Leeds, August 2002.

12 March 2002
Version No: **1.1**
Authors: as above.

PROJECT INFORMATION

Contract no: 1999-AM.11157:
UNIfication of accounts and marginal costs for Transport Efficiency

Website: www.its.leeds.ac.uk/unite

Commissioned by: European Commission – DG TREN; Fifth Framework Programme
Lead Partner: Institute for Transport Studies, University of Leeds (UK)

Partners: ITS/UNIVLEEDS (UK), DIW (De), NEI (Nl), CES/KUL (Be), TIS.PT (Pt), IWW/UNIKARL (De), VTI (Se), IER/USTUTT (De), CERAS/ENPC (Fr), HERRY (Au), EIET/ULPGC (Es), ISIS (It), STRATEC (Be), SYSTEMA (Gr), JP-TRANSPLAN (Fi) VATT (Fi), ECOPLAN (Ch), INFRAS (Ch), EKONO (Fi), EKI (Se)

DOCUMENT CONTROL INFORMATION

Status: 1st submission to DG TREN
Distribution: UNITE partners
Availability: Public (only once status above is “Accepted”)
Quality assurance: C A Nash, University of Leeds
Co-ordinator’s review:

Signed: **Date:**

SUMMARY

The costs of transport infrastructure services are site-specific to a much higher degree than the costs of, say, manufacturing goods, and this is particularly true about port services. The geographical conditions for the harbour are one factor which can make port costs rather different, and in modern times, as the city has expanded around its original port, the availability of the necessary backup land, and opportunity cost of the land are additional important factors for port economy, and the level of the marginal cost of port services. Where land is very scarce, the other factors of production, which are variable in the short run, tend to be overstrained, as it were, in comparison to ports where there is an abundance of land for port operations. The main short-run variable factors are stevedoring labour and laytime of ships. One represents a producer cost, and the other a user cost. In the general expression for the price-relevant marginal cost (MC) below, the first term mainly stands for stevedoring costs, and the second term, representing the price-relevant user cost component, for expected additional laytime costs of ships caused by a new arrival.

$$MC = MC_{prod} + Q\frac{dAC_{user}}{dQ} + MC_{ext} \qquad (1)$$

The third term of the general expression, the system-external marginal cost is normally less important in seaports, and has been discussed only briefly in this study.

The output measure, Q stands for “throughput”, i.e. tons of goods carried through the port between sea and land transport vehicles. The heterogeneity of the cargo can be great (unless everything comes in containers of a standard size), and the differentiation of port charges by commodity type can be intricate, if a marginal cost-based tariff of charges is aimed at. In this study this problem is only touched on: a throughput aggregate is mainly assumed in the calculations.

Two port categories

With reference to expression (1) above, we could soon establish that the result of short-run marginal cost pricing would be very different in two distinct categories of seaports.

A) Where the port is still in its original place, along the river, more or less in the middle of the city, or at a bay constituting the city waterfront surrounded by the buildings of the old town, the marginal cost of port throughput can be relatively high; especially the middle term of (1) can go up very much, in case the demand for port services has kept up with the general economic development.

B) Where the port has already moved out of town, to a site where abundant backup land exists, and the water depth is greater, port operations can be characterized as a pronounced decreasing-cost activity; short-run marginal cost pricing would not, by a long way, pay for the fixed port facilities.

Seaports of category A are rare nowadays in Sweden. Where the port still remains in the original location in, or at the edge of the old town, demand has typically been on the decline. Since industry and wholesaling activities have anyway moved out of town, competing “out-ports”, miles away from the old seaport, have taken over most of the business. There are hardly any advantages any more of a central city location for ports.

The case study port

Therefore, it seemed appropriate to consider a port of the latter category for the main case study; the port of Norrköping was chosen. Like many other Swedish ports, this port is a fully integrated joint stock company. The previous stevedoring companies, and the port authority have merged in 1990. However, this joint stock company does not keep cost records like olden day port authorities, and can legally keep some information that we wanted secret, e.g. on actual prices charged to different customers. When we started to inquire about short-run variations in stevedoring costs with respect to throughput volume, which could be interesting to observe on a monthly, or quarterly basis, we found that no figures of throughput were available except annual records. Moreover, concerning the second, potentially important item of the price-relevant marginal cost, the queuing costs of ships, we were told that there is no queuing of ships in the port of Norrköping, and, anyway, if it in fact occasionally happens that a ship has to wait for a berth, no one keeps a record of this. This state of affairs is not a reflection of a stagnant demand for port services in Norrköping. On the contrary, total throughput has been on the increase most of the time which could be surveyed, and especially so in the last ten years. This made us decide to try out the estimation of “the development cost” of port services in this case study, that is a variant of the long-run marginal cost. This has often been advocated as an alternative to short-run marginal costing as regards transport infrastructure services, where a steady increase in demand leads to a continuos expansion of the system.

Evidence of the price-relevant short-run marginal cost of port services

Before the development cost calculation for the port of Norrköping was carried out, however, we gathered from the literature, and our own previous work evidence of the price-relevant short-run marginal cost of port services, in order to have something to compare with. This tends to be a treatment of ports belonging to category A, first because the relevant literature is mainly operations research in the form of queuing model applications to seaports, and, secondly, because our own previous work, which is most interesting in the present connection, is a study of the port of Uddevalla, which is one of the few category A ports in Sweden.

The port of Uddevalla used to be organizationally divided mainly between a port authority responsible for the port capital including the approach channel, quays, cranes and storage facilities, and a stevedoring company doing the cargo handling. In this case hardly any short-run producer marginal cost worth mentioning is directly assignable to the port authority, i.e. the capital owner. Optimal port charges (as distinct from stevedoring charges) are, in this case constituted by (i) the “quasi-rent” included in marginal cost-based stevedoring charges, and (ii) the price-relevant queuing cost of the ships, and (iii) a “congestion cost” which should be separated from queuing cost. The queuing cost is clearly defined by the time cost of ships, which do not find a vacant berth, but have to wait in the roads, while the congestion cost takes the shape of prolonged service time for the ships at the berths, and/or lower stevedoring labour productivity caused by congestion on the quays and in the sheds for transit storage. With reference to the MC- expression (1) above, which applies to integrated port operations, the first term, MC_{prod}, stands for cargo handling marginal cost. Normally this marginal cost is above the average cost of cargo handling. The difference constitutes a quasi-rent, or “contribution margin”, which results from marginal cost pricing of the stevedoring services. This should be regarded as contributing to the coverage of the fixed capital costs.

Where port capital and labour are organizationally separated, the quasi-rent/contribution margin could be expressed as a rent of quay space to be paid by the stevedoring company.

The price-relevant queuing cost is one part of the second term of (1) above, which could be charged on the ships per day of berth occupancy. Another part of the second term of (1) is the "congestion cost". It should be charged on the shipowner, but preferably per ton of cargo, because in this case it is not the berth occupancy as such, which causes the cost, but the amount of cargo to be (un)loaded.

The sum of the total rent paid by the stevedoring company, and the berth occupancy charges, and congestion charges on the goods paid by the shipowner is to be set against the total port capital cost. In case constant returns to scale apply in port operations, the revenue from rents and occupancy charges should just cover the total capital costs, and, in addition, the marginal cost-based stevedoring charges on the shipowners and/or ware-owners should just cover the total costs of the stevedoring company consisting of rent of quay space (and cranes), and labour wage costs.

For the port of Uddevalla these three different charges were estimated, and the corresponding revenue from optimal pricing was calculated. Setting total port capital cost equal to 100, the relative total revenue, and its components are given in the table below.

Table 1: **Total cost and revenue from optimal pricing in relative terms in the port of Uddevalla**

Total port authority cost	**100**
Total revenue from optimal port charges	**87**
Of which berth occupancy charges	22
Congestion charges	20
"quasi-rent"	45

As seen total cost recovery would not be obtained from charges based on the price-relevant marginal cost, although the deficit to be covered by other means is, relatively speaking, not very large.

The structure of charges

In the end it is, of course, the ware-owner who pays it all. In fully integrated ports, it is possible, and often considered very desirable to simplify the port tariff, traditionally consisting of numerous separate charges, which each can be differentiated in elaborate ways. In the Swedish discussion in connection with the organizational streamlining of ports (integration of capital and labour), taking place in recent decades, it was often held up as an ideal that just one price per ton of goods of a limited number of commodity types should be charged, on one hand, on the shipowner for the cargo movement between the approach channel and transit storage, and, on the other hand, on the ware-owner for the cargo movement between transit storage and the gates at the land-side of the port. In retrospect two things can be said about this urge to simplify port tariffs.

First, price differentiation that really reflects price-relevant marginal cost differences should not be swept out of the way, unless the price differentials could be judged to have a negligible effect on the behaviour of actors in the throughput process, including shipowners and ware-owners. It can be argued that in some instances *more* price differentiation than is normally practised should be beneficial, for example, the kind of peak-load pricing described in the case study in chapter 6.

Secondly, it is interesting that recently the EU commission has indirectly argued for a return to a more disintegrated port organisation, by holding up the importance of competion in port operations. In particular, the traditional separation of port capital and stevedoring labour has been pointed out to be a prerequisite for cost efficiency in the cargo handling, i.e. by putting pressure on stevedoring companies, either by new entries (of more than one stevedoring company per port) or by the possibility to introduce competition *for* the market by a tendering process.

The “development cost” of the port of Norrköping

The financial result of optimal pricing calculated for the port of Uddevalla may, or may not be characteristic of other seaports. It depends above all on whether category A or B applies, and in the former case, on to what extent the central city location is a very strong restriction on expansion of backup land. The main case study of this project is a “development cost” calculation for the port of Norrköping, which during the period of observation – from 1960 to 2000 – has expanded considerably, and moved most of its business out from the mouth of the small river “Motala ström” at the edge of the central city of Norrköping to sites further out in the bay Bråviken. Queuing and congestion are reported to be nearly non-existent in the port, in spite of the fact that total throughput has trebled in the period of observation.

The “development cost” is an alternative name for the long-run marginal cost which indicates that capacity expansion over time is in focus. The implicit idea is that pricing policy should prevent over-expansion, which might follow from not taking the capacity development cost into account, which is seemingly done by only considering the short-run cost. By long time-series analysis the full effect on the costs of the capacity expansion “caused” by growing demand could be ascertained.

In principle, expression (1) above for the price-relevant marginal cost has exactly the same appearance in the short run as in the long run. Each term, however, has a completely different content in the long-run price-relevant marginal cost. The first term, MC_{prod} now contains the port capital cost as well as the running cost including cargo handling labour cost. The middle term, $Q\ (dAC_{user}/dQ)$ is now a product of total throughput and the change in the shipowners costs per ton as a result of an increase in throughput accompanied by the actual capacity expansion taking place. It turns out – here like in other better known cases of transport infrastructure expansion – that the middle term of (1) will be a relatively large *negative* component of the price-relevant marginal cost, which tends to make the final value coming out of the development cost calculation quite low; much lower the average total port cost.

In the case of the port of Norrköping the development cost was found to be virtually zero, which is consistent with the reported absence of queuing ships and congestion in the port. We thus conclude that a clear case for Ramsey pricing exists for seaports of category B. In the port industry this is a well-known, and, since long a widely practiced pricing principle under the motto of “charging what the traffic can bear”.

CONTENTS

1. PURPOSE, BACKGROUND, AND CURRENT ISSUES

1.1 Purpose and plan of the study

This study aims at estimating the price-relevant marginal cost of seaport services. It starts by describing the seaport production process, and outlining the historical background to current policy issues (chapter 1). Then the methodology for the marginal cost analysis is presented (chapter 2). First comes the short-run marginal cost analysis (chapter 3-6), and then the long-run marginal cost analysis (chapter 7). The last-mentioned chapter is the most substantial and original piece of research based on observations over 40 years in the port of Norrköping. It is a special challenge for this study to show empirically that long-run marginal cost pricing would give basically the same result as short-run marginal cost pricing, as elementary microeconomic cost theory proves, but which nevertheless is often denied.

The last chapter, (chapter 8) contains a summary and conclusions, and in addition a brief discussion of the inevitable attendant question, when it has been found that marginal cost pricing would not by a long way cover the total producer costs: what is the second-best pricing policy for seaports under a strict budget constraint?

1.2 Basic characteristics of the production process of seaport services

1.2.1 Facilities for change of mode of transport vs. ”terminals”

The main function of the port is to transfer goods from land to sea transportation and vice versa, i.e. to provide the facilities for changing the mode of freight transport. There are also “seaport terminals” (Bird, 1971), i.e. industrial plants located at the waterfront, where ships unload their cargo more or less directly into the goods manufacturing process. A seaport terminal is literally the end of the journey for the input goods. Seaport terminals are sometimes called “industrial ports”; they are usually private, for the exclusive use of the industrial plant owner. In this study only common user ports are discussed, where the goods are not processed, but carried through the port.

The complete process of getting cargo transferred through the port can be schematically divided into seven principal links, summarized in figure 1 below. Besides the seven links shown in this figure, there are a number of other functions, such as customs inspection, warehousing in the port area, and preparation of cargo by preslinging, stuffing containers, and so forth. Such functions are supplementary by nature, rather than intrinsically part of the transfer between sea and land, and could in principle be performed elsewhere. It is also important to note that port design has consequencies for further links in the total door-to-door transport chain. Water depth in the port approach channel, and along the quays, for example, can have a profound effect on sea transport costs by making the use of bigger ships possible.

The output of the port is most simply defined in terms of tons per unit of time passing through the port in both directions (export and import), that is *throughput.*

1.2.2 The chain is as strong as its weakest link

This is an outworn phrase but nevertheless a useful point of departure for describing basic port production technology. In every activity divisable into stages, or links, it is generally desirable to equalize the capacity of each link. When the potential capacity of one link has been increased relatively by some innovation or other, other links of the chain should also be improved in order to realize the full potential of the original innovation.

Technical developments should be so canalized that the rate of increase in strength of each link of the chain is, on average, the same. In the short term, however, disharmony may occur from time to time on account of the inevitable short-term random occurrence of major innovations. Containerization, for example, has broken the bottleneck in seaborne general cargo transport that has prevailed for centuries – the stowage and unstowage of cargo in ships' holds. Now general cargo can be loaded and unloaded from ships almost as rapidly as bulk cargo.

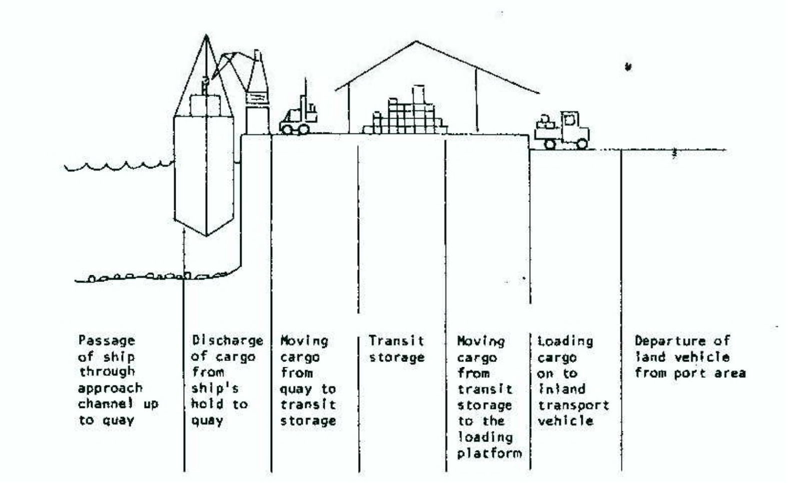


Figure 1.1 **Schematic picture of import berth**

The following identity of throughput demand and throughput supply is a simple starting-point for a brief discussion of seaport development.

$$Q = ⎛ \cdot n \cdot \mu, \qquad (1.1)$$

Where

Q = total port throughput;

⎛ = expected berth occupancy rate;

n = number of berths

μ = expected throughput capacity per berth.

In the short run, a sudden increase in demand is met simply by an increase in the occupancy rate. This way out has an obvious limit, given the more or less random pattern of the arrivals of ships; well before an occupancy rate of 100% is attained, the queuing of ships will be intolerable.

In the past, fairly long periods of port congestion have been experienced due to sharp increases in demand. The import boom of the *nouveau riche* oil producing countries in the 1970s is an extreme example. The port of Lagos (Nigeria) was heavily congested during the years 1974-1977, where in 1975 the average waiting time of ships was 240 days! Similarly, in Saudi Arabia, to relieve pressure on the heavily congested ports, helicopters were used in those days to unload ships and carry cargo to the inland destinations. However, in the long run, the rate of growth in demand has to be matched by an expansion of capacity. This can either take the form of additional berths (n) or increases in throughout capacity per berth (μ). When the cargo handled by the port is rather heterogeneous, a third means of increasing capacity may be to differentiate the port into specialized sections for different types of cargo such as bulk, unitized cargo, or break-bulk cargo. The main present seaport policy issues, however, are bound up with the fact that the origins of export and destinations of import through the old port have to a large extent moved away from the central city. Manufacturing industry and wholesaling are no longer typical central city activities. And if, in addition, the water depth of the old port is insufficient for modern bulk carriers and containerships, the option of port relocation becomes very relevant.

1.3 Historical background to current seaport policy issues

Historically, the development of many ports can be roughly divided into three eras: a long period of capacity expansion by increasing the number of berths, followed in the two first postwar decades by a period of berth capacity improvement, and after that the present era of relocation of the old port. Some ports have entered the third era long ago, and others are just entering. Still others may not enter it at all, in case the throughput is stagnant, but it will surely be increasingly difficult to accommodate a steadily expanding throughput within the limits of the historic port site.

1.3.1 The era of berth number expansion

The loading and unloading of ships have long been a severe bottleneck for seaborne trade. A port that was situated along an estuary responded to increases in throughput by expanding in a lineal fashion along the river banks until it was constrained by some natural obstacle or until internal transport costs within the port area became excessive. Where the port town had been built up around a natural harbor – a protected firth with fairly deep water – the stage when the extension of capacity in this fashion could not go on any longer was reached sooner.

If the width of the port cannot be further extended, how can capacity match demand as trade continues to grow? The solution found was to increase the number of berths within a given width of the port area. This was achieved by the *finger pier* configuration that can still be found in old sections of long-established ports.

In the early nineteenth century, sailing ships were still relatively small, and so dock engineers strove for the maximum quay length possible in a given area. If the water site permits … peninsular jetties afford a great length of quay line within a small compass. (Bird 1971, p69)

In the old days, exchanges between sea and land transport vehicles were made in the most direct of ways. When a ship had arrived and berthed, the merchants sent down their wagons to the ship’s side and took direct delivery of their goods as they were handed over the side.

With further increases in Q, which could be met by increases in n, congestion *ashore* became a problem. The port had come to be situated in the city nucleus, so that the backup land of the quay was restricted. The traffic problem ashore became more and more serious as the number of ships admissible to the port at one time increased. Interference between streams of traffic to and from the ships increased as the quayage became more elaborate. This affected also the turnaround times of ships.

1.3.2 The era of berth capacity improvement

Increases in loading and unloading capacities at general cargo berths were a main prerequisite for the continuous growth of ship sizes, which induced the ports of industrial countries increasingly to choose an indirect route for the throughput. The indirect route between sea and land transport is through transit storage.

The basic idea of the indirect route is that it enables both ships and land transport vehicles to maximize the speed of handling and to minimize turnaround times by making the operations independent in the short run. In the case of imports, cargo is transferred from ships’ holds to the transit shed, and after a shorter or longer period of time the shipments are collected by the importers.

The introduction of the indirect route meant in general that μ could be increased considerably. This did not necessarily mean, however, that the total capacity of the port was increased commensurately. The port was no longer a one-stage production plant. The cargo had to pass through transit storage and could be held up there. A common problem was the restricted backup land for storage (and traffic): Given the average time of storage of shipments, the holding capacity of the storage facilities has to be roughly proportional to throughput.

One solution to the problem of storage area scarcity was to build multistory sheds and warehouses. This obviously puts further demands on the already labour-intensive cargo handling. And rising dock labour wages – a consequence of the growing strength of unions – were to make multistory sheds uneconomical.

In parallel with stevedoring labour wages, seamen's wages were steadily increasing. The need to reduce the labour content in shipping was mainly met by making ships bigger (the increase in the number of seamen is much less than proportional to the increase in ship size). But this expedient depended to a large extent on the possibility of reducing turnaround time in port.

Preslinging and palletization were important means of reducing both turnaround times and port handling costs by combining small packages into large units. But the bigger units could no longer be handled manually; the forklift truck made its appearance. But the efficient use of forklift trucks required much more space than was previously necessary. To avoid delay of ships, cargo had to arrive at quayside before the loading started. This created additional demands for space, both in the shed and in open areas. The finger pier configuration became a severe bottleneck and was gradually eliminated by replacing the slips and piers by wide lineal wharves constructed parallel to the waterfront. This also provided the necessary space for mechanised cargo handling. Each new berth occupied the space previously used by about three berths of the finger pier type. Berth throughput could be increased three times or more. Much bigger (longer and wider) ships could now be accommodated. At the same time, the input of a ship's time and dock labour per ton throughput was reduced substantially.

This development was not without its problems. Given the conventional gear for cargo handling, the bigger units increased the danger of the work and the risk of damage. Furthermore, as cargo had to lie in the open for rather long periods without protection, unpredictability of rain became an awkward problem (in the temperate zones).

The container was the logical continuation of the trend toward bigger units. It was an answer to several demands. By standardization of container dimensions, expensive tailor-made container cranes of very high capacity could achieve sufficiently high degrees of capacity utilization to be profitable. Now the weight per crane cycle could be raised to 10-30 tons (depending on container module); and by eliminating the need for stowing the cargo in the ship's hold, the continuos work of the crane was not impeded by that notorious bottleneck.

When two or, as sometimes happens, three container cranes are working at a time, handling speed may well be some ten times greater than can be achieved by conventional break-bulk cargo handling. Roll-on/roll-off methods of loading and unloading big standardized units constitute, for similar reasons, an equally dramatic improvement of capacity.

Concerning bulk cargo, there has been a similar rise in throughput capacity. But bulk handling has always been more amenable to mechanization and to continuous handling, so the rate of increase in throughput capacity per quay length has been less dramatic.

Handling capacity multiplication for general cargo

How is it possible that the handling capacity per crane in a container system can be almost 25 times higher than the capacity per crane in a traditional break-bulk system? So far as break-bulk cargo handling is concerned, the unloading operation consists of three sublinks:

1. the hold operation, which makes up the set to feed the hook;
2. the hook operation, which brings the set from hold to quay apron;
3. the apron operation, which takes the cargo away from the quay to a temporary place of rest (or directly to transit storage).

A good deal of operations research work and numerous time studies have been devoted to identifying and eliminating bottlenecks at break-bulk berths. Interest was focused on these three sublinks, as it was thought that the prevailing bottleneck of the entire throughput system was there. The intricate stowage and unstowage operation in the restricted area of a ship's hold was found to be the inherent limiting factor of break-bulk cargo handling.

Table 1.1 *Break-bulk handling capacity in the 1970s*

Port	Tons per gang-hour
Antwerp and Rotterdam	18-25
Bremen and Hamburg	14-22
Gothenburg	12-20
New York	12-18
Chicago	12-20
Long Beach and San Francisco	10-15
Karachi	6-22
Valparaiso	10-18
La Valletta	12-15

In the beginning of the seventies the performance per gang-hour (or crane-hour) in different ports seemed to be in the range of 10-25 tons/gang-hour. Some examples are given in table 1.1. This is to be compared to a capacity of about 250-400 tons/container crane-for at container berths. One has to bear in mind, however, that whereas as a rule one or two container cranes are used per container ship, at least three to five cranes were user for (un) loading a conventional break-bulk ship.

The basic point of cargo unitization is the strong correlation between handling capacity and the package unit weight (figure 1.2). Especially stowage in ship's hold, was very time-consuming when the goods came in package units that were sufficiently small to be man-handled.

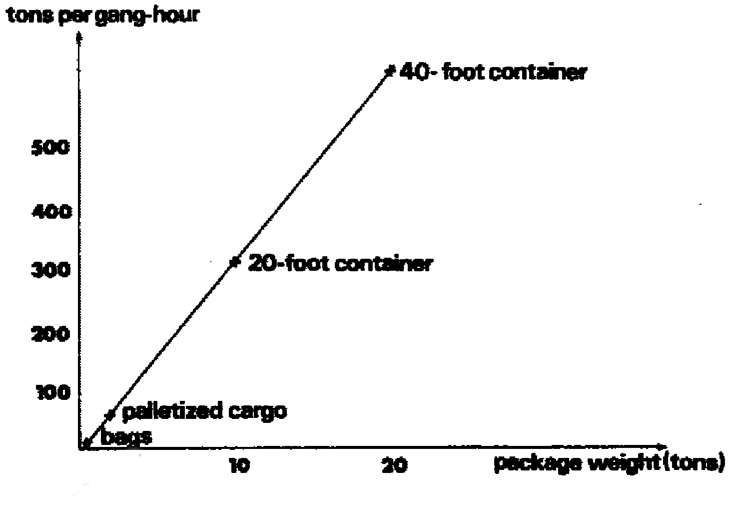


Figure 1.2 **Schematic linear relation between stevedoring productivity and package unit weight**

Hoisting crane capacity has never been a bottleneck. When containers tailor-made to cellular container ships made the stowage operation redundant, the full potential of the hoisting capacity could be profitably realized. The elimination of the need for breaking bulk in the ship's hold as a link in the loading operating, and for making up a set of many small parcels in the ship's hold as a link in the unloading operation, were the key innovations of containerization.

1.3.3 Increasing demand for backup land – the era of seaport relocation

The picture is incomplete as long as the land requirement per throughput ton is ignored. Land requirements for container berths, roll-on/roll-off, and packaged timber and other big unit loads are very much greater than for conventional break-bulk berths. Table 1.2 gives estimates of throughput, for various cargo-handling techniques and land requirements.

Table 1.2 *Rough estimates of land requirements per berth and per throughput ton*

Cargo-handling technique	Land area per berth (hectares)	Throughput per berth and year	Throughput per hectare and year
Conventional Break-bulk	1-2	100.000	75.000
Palletized cargo	3-4	200.000	60.000
Containers	7-10	500.000	60.000

In terms of relative strength of the port links, a new bottleneck has arisen owing to cargo unitization, and especially containerization. The capacity to load and unload ships – the traditional bottleneck – cannot now be matched by the capacities of the other port links designed for traditional break-bulk handling. What is needed more than anything else is supporting land to store the cargo in transit and to move the big unit loads about in feeding the container cranes. In particular, if the obsolete finger pier configuration still remains with hardly any backup area at all except for some narrow sheds, the problems of adaptation to the container age will be very pressing. In addition, the water depth is insufficient in many older ports to accommodate the new big ships.

The problem is manifest in the tendency of ports to move to cheaper sites and deeper water. In some old ports, where cities have grown up around the port, this is not always possible. In San Francisco, for example, no site for a modern container port was available at the right moment, and as a result the site for a container port was chosen in Oakland – across the San Francisco Bay from its historical port – where appropriate land was plentiful. In New York City all major improvements to cargo piers on the Manhattan waterfront have stopped long ago. New facilities are located chiefly at Port Elizabeth, New Jersey. The development of the port of London is a well-known example of the need to build new berths farther and farther down the river estuary, and now Tilbury at the mouth of the river Thames is the only remaining terminal.

The shipload size determines the required backup area. The second-generation container berths were built to comprise a paved area twice as large as the first-generation ones – about 16 instead of 8 hectares. The practical capacity of second-generation berths is like-wise about twice that of first-generation berths, which leaves throughput per hectare roughly unchanged.

The reason why ports build larger and larger container berths is simply that containerships are getting bigger and bigger, and so are the shiploads of containers. Space has to be provided that is at least proportional to the size of the shipload. There are usually two resting areas for the containers, one close to the quayside and the other farther away. A substantial additional area, which is required for internal movements of the containers, has to be considerably more than proportional to the shipload size for smooth and safe carriage of containers to and from the stacking area. This is a principal source of diseconomies of shipload size.

It is rare, however, that a whole shipload of containers of the largest size is loaded or unloaded in one port. Usually more than one port at either end of a route is called at. For example, in the Northern Europe – Far East trade, the big container ships call at Gothenburg, Hamburg, and Rotterdam in Europe, and at Singapore, Kobe and Tokyo in the Far East.

Alternatively shuttle services between pairs of central ports supplemented by feeder services at either end can be organized. In that case ports have to cope with container flows coming intermittently in lumps of 2,000 containers or more. Crane capacity is not a great problem; it is possible to load and unload 2,000 containers in 30-80 hours, depending on the number of cranes employed to achieve a reasonably quick turnaround. The question, however, is, How soon the next ship can be received? How quickly can a sufficient amount of the backup area be cleared so that another container "avalanche" can be coped with?

1.4 Organizational change in seaport: implications for optimal pricing

The most striking feature of traditional (internal) port organization was the considerable number of bodies that participated in the transfer of goods between ships and inland transport vehicles. For example, the loading and unloading of ships used to be the joint responsibility of the port authority, and a stevedoring company. The storage of cargo and the delivery to/from inland transport vehicles typically involved separate cargo handling companies (besides the stevedoring companies). Other port services could vary greatly from port to port in respect of who is carrying them out. A worldwide survey comparing port service organization made by the UNCTAD Secretariat, brings out this point:

Table 1.3: *Percentage distribution of agents responsible for various port operations around 1970 between the port authority, and other public and private enterprises*

Percentage of cases where the providers of main port service were			
	Port authority	Other public bodies	Private firms
Services to ships			
Aid to navigation	60	29	11
Aid to (un)berthing	51	5	44
Repairs	12	11	77
Services to cargo			
Stevedoring	16	15	69
Other cargo handling	38	19	43
Storage	53	17	30
General services			
Police, fire protection	41	59	0

Source: *Port Pricing*. Report by the UNCTAD Secretariat TD/B/C. 4/110 (United nations. New York, 1973)

What is the explanation for this disintegrated pattern of port organization?
Particularly, why is it that capital (port facilities) and labour (stevedores) are still split up organizationally in many ports of the world? A brief look at historical developments may throw some light on this question.

In the time of the sailing ships, the ship crew was sufficiently large to do all the loading and unloading at the port. When the steamship appeared, the number of crew could be reduced radically. This meant that shore-based labour had to supplement the ship crew in the loading of cargo. This was first done under the direction of the ship's master. However, the foreign shipowner soon wanted to have the cargo-handling operation supervised by a local man who was more familiar with local conditions. Such people became known as *stevedores*. At first, stevedoring firms were not very permanent as a result of the ease of the entry into stevedoring and the reserve of willing hands. Then, toward the end of the nineteenth century, the first stevedoring labour union emerged. The unions quickly gained substantial power. Ships that tried to use unorganized labour were blocked and fines had to be paid to lift the ban. At the same time stevedoring companies of more stable organizational forms were created. Joint stock stevedoring companies with shipowners, and ship brokers as main owners, emerged.

The port authority, however, typically did not engage in stevedoring. This is a puzzling feature in view of the fact that all capital equipment was owned by the port. There are two explanations for this. First, the port facilities do not represent the most important capital input in the process of transferring goods from sea to land. The ship berthing along the quay represents a greater cost per unit of time than either the berth or auxiliary assets. The costly presence of port users in the process makes it natural that the complementary factor of labour should be under the control of the shipowner. The ship operator knows best how many men are required at which times and has a greater incentive to see to it that the fastest and cheapest service is used. His lack of detailed knowledge of local conditions, together with the difficulties of maintaining his interests when at sea, made it desirable to have a man on the spot who could provide casually employed manpower when required. This explains the close association between shipowners and stevedores. Second, most major ports used to be considered as *natural monopolies*. Today, the development of inland transport has diminished the monopoly power of ports. But then, if the port authority were also acting as stevedores, the monopoly power of the port would have been even greater. Shipowners (and wareowners) have therefore safeguarded their interests by direct association with the labour force so as to reduce the monopoly power of the port.

In more recent times the balance of power in ports shifted to the labour unions, and the main problem of shipowners were to find ways of counterbalancing labour union monopoly power.

In one century stevedores have moved from the bottom to near the top of the wage rate scale. Labour relations have replaced port monopoly abuses as the primary concern of shipowners, exporters, and importers.

The mechanization of cargo handling in ports, the growth in ship size, and the increasing use of the more expensive container ships have resulted in some interesting, more recent organizational innovations. There has been a tendency to streamline and integrate internal port organization. This tendency has been accelerated by the concept of door-to-door through service. As a result, “terminal companies” have been created in a number of ports – in practically all ports in Sweden. These are owned by the port users and – typically as a minority shareholder – the municipality. The terminal companies usually control both port capital and labour of the three main links of port operations. They usually do not own port capital, but rather enter into long-term leasing of berths and transit sheds.

1.4.1 Separate port charges and stevedoring charges?

The traditional separation in most ports between stevedoring companies providing cargo handling labour and the port authority providing the capital inputs into the throughput process, and the corresponding division of the total price for transferring goods between sea and land transportation into cargo handling charges and port charges is still the common practice. There is a good number of ports where cargo handling labour and port capital are in fact organized under the same hat, but where the practice of separating the cargo handling charges and port charges is nevertheless maintained. For people of the port industry, who have grown up with this practice, it will probably sound odd that the natural charging practice – “natural” in the sense that practices in other industries would lead an outsider to expect this – would have stevedoring charges per ton of different commodities cover also port capital costs.

Under competitive conditions, the price p is set where the demand curve intersects the short-run marginal cost (SRMC) curve. As the capacity limit is approached, the average variable cost (AVC) will rise. If the output produced is to be found in the range where AVC is rising, which is the normal case, p = SRMC will exceed AVC. A “contribution margin”, p – AVC, will accrue, which may or may not cover the short-run fixed capacity cost. The total contribution to the fixed cost amounts to Q(p – AVC). It will be just sufficient in case constant returns to scale exist. We shall go no further into this matter now; the relevant point in the present connection is that the normal case in most industries is that a single price is levied per unit of output. There is no question of dividing the price in two parts – one for the capital services and one for the short-run variable factor services embodied in a unit of output.

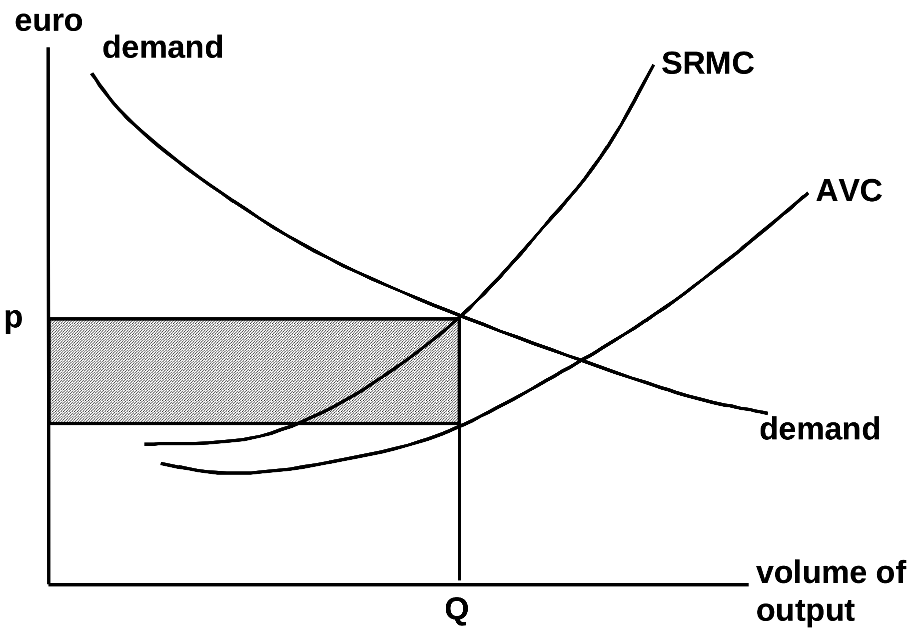


Figure 1.3 **The standard case of covering the capital costs: the residual of total revenue after remunerating the variable factors of production is a "quasi-rent", which should cover short-run fixed cost.**

When it comes to ports, a need for two separate prices may arise for organizational reasons. In this case a natural arrangement, which would be more in line with the standard case, illustrated in figure 1.3, would be for the port authority to *lease* the berths including cranes to stevedoring companies. The stevedoring charges for loading and unloading cargo would then be analogous to the standards case; that is, the handling charge per ton in the first place would be set to equate supply (represented by SRMC) and demand for port throughout. The price would leave a certain contribution toward the recovery of the berth rent. Under competitive conditions, it could be expected that in the long run the contribution margin would just suffice to cover the berth rent and leave a normal profit to the stevedoring company as well.

Interestingly, there is a recent move in the thinking on seaports of the European Commission towards a return to the organizational separation of port infrastructure and stevedoring and other services (European Commission, 2001). This new/old idea seems to be entirely motivated by the objective of creating competitive conditions in markets for port services. Stevedoring services, for example, could more readily be exposed to competition *on* the market, or *for* the market by tendering, if the port infrastructure is owned and run by a separate organization.

1.4.2 Centralized versus decentralized industrial organization

In a national perspective, the port industry has been organized in one of two ways. In some countries ports are organized by a central port authority, which in most cases regulates both investment and pricing in individual ports. In other countries, ports do not obey a central authority but compete with each other. Ports in these cases are usually owned by local authorities, but private ports exist too, and a strong tendency towards privatization can be observed in recent times, where the municipality mainly acts as the landlord.

Both forms of industrial organization have their drawbacks. Opponents of the central port authority concept argue that decision making is inefficient, as local conditions are all important, and the impact of port investment and pricing is regional rather than national. Others claim that lack of central control has led to excess capacity and waste. The competition has taken the form of adding berths and equipment, resulting in excess capacity.

Irrespective of the form of organization, the basic economic principles for pricing and investment with a view to social benefit maximization can be condensed into these two rules:

In the short run, the utilization of the fixed capacity of the facility concerned should be such that the social marginal cost equals the marginal benefit of the users of the facility.
In the long run, or the planning stage, capacity investment should be planned so that the long-run social marginal cost is equal to the marginal benefit.

2. METHODOLOGY

2.1 Theoretical foundation

The last-mentioned, two golden rules for allocative efficiency can be diagrammatically illustrated like in fig 2.1 below. When both efficiency conditions are fulfilled, the short-run and long-run marginal costs intersect.

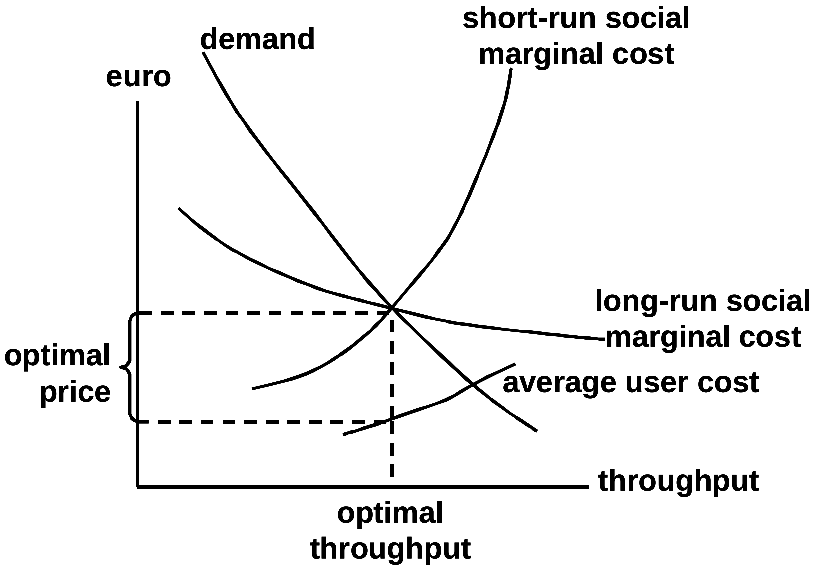


Figure 2.1 **Short-run and long-run efficiency conditions for port operations**

An important characteristic of port services makes it practical to give these rules a slightly different, and more specific, formulation. First, the costs of the port users – shipowners in the first place – are normally a dominant item in the total costs incurred in ports. This fact speaks for adopting an approach well known from road pricing and investment theory, namely: Put the user and producer costs on an equality, and define the total port costs as including both. Figure 2.1 incorporates this procedure. One has then to bear in mind that demand should be viewed as a function of the *generalized cost* (= *charges* + *user costs*) rather than simply the price; consequently, along the vertical axis of the diagram, generalized cost is measured. Another consequence to bear in mind is that the equality between “price” and “social marginal cost” is not applicable when the social cost is the sum of user and producer costs. The pricing rule instead reads “price should be equal to the social marginal cost minus the average user cost”. The distinction merits a name of its own. Thus we introduce the *price-*

relevant marginal cost, MC, which is defined in this way, including also the transport-system external marginal cost.

$$MC = MC_{prod} + MC_{user} + MC_{ext} - AC_{user} \tag{2.1}$$

The difference $MC_{user} - AC_{user}$ can mathematically also be written as the product of total throughput, Q, and the derivative of AC_{user} with respect to Q

$$MC = MC_{prod} + Q\frac{dAC_{user}}{dQ} + MC_{ext} \tag{2.2}$$

2.2 Empirical considerations

Empirical marginal cost analysis can focus on four different areas in accordance with this 2 x 2 matrix:

	cross-section analysis	Time-series analysis
SRMC	1	2
LRMC	3	4

In the 1970s a great deal of empirical research on seaport and shipping cost functions was carried out in Sweden and Israel which is reported in Jansson (1974), Shneerson (1976) Jansson and Rydén (1979), Jansson and Shneerson (1978,1982, 1987). So far as statistical cost analysis the three fields 1, 2, and 3 were all investigated with mixed success. Now it is time to look at 4. The main original contribution of the present study is in this field. It also turned out to be almost a necessity for empirical reasons.

With reference to expression (2.2) above, we could soon establish that the conditions for marginal cost analysis would be very different in two distinct categories of seaports.

A) Where the port is still in its original place, along the river, more or less in the middle of the city, or at a bay constituting the city waterfront surrounded by the buildings of the old town, the marginal cost of port throughput can be relatively high; especially the middle term of (2.2) can go up very much, in case the demand for port services has kept up with the general economic development.

B) Where the port has already moved out of town, to a site where abundant backup land exists, and the water depth is greater, port operations can be characterized as a pronounced decreasing-cost activity; short-run marginal cost pricing would not, by far, pay for the fixed port facilities.

Seaports of category A are rare nowadays in Sweden. Where the port still remains in the original location in, or at the edge of the old town, demand has typically been on the decline. Since industry and wholesaling activities have anyway moved out of town, competing "out-ports", miles away from the old seaport, have taken over most of the business. There are hardly any advantages any more of a central city location for ports.

Therefore, it seemed appropriate to consider a port of the latter category for the main case study. The port of Norrköping[*] belongs to category B. Like many other Swedish ports, this port is a fully integrated joint stock company. The previous stevedoring companies, and the port authority have merged in 1990. We soon found out that this joint stock company does not keep cost records like olden day port authorities, and could legally keep some information that we wanted secret, e.g. on actual prices charged to different customers. When we started to inquire about short-run variations in stevedoring costs with respect to throughput volume, which could be interesting to observe on a monthly, or quarterly basis, we found that no figures of wage costs were available except annual records. Concerning the second, potentially important item of the price-relevant marginal cost, the queuing costs of ships, we were told that there is no queuing of ships in the port of Norrköping, and, anyway, if it in fact occasionally happens that a ship has to wait for a berth, no one keeps a record of this. This state of affairs is not a reflection of a stagnant demand for port services in Norrköping. On the contrary, total throughput has been on the increase most of the time which could be surveyed, and especially so in the last ten years. This made us decide to try out the estimation of "the development cost" of port services in this case study, that is a variant of the long-run marginal cost. This has often been advocated as an alternative to short-run marginal costing as regards transport infrastructure services, where a steady increase in demand leads to a continuos expansion of the system. The interurban road network in the first half of the post-war period

---

[*] Originally, the port of Sundsvall was chosen (which also belongs to category B), i.a. because the general manger was very obliging. Unfortunately, he was forced to retire soon afterwards, and we had better turn to another port.

when the car ownership diffusion was at its very strongest is a case in point. The main hypothesis to be tested by estimating the marginal development cost of the port of Norrköping is that the price-relevant short-run and long-run marginal costs, which are theoretically equal, provided that the optimal investments are made, would also in practice tend to come to the same. In view of the reported absence of queuing and congestion costs in the port, the expectation was that the development cost would be at a comparatively low level relative to the average total cost.

2.3 Evidence of the short-run marginal cost of port services from previous work

Before the development cost calculation for the port of Norrköping was carried out, however, we gathered from the literature, and our own previous work evidence of the price-relevant short-run marginal cost of port services, in order to also represent ports belonging to category A. The relevant literature is mainly operations research in the form of queuing model applications to seaports, which clearly is applicable in the first place to category A ports. This is discussed in the next chapter. The most interesting of our own previous work is a study of the port of Uddevalla, which is one of the few category A ports in Sweden. In the following chapters 4-6 the old data of the port of Uddevalla as well as some other data are the basis for a fairly comprehensive discussion of short-run queuing and congestion costs in seaports, which finally is tied together by an illustrative, numerical example of peak-load pricing in the port of Uddevalla.

2.4 Transport system-externalities

The third term of the general expression (2.2) above for the price-relevant marginal cost , the system-external marginal cost, MC_{ext} is less important in seaports, given that the ships are at rest most of the time, and there are adequate facilities in the port to take care of oil spill, waste water, and other litter. However, in many cases ships are keeping engines going also while staying in the port for electricity generation, and other purposes, which makes MC_{ext} well above zero. For example, RoRo ships in the port of Norrköping normally use 25-50 percent of the full effect of the engines, and tankers run on full engine effect also in the port. On their way in and out of the port, all the ships certainly emit harmful fumes, and cause other externalities as well. This is obviously the case also for the trucks coming and going at the landside of the port. The latter externalities should, in principle, be reflected in adequate road pricing, which is outside the scope of this analysis.

However, in practice there is neither proper road pricing, nor adequate fuel charges on the bunker oil used for international sea transport, because the ship can avoid the charges by bunkering in “tax havens”. Therefore an interesting possibility is to make use of port charges as surrogate urban road user charges on trucks, and surrogate emission charges on ships, in view of the fact that the itinerary of the ships calling at the port are well known. The potential of this kind of second-best pricing is a topic of a coming project, and will not be further discussed in the present study.

3. SHORT-RUN MARGINAL COST ANALYSIS: THEORY OF QUEUING AND CONGESTION COSTS

Short-run marginal cost pricing proposals for seaports used to rely on queuing theory. In the following chapter, the application of queuing theory to a port facility consisting of a number of exchangeable common user berths is outlined. That is a picture of a traditional port of given location, but it is a less relevant description of the modern seaport on the move away from its original site in the middle of the central city.

In the short run port infrastructure is given. The wear and tear from use of this infrastructure is almost negligible, so the short-run marginal cost analysis can focus on the queuing and congestion costs, besides the direct cargo handling, or stevedoring costs.

3.1 Queuing model application

The common user seaport supplies services to ships, which arrive to the port largely at random, and which have differing requirements of port resources. Therefore, the short-term demand for port services varies – one week all resources may be occupied and ships be waiting in the roads, the next week there may be no ships at all in the port. The service time of ships is also highly variable.

If it can be assumed that the pattern of arrivals and service times adhere to some well-known probability distribution, the application of queuing theory to problems of port operations can be useful.

We begin with the simplest model, a so-called single-channel facility. We then present the more complicated model of a multichannel facility. This model is particularly interesting in pinpointing the importance of economies of scale in port operations. It will be demonstrated that if the total throughput increases, the number of berths can be expanded at a considerably lower rate without increasing the expected queuing time per customer.

3.1.1 Queuing time at a single-berth facility

The expected queuing time of customers arriving at random at a facility consisting of just one service station – a bank teller, for example, or a berth – rises quite sharply even at modest occupancy rates, If, in addition, customer requirements differ significantly, or if individual service times vary substantially, this tendency will be reinforced.

Using elementary queuing theory, this can be represented quite well. Let us make the following assumptions regarding the formation of the queues:

1. Customers (ships) arrive at random, with the distribution of arrivals described by the Poisson probability distribution.
2. Similarly, the service time is a random variable fitting the negative exponential probability distribution.
3. There is no upper limit to the queue length. Customers are patient and there is no restriction on waiting room space.

Under these conditions, the expected (mean) queuing time in a statistical equilibrium can be written

$$q=\frac{As^{2}}{1-As}=s\frac{\phi}{1-\phi} \qquad (3.1)$$

where

q = expected queuing time per ship in days;

A = expected number of ship arrivals;

s = expected service time per ship in days capacity;

ϕ = expected occupancy rate.

The mean queuing time of ships stars to rise already at rather low levels of capacity utilization, and will rise more and more sharply as the level of full capacity is approached.

Table 3.1 *Average and marginal queuing times relative to the service time at different occupancy rates*

Expected Occupancy rate: ϕ	The *average* queuing time as a proportion of the service time: $q/s=\frac{\phi}{1-\phi}$	The *marginal* queuing time as a proportion of the service time: $\frac{\partial(Aq)}{\partial A}/s=\frac{\phi(2-\phi)}{(1-\phi)^2}$
0.1	0.1111	0.2346
0.2	0.2500	0.5625
0.3	0.4286	1.0408
0.4	0.6666	1.7778
0.5	1.0000	3.0000
0.6	1.5000	5.2500
0.7	2.3333	8.5944
0.8	4.0000	24.0000
0.9	9.0000	99.0000
1.0	∞	∞

The rise in the mean queuing time is indicated in the middle column of table 5.1, which gives the ratio of q to s for different values of ϕ.

Marginal queuing time – the additional total queuing time that results from another ship arrival – rises even faster. The marginal queuing time is obtained by taking the partial derivative of the total queuing time Aq with respect to A:

$$\frac{\partial(Aq)}{\partial A}=\frac{As^2(2-As)}{(1-s)^2}=s\frac{\phi\ (2-\phi)}{(1-\phi)^2} \tag{3.2}$$

As can be seen from the right-hand column of table 3.1, the ratio of marginal queuing time to s increases much faster. (Another characteristic of this as well as of many other, more complicated queuing models is that, given the occupancy rate, q is proportional to s. A doubling of the mean service time, for example, will also double the mean queuing time.)

The importance of service time variability. The root cause of the queuing that occurs is the variability of A and s. How would different assumptions about the service time distribution affect the mean queuing time? The so-called Pollaczek-Khintchine formula provides a general answer to this question. For any arbitrary distribution of the service time s, the steady state mean queuing time q can be expressed as a function of the mean and the variance of the service time and the arrival rate:

$$\mathrm{q} = \frac{A\left[\mathrm{s}^2 + \mathrm{var}\,(\mathrm{s})\right]}{2\,(1 - A\mathrm{s})} \tag{3.3}$$

Inserting ϕ, which equals *As*, and denoting the relative variance var(s)/s by V(s), equation (3.3) becomes

$$\mathrm{q} = \frac{\phi\,(s + V(s))}{2\,(1 - \phi)} \tag{3.4}$$

Given the occupancy rate, the mean queuing time is proportional to the sum of the service time and its relative variance. To reduce queuing time, it is as important to reduce the variance of service time as it is to reduce the mean service time itself.

If *s* is distributed negative exponentially, its variance equals s^2, and by inserting this result (3.1) above is obtained:

If the variance of the service time is very small, the case of constant service time, $\bar{\mathrm{s}}$ becomes applicable. Setting $V(s) = 0$ in the general formula gives

$$\mathrm{q} = \frac{\bar{s}}{2} \cdot \frac{\phi}{1 - \phi}. \tag{3.5}$$

Eliminating the variability of service time will apparently reduce the mean queuing time by half.

3.1.2 Queuing time at a multiberth facility: economies of scale in port operations

In situations when an arriving ship may use one of several berths, it is appropriate to apply the multichannel variants of queuing models. These models are mathematically more involved. To simplify the exposition, the general formula in the multichannel case is derived in two stages. When *p* is introduced as a symbol for the probability that a ship arrival will find all berths occupied, the mean queuing time can be written

$$\mathrm{q} = \frac{\mathrm{s}}{\mathrm{n}(1\text{-}\phi)} \cdot p, \tag{3.6}$$

where n is the number of service stations, s is the mean service time of each station, and ϕ, the mean occupancy rate, now equals As/n. The term $s/n(1 - \phi)$ represents the expected queuing time of those customers who actually meet with a delay, while p is the probability that a delay will occur.

The total effect on queuing time of adding berths comes from both these terms. The first term is inversely proportional to n. If ϕ is held constant, that is, the number of berths increases in proportion to demand, the *total* queuing time is equal to a constant times p. The relation between p and n strengthens in fact the advantage of a multiberth port:

$$p = \frac{(n\phi)^n}{n!(1-\phi)} \frac{1}{\sum_{i=0}^{n-1} \frac{(n\phi)^i}{i!} + \frac{(n\phi)^n}{n!(1-\phi)}}. \tag{3.7}$$

Given the occupancy rate, p falls slightly as n is increased. Hence, the combined effect of the two factors in (3.6) means that there are important economies of scale inherent in multichannel service facilities. These economies of numbers are truly remarkable, as total queuing time will actually decrease when demand and capacity are expanded at the same rate.

3.1.3 Toward a multistage, multichannel model

A single-stage queuing model is often inadequate as a means of representing a seaport. This has been demonstrated very clearly during the oil boom in ports in the Persian Gulf and in Africa, where a common experience was that transit sheds were crammed while the quay occupancy did not always reach the same excessive level. Due to the relatively fast productivity increase in the loading/unloading operation proper, the transit storage stage in the throughput process became the bottleneck. The transfer of goods between sea and land transport is a multistage process whenever the indirect route is chosen, that is, when transit storage comes in between. The “chain of links” metaphor is used in two different contexts. One is that of detailed operation analysis in which every single action in the handling of cargo is defined as a link in the chain. A substantial number of links can be identified, particularly as far as the handling of general cargo is concerned.

Another context in which the chain-of-links concept is useful is in connection with the long-run problem of optimizing port capacity, that is, choosing the optimal number of service stations. For such investment problems a very fine division into links of the throughput chain would become unmanageable, as the number of factors to be taken into account becomes very large when the entire design of the port is variable. Where the indirect route of cargo predominates, it may seem appropriate to regard the throughput process as consisting of just two main stages with a “waiting room” in between. The transit storage space is a waiting-room between the two throughput stages of loading/unloading ships and loading/unloading land transport vehicles. The waiting room acts as a buffer between these two stages, making them independent, and thus improving the efficiency of both. According to this view, the multichannel queuing model is applicable separately to the loading/unloading of ships and to the loading/unloading of land transport vehicles. This approach is sound, provided the holding capacity of the waiting room is practically unlimited. However, it is well known that in the transit storage facilities of many ports this is not so.

The cost of the transit storage capacity depends to a great extent on land values, In older ports situated in or near the core of a city, storage space is a chronic bottleneck. And as soon as the capacity of the waiting room is no longer unlimited, the performances of the two service stages will be connected. When the waiting room happens to be completely filled by customers because of some holdup in production in the second stage, the preceding service stage cannot pass on customers who have been served. The customers have to remain in the first stage, blocking the way for the customers behind them. It should be mentioned here that one branch of general queuing theory deals with two-stage processes with waiting room of limited holding capacity between them.

However, this is not the whole story of transit storage in a port. An import consignment that arrives in transit storage cannot move to the next stage as soon as a service station (freight delivery station) is vacant, unless the right truck is there to pick it up. This is obviously a decisive condition. The time spent by a consignment in transit storage is not determined primarily by the capacity of the following service stage, but by the time that elapses before it is collected by the importer or, in the case of export cargo, by the interval between its receipt at the port landside and the arrival of the relevant ship.

Thus, in one respect, the storage facility does not correspond exactly to a waiting room in the limited sense relevant to queuing models. This can be allowed for by defining transit storage as an independent service stage as well as a waiting room, so that the time spent by the cargo in transit storage can then be regarded as service time rather than pure waiting time. This would require a three-stage multi-channel queuing model.

We end the discussion at this point. For a further discussion and development of a multistage, multichannel model of the throughput process, the reader is referred to Jansson and Shneerson (1982), and Jansson (1984). For the present study, the conclusion is that the modelling work so far has produced some new insight into the complexities arising on account of the large stochastic element in seaport supply and demand, but for practical purposes rough and ready statistical cost analysis has to be resorted to in order to bring us further. As a surrogate for a penetrating, multistage queuing model, we will now introduce the notion of "congestion costs" as distinct from queuing costs, in order to make the short-run price-relevant cost picture more complete.

3.2 Congestion costs in addition to queuing costs?

A different approach, without a firm basis in operations research, is simply to rely on regression analysis, provided, of course, that relevant data for cost and output are to be found. It is also true that certain basic assumptions of the queuing models can be questioned.

A particularly critical assumption of standard queuing models is that the service time s of ships is constant, i.e. independent of the occupancy rate, ϕ. The rationale of this very simplifying assumption is that a port seemingly belongs to what can be called the pure departmentalized case. A departmentalized facility consists of several service stations for which the activities in one station has no influence on the activities in the other stations. The service one gets is the same irrespective of the occupancy rate of the other services stations. Is this true about port services? It can be shown that even a modest positive relationship between s and ϕ will result in substantially longer queuing times.

A related issue of importance for the shape of the total short-run cost is whether the cargo handling cost per throughput ton – the stevedoring cost – is affected by the rate of port capacity utilization. This issue has no direct bearing on the question of the validity of the standard queuing model, but is important in its own right. Since it can be assumed that there is some substitution between service time of ships and stevedoring labour, the empirical analysis had better deal with these two costs at the same time.

A quasi-departmentalized case exists in a number of different branches of the service sector. It is characterized by the fact that the total manning and/or equipment of the facility will be more and more *diluted* per user as the occupancy rate increases. A familiar case is the restaurant, where each waiter serves a certain number of tables. The more customers there are in the restaurant, the less will be the share of the waiter's total time per customer. When the restaurant is full or nearly full, a meal will, on average, take a longer time than the customers would prefer. Actual queuing, however, will not occur until all tables (chairs) are occupied. In the transport sector the quasi-departmentalized case may be represented by seaports. If there is a more or less fixed pool of stevedores and cranes (which can be moved between berths), it seems to follow that the more berths that are occupied, the smaller will be the maximum possible input of cargo handling labour and equipment per ship, and the longer the mean service time. One does not speak about queuing until all berths are occupied and ships have to wait in the roads.

On the other hand, it can be shown that dilution of resources is inoptimal under certain circumstances. Take crane allocation as an example. Suppose that a particular quay line is equipped by three quay cranes and that all three cranes may be used simultaneously at a ship. Suppose further that the marginal productivity of a crane is constant, independent of whether one, two, or three cranes are working at the same ship. Under these conditions it is optimal – in the sense that the sum of the total queuing time and service time will be minimized – to use three cranes on a ship irrespective of whether other ships are waiting to be served. Dilution of cranes – one crane for each of three ships, two cranes for one ship, one crane for another ship, and so forth – is an inferior policy of crane allocation.

Another fact that speaks against the resource dilution hypothesis is that *overtime work* should be more frequent when other ships are waiting in the roads. This may tend to reduce the service time of ships in periods of high berth occupancy. This possible reduction is obviously not costless. Additions of 50-100% or more of the basic rates have to be paid for stevedoring work in the evening and on holidays. If the overtime work is ordered by the shipowners themselves, overtime payment should not be viewed as an extra cost due to a high level of total demand, because it can be assumed that the saving in laytime for the ships on which overtime is worked is a sufficient compensation. On the other hand, to the extent that the Overtime work is ordered by the port authority or stevedoring company with a view to reducing the queuing time of other ships, it represents a partial extra cost that should be attributed to the high level of total demand.

There is a possible third negative effect of a high traffic intensity, which is due to the less than fully “departmentalized” character of a port section. The occupancy rate of the storage facilities can be expected to have some influence on the internal transport and handling costs. In fully occupied storage facilities all operations become much more laborious than in less crammed facilities. Overfull facilities, meaning that the storage area is overflown to the extent that the traffic area is reduced, can be a very unfavourable state; it may even reduce the throughput capacity by making all movements slower and the average distance of internal transports longer.

3.2.1 Distinguishing congestion costs and queuing costs

A name for the possible negative effects of a high rate of capacity utilization mentioned above is *congestion costs*; these are to be distinguished from the queuing costs. The latter do not appear until actual demand exceeds capacity. Congestion costs exist if the other short-run costs of port operations, per unit of throughput, are an increasing function of the actual capacity utilization. When actual demand exceeds capacity, extreme congestion costs arise, which we call *queuing costs*. When a port is said to be congested, it is commonly meant that ships are queuing, waiting for a vacant a berth.

We prefer to restrict the term congestion costs to effects that show up as increases in the cargo handling costs and/or the service time of ships (which, in turn, will influence the queuing time of ships).

It is an important empirical task to examine (i) whether congestion costs, defined as above, exist at all in ports, and, if they exist, (ii) how they are related to the level of throughput, or port capacity utilization. It is much more difficult to identify and measure possible congestion costs than queuing costs. The latter are clearly visible in the form of idle ships waiting in the roads (or trucks, etc., queuing at the land side of a port). The congestion costs, on the other hand, take the form of increased service time and/or higher stevedoring costs for a given cargo. If empirical observations showed that the short-run total variable costs (excluding queuing time costs) increase *progressively* with rises in port throughput, this would be an indication that congestion costs exist.

4. EMPIRICAL EVIDENCE OF QUEUING COST FUNCTIONS OF SEAPORTS

In the 1970s a great deal of empirical research was carried out jointly at the Economic Research Institute of the Stockholm School of Economics, and the Israeli Shipping Research Institute and University of Haifa with a view to answering the question: can the results of the standard queuing models be trusted so far as seaport operations are concerned? The parts of this work which are directly relevant for the present purpose will be summarized below. Since then no similar work has, to our knowledge, been carried out elsewhere, and it has also proved to be very difficult to do the same thing again in Sweden because of data unavailability, which partly depends on a general cut in the public production of statistics, and partly on the transformation of many ports from public administrations to joint stock companies.

First we take up a main, general problem of statistical curve fitting when it comes to queue formation processes.

4.1 The observation period – a month, a quarter, or a whole year?

The result of the queuing model discussed in the previous chapter assumes that the system is in statistical equilibrium. The applicable time period is assumed to be sufficiently long to allow for the attainment of statistical equilibrium. Averaging of subperiod outcomes is inherent in the concept of a queuing time function. When regression analysis is applied to real-life observations, the period of observation selected should be sufficiently long to allow for the attainment of statistical equilibrium.

This point can be illustrated by a schematic example. Suppose that the service time at a particular service-providing facility is one day and that customers always arrive in the morning. Those who have received service leave the facility the same evening. Those who find all service stations occupied have to wait until the following morning or even longer. Now assume that we shall estimate the relationship between total queuing time Z and the occupancy rate of the facility ϕ. Queuing time will be suffered by customers only on days when 100% occupancy is reached. This means that if the observation period is just one day, zero queuing will be found for all occupancy rates less than 100%. At the other extreme, if the observation period is very long – one year or more – the scatter is likely to form a smooth rising curve. In this case the queuing time and the occupancy rate are, of course, *average* values taken over a whole year. The $Z(\phi)$ function, based on these average annual observations, are comparable to the curves derived from the queuing theory formulas.

For practical reasons, very long observation periods are infeasible because the number of observations will be too few, and various external factors of influence, which are difficult to control for, will not remain the same. In practical work, the period of observations has to be shortened. If, for example, weekly observations are relied upon, a rather dispersed scatter is to be expected. A low $\overline{R}^2$ for the fitted curve is not necessarily an indication that a systematic influence of other explanatory variables exists (other than the average occupancy rate). It may simply mean that the observation period is too short for a statistical equilibrium to work out.

This problem will be further discussed in more concrete terms in the following presentation of the results of regression analysis.

4.2 The queuing time function of the port of Uddevalla, Sweden

The best data we had access to were from the port of Uddevalla, Sweden. It used to be a comparatively busy port with an occasionally high rate of capacity utilization. Monthly data on throughput, berth occupancy rate, and the queuing time of ships could be obtained. Observations were collected for 42 months from January 1973 to June 1976 for the following variables:
the mean occupancy rate of all the berths ϕ,
total queuing days of ships Z,
total throughput of 20 groups of commodities $Q_1, \ldots Q_{20}$.

The 20 different throughput quantities were aggregated into a single throughput quantity Q for each month. This was done by taking a weighted average of the 20 commodities. The most suitable weights available were stevedoring charges, which are by and large related to the relative ease of handling the goods.

A crucial question is whether a month is a sufficiently long period for the measurement of Z. On the other hand, it is important that the figures of Q or ϕ are sufficiently wide-ranging to make curve fitting possible. By choosing a much longer observation period, the random fluctuations in demand will tend to level out in the averages, making the seasonal variations the only source of variations in the data. The monthly variations in the traffic of Uddevalla in the years 1973-1976 are greater than what would be expected if strictly random fluctuations are in effect. Given that the average number of ship arrivals was 83/month during the total period of observations, and assuming that the arrivals of ships to Uddevalla obey the Poisson distribution, the expected coefficient of variation of monthly arrivals is $(1/\sqrt{83})\ 100 = 11\%$, provided that no seasonality exists. The actual coefficient of variation, however, is 32% so far as arrivals of ships are concerned, and 26% for the aggregate throughput of goods. This is explained by seasonality. An inspection of the monthly figures reveals a cyclical pattern with March as the peak month, and the summer months constituting a trough.

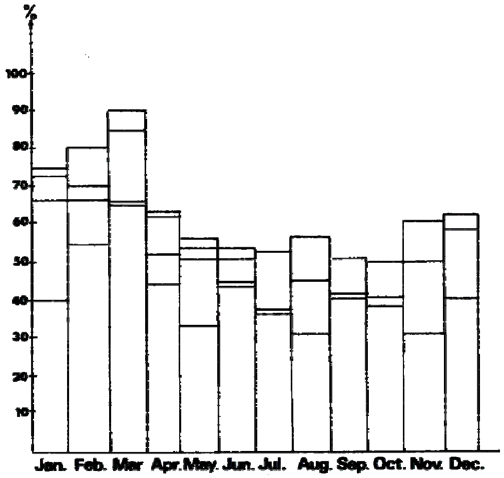


Figure 4.1 **Berth occupancy rates in the port of Uddevalla 1973-76**

Table 4.1 *Average values for the aggregate throughput and the berth occupancy rate for each month*

Month	Q (tons)	ϕ (%)
Jan.	129,751	64
Feb.	134,939	68
Mar.	160,537	76
Apr.	119,787	56
May	104,494	48
Jun.	96,925	50
Jul.	96,784	41
Aug.	90,257	44
Sep.	110,221	43
Oct.	109,923	42
Nov.	94,798	46
Dec.	104,406	52

Table 4.2 *Total monthly throughput of timber-equivalent tons of consecutive years*

1/1/1973-31/12/1973	127,671
1/1/1974-31/12/1974	119,292
1/1/1975-31/12/1975	101,569
1/1/1976-30/6/1976	103,750

Another source of systematic differences in demand between different months included in the sample is that, looking at annual figures, the total traffic has been somewhat on the decline in the period 1973-1976 (table 4.2). Anyway, regression analysis was tried for explaining the monthly variations in queuing time.

In figures 4.2 and 4.3 scatter diagrams of observations of Z and Q, and of Z and ϕ, respectively, are shown. As seen, no queuing time at all will occur below a level of aggregate throughput of 80,000 tons. This corresponds to a berth occupancy rate of almost 40%. As the traffic is increasing beyond this level, total queuing time will rise quite sharply.

The choice of form of the regression function to fit the observed data on queuing times is not so straightforward. The relation between queuing time and the occupancy rate of a multichannel facility is rather complex, and on theoretical grounds a linear relation is clearly out of the question. Instead, exponential relations both between Z and Q and Z and ϕ have been assumed. Such constant elastic curves, however, are not completely satisfactory in view of the fact that a capacity limit exists. The occupancy rate ϕ can clearly not exceed unity, and as the capacity limit is approached, the elasticity of Z with respect to ϕ or Q is bound to increase. As an alternative, $1 - \phi$ has been tried as an explanatory variable. In this case we expect, of course, a negative regression coefficient (elasticity). From queuing theory we know that the inverse of this variable appears as a central argument in practically all queuing time formulas.

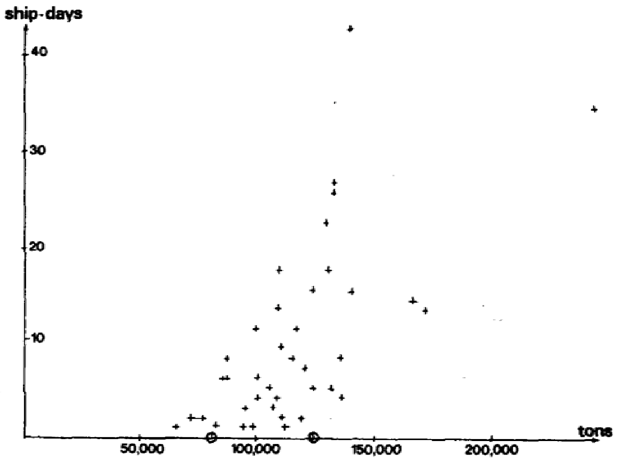


Figure 4.2 **Total queuing time versus port throughput**

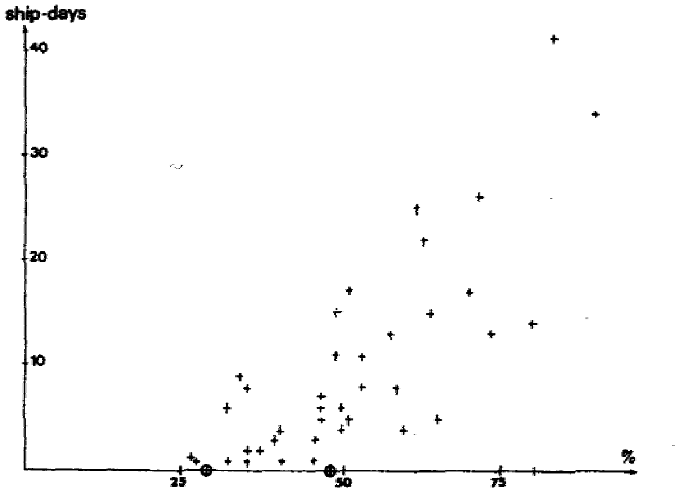


Figure 4.3 **Total queuing time versus berth occupancy**

Table 4.3 *Result of regression analysis of total queuing time Z on port throughput and berth occupancy rate, respectively*

Number of observations	Independent variable	Constant	Regression coefficient (elasticity)	t-value	$\bar{R}^2$
40	Q	$3.55 \cdot 10^{-14}$	2.82	5.34	0.31
39[a]	Q	$3.03 \cdot 10^{-15}$	3.03	4.90	0.29
40	ϕ	42.95	2.98	7.48	0.69
40	$1 - \phi$	1.31	-1.83	-6.45	0.52

[a] The odd observation Q = 240.000 tons is excluded

As follows from table 6.3 the best variant for explaining the variation of Z is a simple exponential relation between Z and ϕ. In round numbers it turn out to be: $Z = 43\phi^3$.

In comparison to this function, the alternative of using $1 - \phi$ as independent variable exhibits a substantially lower degree of explanation.

Replacing ϕ by Q as the independent variable yields very similar regression coefficients (close to 3 in each case). It is to be expected (and it was found) that Q and ϕ are very nearly proportional. The value of $\overline{R}^2$, however, becomes substantially lower when Q is used as explanatory variable. This can be interpreted to mean that the occupancy rate ϕ is the main direct determinant of the total queuing time Z. The throughput Q influences Z only to the extent that Q and ϕ are correlated. As Q does not explain all variations in ϕ, it performs correspondingly worse in explaining Z.

4.3 Queuing time functions of the ports of Haifa and Ashdod, Israel

We can also report results of regression analysis of queuing time functions from the ports of Haifa and Ashdod in Israel. Both ports have worked at levels of high occupancy rates (as well as low occupancy rates) during various periods over the years, which makes curve-fitting possible. The period of time for which data were collected is longer than at the port of Uddevalla – 9 years for both ports (from 1966 to 1974). This makes it possible to select a longer (than a month) period of observation. Quarterly data were used, but to elucidate the importance of the observation period, the results obtained were compared to the results that followed from using monthly data.

Observations were collected for queuing time, occupancy rate, and throughput. Occupancy rates were calculated by dividing the product of the average length of ships and the number of ship-days of berthing ships by the product of the length of the berths and the working days during a period of observation.

Table 4.4 *Result of regression analysis of total queuing time on berth occupancy in the ports of Haifa and Ashdod – quarterly observations*

Type of cargo	Name of port	Number of observations	Independent variable	Constant	Regression coefficient (elasticity)	t-value	$\bar{R}^2$
General Cargo	Haifa	36	ϕ	718.3	2.949	2.20	0.48
General Cargo	Ashdod	36	ϕ	1071.0	2.99	3.02	0.29
Grain	Haifa	36	ϕ	781.9	3.172	3.79	0.70

Two separate sets of data were collected and two separate regressions were run – for general cargo and for bulk. Table 4.4 summarizes the regression results. The best fit was obtained by a simple exponential equation. The proportion of the variations explained by the regression equation ($\bar{R}^2$) is rather low for the port of Ashdod. This can be attributed to irregularities due to the two wars (1967 and 1973) and to strikes. The port of Ashdod was notorious for strikes during this period.

The similarity of the regression coefficients for the two ports as well as for the port of Uddevalla is striking. However, the value of the constant is more important for the queuing time in the relevant range than the elasticity, and the constant represents i.a. the average service times, which is much longer in the Israeli ports due to bigger shiploads by bigger ships.

Concerning general cargo through Haifa, monthly observations were also used. The results are summarized in table 4.5. Shortening the period of observation – from quarterly to monthly figures – gave, as seen, much worse results. This proves the point made earlier of how important it is to select a period sufficiently long to allow the attainment of statistical equilibrium.

Table 4.5 *Result of regression analysis queuing time on berth occupancy in the ports of Haifa – monthly observations*

Type of cargo	Number of observations	Independent variable	Constant	Regression coefficient (elasticity)	t-value	$\bar{R}^2$
General cargo	108	ϕ	79.4	1.08	0.59	0.13

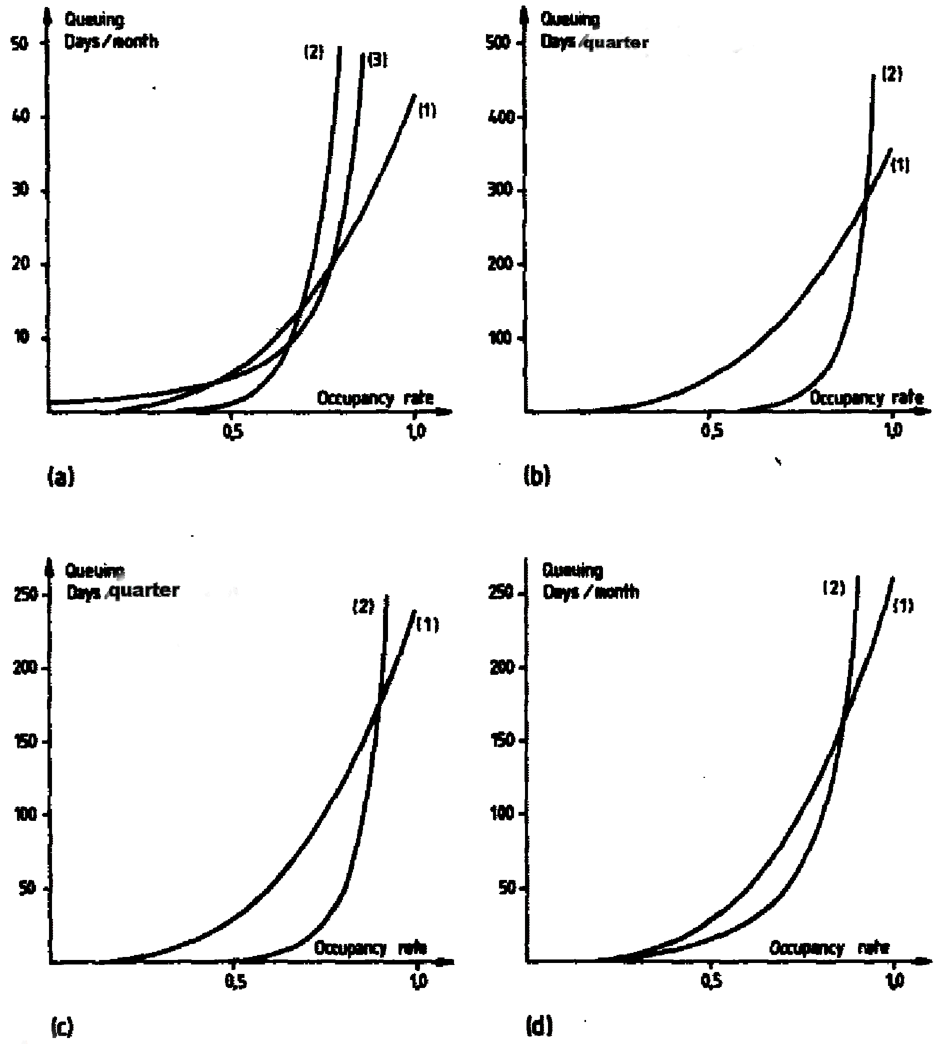


Figure 4.4 **Total queuing time Z as a function of the occupancy rate ϕ in theory and practice: (a) Uddevalla, 10 berths; (b) Ashdod, 13 berths; (c) Haifa, 10 berths (general cargo); (d) Haifa, 1 berth (grain). Curves: (1) empirically estimated relation between Z and ϕ; (2) relation between Z and ϕ derived from queuing theory; (3) empirically estimated relation between Z and 1 - ϕ.**

4.4 Comparison of empirical measurement and queuing model results

We shall now consider how the previous findings agree with the predictions of standard queuing theory. In each of figures 4.4a – d two curves representing (1) the empirically estimated function *Z* (ϕ) and (2) the corresponding relation obtained from queuing theory are drawn. As can be seen, in all four cases the curve obtained from the multichannel queuing model is well below the empirically estimated curve to begin with. Curve 1 rises at an earlier stage but does not accelerate like curve 2. As the capacity limit is approached, however, the two curves intersect, and curve 2 becomes the highest. The discrepancy between the two curves is quickly getting larger for very high occupancy rates.

An explanation of this latter disagreement between the two curves is close to hand: customers are by no means as patient as assumed in the queuing models. If a queue of ships is building up, waiting ships sooner or later will go to other ports. We were told in the port of Uddevalla that it also happens from time to time that ships that plan to call at the port of Uddevalla, but learn by cable that ships are waiting in the roads, choose to go to a neighbouring port, for example, Gottenburg. Such deviations are not costless, although presumably less costly than the alternative of incurring actual queuing time. It was not possible to estimate the total magnitude of those latent queuing costs.

It should be observed, however, that rather few observations are made in the very high range for the berth occupancy, and, given the form of the regression equation ($Z = a\phi^b$) it is impossible to obtain the sharp upward turn characteristic of the queuing time function of the multichannel model, even if the observations in the 0.8 – 1 range of ϕ would speak for that. We therefore “forced” the curve fitted to the data to bend sharply as the capacity limit is approached by also trying $1 - \phi$ as explanatory variable. The proportion of variation explained was considerably less in this case, but, as can be seen, curve 3 in figure 4.4a is somewhat more in line with the queuing model.

The opposite discrepancy in the range below the very high occupancy rates is more difficult to explain. It cannot be just a coincidence in view of the fact that the same picture appears in all cases. A conceivable explanation is that the service time of ships does not stay constant as the port capacity utilization rises, because bottlenecks appears in other stages of the throughput process, beyond the loading and unloading of ships, which can give rise to queuing, also when vacant berths exist. This possibility will be considered in the next section.

5. EMPIRICAL EVIDENCE OF CONGESTION COSTS IN PORTS

The queuing time of ships waiting in the roads is not the only, and perhaps not even the most important negative effect of a very high rate of port capacity utilization. The question is if ports are regularly operating at a sufficiently high rate of capacity utilization for more appreciable congestion costs (besides ships' queuing time) to appear? In many Swedish ports this question have mostly been answered in the negative by port managers. If total variable costs increase proportionally to throughput, the inference would be that congestion costs do not exist. When it comes to empirical estimation, the problems are that (1) ships and cargoes are very different with respect to cargo handling efficiency and (2) the cargo handling performance is strongly affected by numerous external factors independent of the level of traffic. Unfortunately, it does not seem to be common practice among ports to keep records that give a clear picture of congestion costs. The results of two case studies from the port of Stockholm, and the port of Uddevalla are the only empirical evidence to hand. The Stockholm study is from 1974. Data were obtained on ship laytimes along the quay (as distinct from queuing time spent waiting for a vacant quay-berth) stevedoring labour inputs into the handling of different cargoes, and weekly rates of utilization of quays and cranes. The problem was that although a fairly wide spread of capacity utilization rates appeared, the variations were restricted to the lower half of the possible range. The traffic of the port of Stockholm had been on the decline, and downward adjustments of capacity naturally lag behind. In retrospect the conclusion may seem inevitable: No congestion costs were apparent in the low range for the rate of capacity utilization.

The port of Uddevalla was selected for the second case study because it had been operating at a comparatively high rate of capacity utilization for some time. If congestion costs exist in any segment of the range of capacity utilization, this should be revealed in the port of Uddevalla. We obtained monthly data for the period from January 1973 to June 1976 concerning the mean berth occupancy rate of the whole port, the total throughput of 20 groups of commodities, which could be aggregated to timber-equivalent throughput tons, as has been mentioned in the preceding section, and in addition we got the stevedoring costs, and the costs of ships' laytime at the berths, which we call "service time costs". In the port of Uddevalla, total stevedoring costs amounted to about $1.6 million/year, and the annual total service time costs could be estimated to be about $3 million.

On the basis of these data, the following relations were investigated:

service time vs. port throughput
stevedoring cost vs. port throughput
the sum of service time costs and stevedoring costs vs. port throughput.

Two alternative forms of the regression equation were used, the linear form $y = a + bx$ and the exponential form $\ln y = \alpha + \beta \ln x$.

As to the interpretation of the linear regressions, we shall take the existence of an appreciable positive intercept to be a rough indication that the “true” relation is degressive, and a negative intercept as evidence of a progressively increasing function. In the exponential case the value of β gives the elasticity of the function. An elasticity greater than unity would indicate that congestion costs exist in one form or another.

A peculiarity of the data that unfortunately made the interpretation of the results somewhat difficult should be mentioned. All observations except one (41 out of 42) fall within a range of aggregate throughput between 65,000 and 170,000 tons. In this range a reasonably even spread of observation’s exists, as is clear from the scatter diagrams of figures 7.1 and 7.2. The only observation outside this range is at a level of throughput of 240,000 tons. This represents March 1973, which was a very exceptional month according to a memorandum about port capacity estimation by the manager of the port of Uddevalla. Everything seemed to go exceptionally smoothly this particular month. The “practical capacity” is estimated to be som 50-70% of this record figure.

Therefore one can have little confidence in the predictive value of the regression equations in the 170,000-240,000 range. It is arguable that the odd far-out observation should be disregarded and the regression analysis restricted to the 65,000-170,000 throughput range. We shall give results obtained by using all 42 observations, as well as results obtained when the odd month is excluded.

5.1 Service time and throughput

The purpose of this analysis is to determine whether there is any tendency of total service time X to increase progressively rather than proportionally as throughput Q is increasing. The data did not give a direct measure of X. Instead, the berth occupancy rate ϕ is used as a proxy variable, since the total service time is necessarily proportional to the berth occupancy rate.

Figure 5.1 shows the mean occupancy rate plotted against total throughput on a monthly basis. The pattern of the scatter hardly indicates anything but a linear relation. The results of the regression analysis are shown in the table below.

Table 5.1 *Result of regression analysis of berth occupancy (as a proxy for total service time) on port throughput*

Form of regression equation	Number of observations	Constant	Throughput coefficient	t-value	$\overline{R}^2$
Exponential	41	0.000243	1.06	10.24	0.72
Exponential	42	0.000655	0.97	10.70	0.71
Linear	41	-3.08	0.0005	9.99	0.71
Linear	42	6.79	0.0004	10.01	0.71

The exponential form of the regression equation shows an elasticity very close to unity, and the linear regression equation bypasses the origin by an almost negligible intercept. This implies that a given increase in throughput will cause a proportional increase in the berth occupancy rate. Hence, there is no evidence of congesiton costs in terms of longer service times. In fact, the impression of the case of 42 observations is that the input of ship time per throughput ton may even *fall* slightly with increases in total throughput. This possibility is not as farfetched as it may first appear. The proportion of overtime work is an important factor in this context. If the port authority and/or stevedoring company have some influence on when ships should be worked on overtime, it can be expected that when other ships are waiting in the roads, the proportion of overtime work will increase. This may make for a reduction of the service time. If the proportion of overtime work goes up when demand is high, this should be revealed by the relationship between the total stevedoring costs and throughput. In other words, if the congestion costs do not show up as longer service times, they may appear in the form of increasing stevedoring costs.

5.2 Stevedoring costs and throughput

In the port of Uddevalla practically all stevedoring was performed on a piecework basis. This speaks for strict proportionality between wage costs and throughput. However, the permanently employed stevedoring labour was guaranteed a certain minimum wage even if no work was available. The total hours of work by the permanently employed labour force were less than a third of the total hours performed by casual labour, but the permanent labour force was used in the first place, as its number of idle hours should be kept to a minimum.

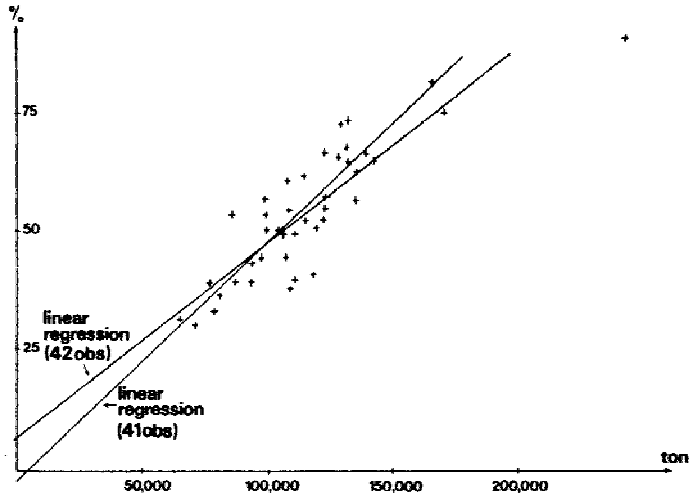


Figure 5.1 **Berth occupancy versus port throughput**

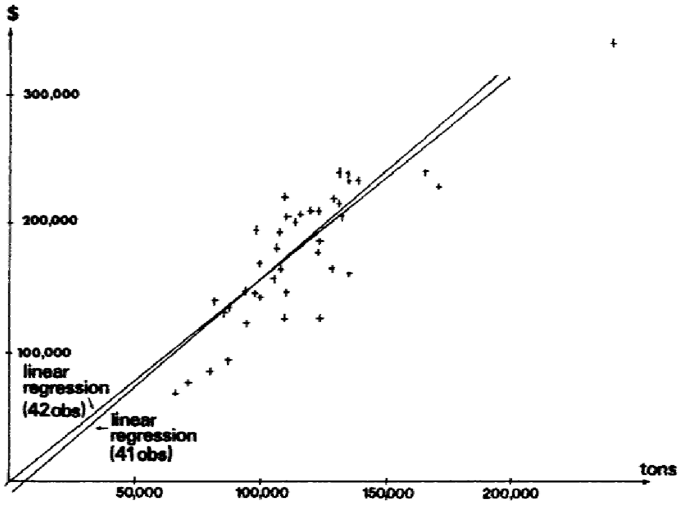


Figure 5.2 **Total stevedoring cost versus throughput**

Thus, we expect that in an initial range the total stevedoring wage costs are constant at the level of the guaranteed minimum total wages of the permanently employed stevedores. This range is not represented, however, in the data. Our data on stevedoring costs included the extra payments for overtime. (It was not possible to distinguish the latter from the ordinary wage costs). Given our hypothesis that the proportion of overtime work rises with increases in the throughput, we expect that the total stevedoring wage costs will eventually increase progressively as total throughput goes up, or, in other words, that the elasticity of the total stevedoring costs with respect to total throughput is somewhat greater than unity in the upper range.

The available stevedoring labour costs were total nominal wages paid every month from January 1973 to June 1976. In view of the high wage inflation during this period, it was necessary to eliminate the effect of inflation in some way. Also, it cannot be assumed that the productivity of stevedoring labour remained constant during the period. As our purpose was to estimate the short-run relation between labour costs and throughput, it is desirable that both these factors should be eliminated. The best method seemed to be to inflate the cost figures of 1973, 1974, and 1975 by the percentage rises in the stevedoring charges up to 1976 rather than by the rises in the wage rates. This is based on the assumption that changes in stevedoring charges will reflect both productivity and nominal wage rate changes.

As is evident from the regression lines that have been inserted in the scatter diagram of figure 5.2, the linear regression indicates strict proportionality between the stevedoring wage cost and throughput. (It is perhaps somewhat misleading to extend the regression lines into the low-throughput range, where no observations exist. Zero throughput will not mean zero stevedoring labour costs in the short run on account of obligations to permanently employed labour). A somewhat different indication is obtained from the result of the exponential regression (see table 5.2): elasticities were greater than unity ($\sum = 1.23$ in the case of 41 observations, and $\sum = 1.16$ in the case of 42 observations). This implies that there were some congestion costs, presumably in the form of extra payment for increased overtime work when throughput was increased.

Table 5.2 *Result of regression analysis of the stevedoring costs on port throughput*

Form of regression equation	Number of observations	Constant	Throughput coefficient	t-value	$\overline{R}^2$
Exponential	41	0.101	1.23	8.47	0.60
Exponential	42	0.234	1.16	9.28	0.67
Linear	41	-13.369	1.658	9.66	0.69
Linear	42	1.183	1.522	7.92	0.61

a. Dependent variable = stevedoring costs; independent variable = port throughput

It is not possible to draw any firm conclusions from the result of the regression analysis although there are some indications that high throughput volumes may entail a small rise in the stevedoring cost per ton due to overtime payments.

5.3 The sum of service time and stevedoring costs versus throughput

The third step in our attempt to shed some light on the existence of congestion costs is to estimate the relationship between the total service time and stevedoring costs and throughput Q. In order to calculate these cost, it is first necessary to transform the data on berth occupancy into service time costs of ships. This is done by taking the product of the occupancy rate, the number of berths, and an average cost per ship-day in port. A figure for the latter cost has been obtained in this way. The average net registered tonnage (nrt) of ships calling at Uddevalla in the period of observation was 884 nrt. The average cost per day in port of a ship calling at the port of Uddevalla is assumed to be represented by an unweighted average of the costs of Swedish, British, and Norwegian all-purpose cargo ships of this size . In round figures this cost amounts to $1,600.

It is noteworthy that the coefficient of determination ($\overline{R}^2$) increases appreciably by regressing the sum of service time costs and stevedoring costs on Q rather than taking each of the two cost items separately. The results from the exponential regression are not conclusive, but in the main range of throughput, that is, between 65,000 and 170,000 tons, some progressivity is indicated since $\sum = 1.12$ without the outlier; if the odd month is included, the elasticity falls to 1.04.

Table 5.3 *Result of regression analysis of the sum of service time and stevedoring costs on port throughput*

Form of regression equation	Number of observations	Constant	Throughput coefficient	t-value	$\bar{R}^2$
Exponential	41	0.932	1.12	11.85	0.78
Exponential	42	0.565	1.04	12.46	0.77
Linear	41	-28.162	4.064	11.24	0.76
Linear	42	33.792	3.488	12.08	0.78

5.4 Conclusions and implications for queuing models

The empirical data do not lend support to the hypothesis that the service time of ships will increase with increases in the capacity utilization. The average service time of ship seems to remain virtually constant right up to the limit when all berths are occupied, presumably in part due to a higher proportion of overtime work in periods when demand is high. Such a cost-raising effect of a high level of demand is possible to discern, although it is certainly not very pronounced. However, as regards queuing models of ports, it is important to conclude that the departmentalized case seems to apply. At least we have found no evidence that contradicts the fundamental assumption of queuing models of seaports, that service time of ships is constant right up to the capacity limit. It is thus necessary to look for another explanation for the discrepancy found in the preceding chapter between actual queuing time in ports and queuing time predicted by the standard queuing model. Perhaps the explanation involves the single-stage character of the standard queuing model. From the theory of multistage, or serial queuing processes, we know that the queuing time increases greatly as the number of equicapacity stages is increased. For ports where the transfer of goods between sea and land transport takes the indirect route involving transit storage, single-stage queuing models are, perhaps, too much off the mark.

6. SHORT-RUN MARGINAL COST PRICING OF PORT SERVICES: AN ILLUSTRATIVE EXAMPLE

Optimal price P* for bringing cargo between ships and transit storage in the port should be equal to the price-relevant marginal cost, which in principle is a three-term expression:

$$P^* = MC_{prod} + Q\frac{dAC_{user}}{dQ} + MC_{ext} \qquad (6.1)$$

The producer marginal cost component mainly consists of cargo handling, or stevedoring costs. The user cost component is a twofold ship time cost, (i) queuing time for a vacant quay-berth, and (ii) laytime at the berth during the process of (un)loading. The throughput system-external costs are normally negligible, given that the ships are at rest in the port, and there are adequate facilities in the port to take care of oil spill, waste water, and other litter. Our numerical example will therefore entirely focus on the two first components of the price-relevant marginal cost.

6.1 Peak-load pricing by queuing charges

Being a user cost, the price-relevant queuing cost is equal to the product of total throughput Q, and the derivative of the average queuing cost, AC_q with respect to Q, or alternatively as the difference between MC_q and AC_q. This can also be written as the product of the applicable cost-elasticity (E) and AC_q.

Queuing model applications to seaports assume that the short-run fluctuations in demand are entirely random. In practice there are usually seasonal fluctuations as well, which produce a systematic pattern of distinct peaks and troughs of port activity. This means that the expected queuing cost may differ considerably over time. There is a case for differentiating the charges according to the seasonal changes in demand. In the short run, given the capacity, a more even pattern of ships arrivals will reduce total queuing costs. A decrease in peak traffic and a corresponding increase in off-peak traffic will reduce the mean queuing time. In the long run, a levelling of the time profile of demand will save capacity investment costs to an amount that should, at least, be equal to the short-run cost savings.

The optimal seasonal price differential is equal to the difference between the price-relevant marginal costs in peak and off-peak *after the introduction of peak load pricing*. In practice, it will be difficult to find the optimal differential. It should also be remembered that what turns out to be the correct differential one year may not be so the next year. These obvious difficulties should not be regarded as too discouraging. The optimal price differential can be found by a process of trial and error. The range of the optimal price differential is limited – it can range from zero to the present difference between the peak and off-peak price-relevant marginal costs. If, for example, demand is thought to be elastic, it may be wise to start conservatively by introducing a price differential close to the lower limit. If this turn out to be too little, the price differential can be gradually raised until the optimal level has been reached.

Table 6.1 *Derivation of the price-relevant queuing cost functions*

On the basis of Z (ϕ)	On the basis of Z (Q)
$Z = 43 \cdot \phi^3$	$Z = 3 \cdot 10^{-15} \cdot Q^3$
$X = \phi \cdot n \cdot t$	
$AC = \frac{v \cdot Z}{X} = \frac{1{,}600 \cdot 43 \cdot \phi^3}{\phi \cdot 10 \cdot 30} = 229 \cdot \phi^2$	$AC = \frac{v \cdot Z}{Q} = \frac{1{,}600 \cdot 3 \cdot 10^{-15} \cdot Q^3}{Q}$
	$= 4{,}8 \cdot 10^{-12} \cdot Q^2$
$E = \frac{\partial AC}{\partial \phi} \cdot \frac{\phi}{AC} = 2$	$E = \frac{\partial AC}{\partial Q} \cdot \frac{Q}{AC} = 2$
$P_x = E \cdot AC = 2 \cdot 229 \cdot \phi^2 = 458 \cdot \phi^2$	$P_Q = E \cdot AC = 2 \cdot 4{,}8 \cdot 10^{-12} \cdot Q^2$ $= 9{,}6 \cdot 10^{-12} \cdot Q^2$

Table 6.2 *Queuing charges at different rates of capacity utilization in the form of berth occupancy charges, and charges on goods, respectively*

Berth occupancy rate	Berth occupancy charges, \$ per service day, P_X	Goods charges, \$ per cargo ton, P_Q
0.3	41	0.06
0.4	73	0.10
0.5	115	0.15
0.6	165	0.21
0.7	225	0.28
0.8	294	0.37
0.9	371	0.48

We shall demonstrate the application of peak load pricing to the port of Uddevalla, Sweden, under the condition that only one charge is applicable in each particular season. The port is characterized by fairly significant variations in demand. The average values of aggregate throughput of each month of the year, and the corresponding berth occupancy rates are presented in table 4.1. Suppose that these monthly variations are due to seasonality – the variations in demand from one year to another for a particular month can be assumed to be mainly stochastic. Then a price differentiation by month is justified. The queuing time functions estimated in chapter 6, based on data for the port of Uddevalla of total queuing time Z and port throughput Q as well as the berth occupancy rate ϕ, provide a basis for illustrating the different levels of the price-relevant marginal costs.

Table 6.1 shows how the price-relevant queuing costs are derived on the basis of the previously estimated functions $Z(\phi)$ and $Z(Q)$, assuming (as before) that the mean value of a ship-day, v = $1,600, the number of berths, n = 10, and the number of days per month, t = 30.

The optimal structure of charges obviously depends on the demand elasticities. Let us take the two extremes. Suppose, first, that all demands are completely inelastic. In this case pricing has, of course, no allocative purpose. Before and after the introduction of peak-load pricing the pattern of demand will be the same. Then the structure of charges is directly obtained by combining table 6.1 and table 4.1. On the assumption that a particular charge is levied for each month, the total revenue from berth occupancy charges is easily calculated. Using the monthly average value according to table 6.1, the total annual revenue would amount to $272,000. This sum is to be compared to the actual revenue from port charges in Uddevalla, which, for example, in 1974 was almost exactly $1 million.

Suppose, on the other hand, that the cross elasticities are very high, with the result that the expected traffic volume each month becomes the same after the introduction of peak load pricing. Suppose also that no traffic is gained from, or lost to, other ports, but that the whole change in the demand time profile is constituted solely by redistribution of traffic already using the port of Uddevalla. It is interesting to note that the total annual revenue from port charges would be almost the same ($260,000) as in the opposite extreme case.

A separate charge each month may be a too detailed differentiation. Two or three, different levels of the charges would probably be more adequate – for example, a peak charge for

March, a somewhat lower charge for January, February, and April, and a low off-peak charge for the rest of the year. Suppose – just to take an example – that an optimal threefold seasonal differentiation of the port charges would reduce the level of traffic in March by 20% and the level of traffic in January, February, and April by 10%, and raise the off-peak level of traffic correspondingly. The charges consistent with this result are given in table 6.3.

Table 6.3 *Hypothetical peak-load pricing structure in the port of Uddevalla in 1974*

Period	Optimal charges per service-day ($)	Occupancy rate	
		Before peak load pricing	After peak load pricing
Mar.	170	0.76	0.61
Jan., Feb., Apr.	149	0.63	0.57
May-Dec.	115	0.46	0.50

In this hypothetical case the total revenue would be $246,000. The absolute values of the calculated prices and revenues are of no particular significance. Only in relation to the port costs do they tell us something. As mentioned, the total revenue from port charges in Uddevalla came to $1 million in 1974. If this sum could be regarded as a proxy for total port costs, it consequently seems that optimal time-differentiated berth occupancy charges would bring in revenue of about 25% of the total costs.

6.2 Congestion and cargo handling charges

It could be argued on the basis of the analysis of section 7 above, that the berth occupancy charges should be supplemented by some sort of “congestion charges”. The empirical evidence of congestion costs in the port of Uddevalla is inconclusive, but let us take it at face value just to illustrate the price theory.

In this case the “congestion” affects two cost categories – the user costs of ships’ laytime along the quay, and the producer costs of stevedoring services, i.e. cargo handling.

Take first the laytime, or service time cost. According to the regression results of table 7.1, the berth occupancy rate (in per cent) is related to throughput in this way:

$$100\phi = 0.243 \cdot 10^{-3} Q^{1.06}$$

Multiplying both sides of the equation first by $10 \cdot 30/100$, to get an expression for total service time per month, *X* as a function of *Q*, and secondly by \$1600, the total service time cost (TC_{st}) as a function of *Q* is obtained:

$$TC_{st} = 1.17Q^{1.06}$$

The corresponding price-relevant cost is according to formula (8.1) above obtained by taking the average cost, and multiplying by the average cost elasticity with respect to throughput.

$$E \cdot AC_{st} = 0.06 \cdot 1{,}17Q^{.06}$$

The relevant range so far as monthly throughput of the port of Uddevalla is concerned is 60.000 – 160.000 tons. In this range, the optimal charge will rise from \$0.135 to \$0.144, i.e. stay almost constant at a level of the same order of magnitude as the optimal queuing charge for berth occupancy rates between 0.4 and 0.5. Unlike the latter charge, there is no need for a peak-load structure of the congestion charge, at least so far as the influence on the service time cost is concerned.

Next comes the congestion cost derived from the stevedoring cost function. This is a producer cost, which means that the whole marginal cost is price-relevant. Cargo handling charges should in the first place cover the direct stevedoring costs. The main component should be equal to the average stevedoring cost. If the marginal stevedoring cost exceeds the average cost, which is the case according to the regression results presented in table 7.2 above, the difference between MC_{ch} and AC_{ch} could be viewed as “quasi-rent”, or a contribution margin towards covering short-run fixed costs (“ch” stands for cargo handling). This contribution margin can be calculated by the same formula (8.1) as was previously used.

$$MC_{ch} - AC_{ch} = 0.23 \cdot 0.101Q^{.23}$$

This “contribution margin/congestion cost” ranges from \$0.29 to \$0.36 per ton in the relevant throughput range of 60.000 to 160.000 tons per month, which also seems to be too narrow to warrant seasonal differentiation.

6.3 The financial result of optimal pricing

Summing up the queuing and congestion costs in the port of Uddevalla the following total result of optimal pricing appears: After covering the cargo handling costs by stevedoring charges, there are three remaining sources for financing the port capital

1) *queuing charges* in the form of berth occupancy charges and/or charges on the cargo
2) *congestion charges*, which also can be levied either on the ship, or on the cargo, or on both
3) a *"contribution margin"* corresponding to the difference between the marginal and average stevedoring cost.

Proper peak-load pricing by queuing charges, where the peak charge would be $0.22 and off-peak charge $0.15 per cargo ton, would give a total revenue of about a quarter of a million US dollar, according to the calculations of section 6.1. The optimal congestion charges computed per cargo ton turn out to be somewhat higher: about $0.14 plus $0.32 on account of increased service time costs and stevedoring costs, respectively.

Altogether total revenue from queuing and congestion charges would be about $870.00 to put against a total port cost of $1 million, that is a degree of cost coverage of 87%.

It is obvious that the latter percentage depends to a considerable extent on the rate of overall capacity utilization – at least so far as the queuing charges are concerned. In the port of Uddevalla capacity utilization was said to be comparatively high for Swedish conditions. There are important physical restrictions on port expansion in Uddevalla. In recent times there has generally been increasing pressure on ports for expansion, and/or relocation out of town to sites where deeper waters and plentiful backup land exist. The consequence of these tendencies for the price-relevant marginal cost of port services is what we will focus on in the following chapters of the report.

7. LONG-RUN MARGINAL COST ANALYSIS: CASE STUDY OF THE PORT OF NORRKÖPING

Among the industries supplying transport infrastructure services, the seaport industry is clearly the one which has undergone the most radical change since the 1960s. Many, probably most old ports of the world are on the move, or have just completed the move out of town to shores with deeper waters, and plentiful back-up land for the cargo handling and storage. In this process the seaport infrastructure has been relocated as well as thoroughly reshaped.

This has been a necessity for several reasons. First, the ship size has grown tremendously in the latter half of last century, and the water depth of old seaports along a river, at the estuary of the river, or at the shallow end of a bay, has proven absolutely insufficient. Secondly, the modern big ships carry shiploads many times larger than traditional tramps and break-bulk liners, and the handling and storage of these large shiploads require plenty of back-up land. Thirdly, increasing traffic of heavy goods vehicles at the land-side of the port in the central city are nowadays both the weakest link in the transport chain, and unacceptable for environmental reasons. Fourthly, dangerous goods in large quantities are involved in sea transport, and should be handled well outside the population concentrations.

Add to this the mechanization of the cargo handling methods, both for general cargo, and bulk cargo, which has transformed the port industry from an extremely labour-intensive to a markedly capital- and land-intensive industry in a few decades, and the complete transformation of the traditional seaport, around which many old towns and cities grew up, is apparent.

7.1 Purpose and plan of the “development cost” calculation

The financial result of optimal pricing calculated for the port of Uddevalla in the preceding chapter may, or may not be characteristic of other seaports. It depends on, above all, whether category A or B applies, and in the former case, to what extent the central city location is a very strong restriction on expansion of backup land. The main case study of this project is a “development cost” calculation for the port of Norrköping, which during the period of observation – from 1960 to 2000 – has expanded considerably, and moved most of its business out from the mouth of the small river “Motala ström” at the edge of the central city of Norrköping to sites further out in the bay Bråviken. Queuing and congestion are reported to be nearly non-existent in the port, in spite of the fact that total throughput has trebled in the period of observation.

The “development cost” is an alternative name for the long-run marginal cost which indicates that capacity expansion over time is in focus. The implicit idea is that pricing policy should prevent over-expansion, which might follow from not taking the capacity development cost into account, which is seemingly done by only considering the short-run cost. Methodologically the idea is that by long time-series analysis the full effect on the costs of the capacity expansion “caused” by growing demand could be ascertained.

There are two main problems with this approach when it comes to empirical estimation, which we have paid special attention to in the case study:

1) The user cost effects of investments in new capacity must not be overlooked.
2) Technical change and growing experience of the technology adopted during the long period of observation have to be allowed for.

Port operation technology has had a period of rapid progress up to the 1960s. Most of the technology applied today in ports was well known 30 years ago, but requires both large throughput volumes, and, above all, large land areas to be economical. The driving forces of the expansion of the port of Norrköping can be characterized as (i) a compliance with the demand for deeper waters, and vacant land required by larger ships and shiploads, by moving out of town, and (ii) efforts to increase total business volume by recruiting new customers, in order to reap the fruits of the new labour-saving technology, which has tended to have a strong bias towards large-scale production. In order to estimate the true economies of scale by time-series analysis, it is necessary to control both for technical progress, and increased skill due to repetition (compare Lieberman, 1994, and Griffith and Wall, 2000).

The user cost aspect is even more fundamental. In principle expression (2.2), in chapter 2 for the price-relevant marginal cost, which is repeated below as (7.1), has exactly the same appearance in the short run as in the long run. Each term, however, has a completely different content. The first term, MC_{prod} now contains the port capital cost as well as the running cost including cargo handling labour cost. The middle term of, $Q\ (dAC_{user}/dQ)$ is now a product of total throughput and the change in the shipowners costs per ton as a result of an increase in throughput accompanied by the actual capacity expansion taking place. It turns out – here like in other better known cases of transport infrastructure expansion – that the middle term of (1) will be a relatively large *negative* component of the price-relevant marginal cost, which tends to make the final value coming out of the development cost calculation much lower than the average total port cost.

$$MC = MC_{prod} + Q\frac{dAC_{user}}{dQ} + MC_{ext} \qquad (7.1)$$

Concerning the third term, let us, state again that the transport-system-externalities are left out of consideration in the present discussion; just the two first terms are considered.

By ”long-run” is implied that most important aspects of port design, if not necessarily all, are variable. Above all, port capacity cost is now a variable cost. As regards the other main producer cost component, the stevedoring cost, the non-fixity of capacity means that the stevedoring cost is not starting to rise with increases in throughput, which is characteristic of the short-run marginal cost, but may instead fall on account of labour-saving capital investments .

In order to focus the empirical analysis on the most interesting aspect in this connection, which is the ratio of MC to AC_{prod}, a slight reformulation of the expression for the price-relevant cost above is helpful.

We are introducing the following cost-elasticity designations:

$$E = \frac{dTC}{dQ}\frac{Q}{TC} = \text{the elasticity of total (= producer and user) cost with respect to throughput}$$

$$E_{prod} = \frac{dTC_{prod}}{dQ}\frac{Q}{TC_{prod}} = \text{the elasticity of total producer cost with respect to throughput}$$

$$E_{user} = \frac{dTC_{user}}{dQ}\frac{Q}{TC_{user}} = \text{the elasticity of total user cost with respect to throughput}$$

Ignoring MC_{ext}, we can first write the price-relevant marginal cost (MC) like this:

$$MC = MC_{prod} + \frac{dAC_{user}}{dQ} \frac{Q}{AC_{user}} AC_{user} = E_{prod} AC_{prod} + (E_{user} - 1) AC_{user}$$

Remembering that E is the weighted sum of E_{prod} and E_{user},

$$E = E_{prod} \frac{AC_{prod}}{AC} + E_{user} \frac{AC_{user}}{AC},$$

we can substitute E for E_{prod} and E_{user} in the MC-expression in this way:

$$MC = E\,(AC_{prod} + AC_{user}) - AC_{user}$$

Forming the ratio of MC to AC_{prod}, we finally have

$$\frac{MC}{AC_{prod}} = E + \frac{AC_{user}}{AC_{prod}} (E - 1) \qquad (7.2)$$

As seen, in order to find out to what extent port prices based on the long-run marginal cost would cover the total port cost, we need to know (i) the total (producer + user) cost elasticity with respect to throughput, and (ii) the ratio of user cost to producer cost. To estimate (i) and (ii) is the main purpose of the case study to follow is to estimate these two unknowns.

7.3 The present situation of the port of Norrköping

The port of Norrköping is situated on the East Coast of Sweden at a bay of the Baltic called Bråviken (see Figure 7.1 below). It is one of 50 public ports in Sweden. Public ports are open for all sorts of traffic. Industrial (private) ports, on the contrary, are only open for traffic between the industries that own the port. Holmen[1] and Djurön are examples of industrial ports in Bråviken. They are linked to Holmen Paper and ODAL (Swedish Farmers).

---

[1] Holmen is marked as Braviken in the upper map of figure 7.2.

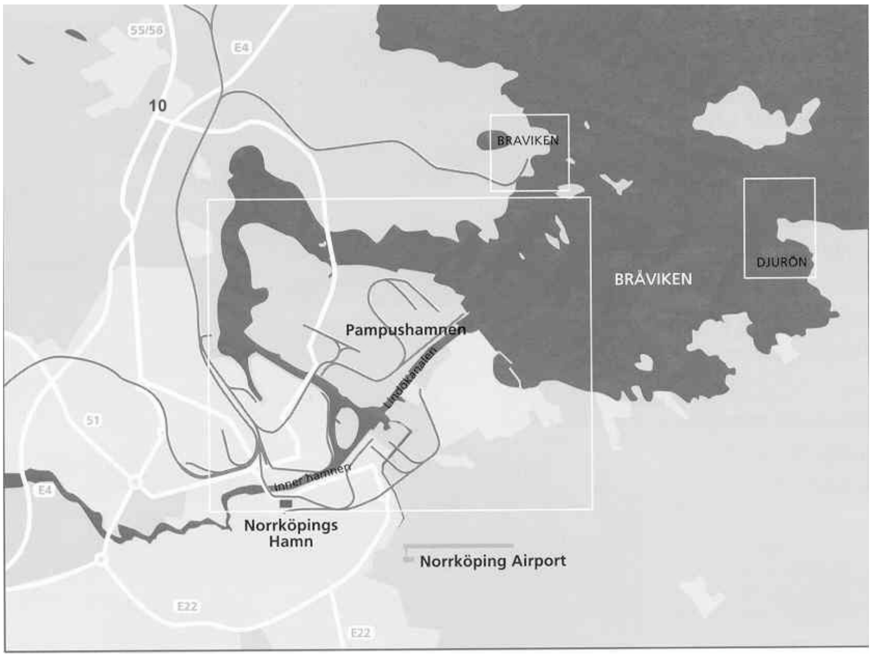
In many public ports the municipality owns the fixed facilities, while a stevedoring company is responsible for cargo handling. This was also the case in the port of Norrköping until 1990. Then the company Norrköping Port and Stevedoring (NHS) was founded. The municipality of Norrköping owns 40% of the shares in NHS, Holmen Paper 27%, ODAL 27.4%, Unér Shipping 2.8% and Novabolagen SB 2,8%.

All kinds of cargo types are handled in the port of Norrköping. Petroleum, forest and agricultural products dominate the cargo volume in the port. The cargo handling of agricultural products at Djurön has made the port of Norrköping one of the most important grain ports in Europe. The port handles over one million ton of petroleum products every year. Norrköping is thanks to Holmen Paper, AssiDomän and Fiskeby Board the leading producer of paper in Sweden. Thus, the port of Norrköping has become an important link for transports of paper products. Moreover, the port is one of Sweden's major ports of shipment for timber and sawn wood products. The investment in new cranes has also made the port of Norrköping one of the best-equipped Swedish ports to handle heavy goods (up to 320 ton).

Figure 7.2 shows two maps of the port area of Norrköping. The top map illustrates the port's location in the inner parts of Bråviken, in connection with the outflow of the small river through Norrköping, Motala ström. The map below gives a more detailed view of the port area, and the different quays.

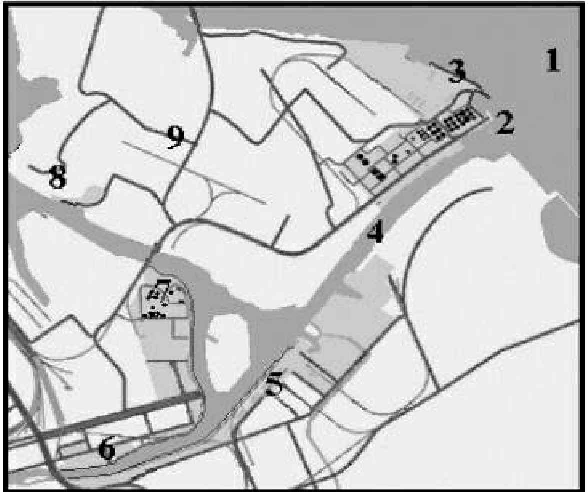
Figure 7.1. **Map of Sweden with some of the more important seaports.**

Figure 7.2a. **Overview map of the port of Norrköping.**



The port of Norrköping is situated within the larger square in the middle of Figure 7.2a (which is enlarged in Figure 7.2b below). The two industrial ports, Braviken and Djurön, are within the two smaller squares.

Figure 7.2b. **Map of the port of Norrköping.**



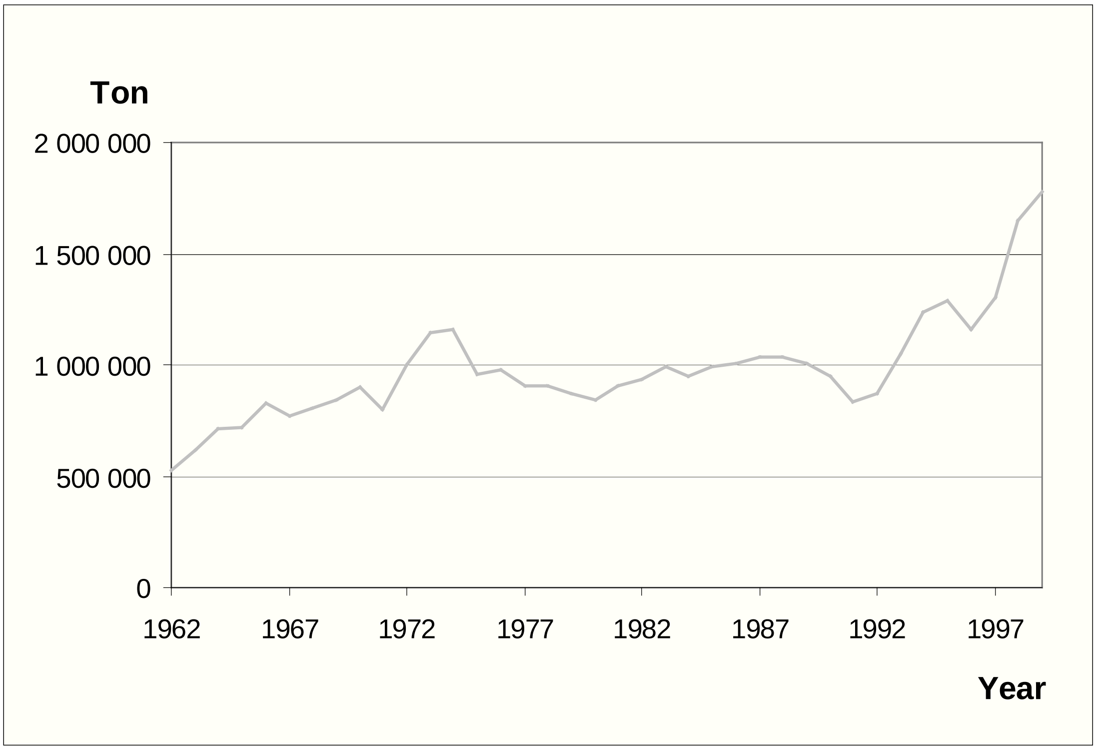
1. Bråviken
2. The Pampus petroleum quay
3. The Pampus dry cargo quay
4. The Lindö channel
5. The Öhman quay
6. The Northern and Southern quays; timber quays
7. Gästgivarehagen; former petroleum quay
8. Ramshäll; the military petroleum quay
9. The island of Händelö
10. Road junction between E4 and the road to Händelö

The quays numbered 5-8 are referred to as the “inner port”.

7.4 The development of port throughput, investments, and costs

In Diagram 7.1 aggregate throughput in the port of Norrköping during the period 1962-1999 is given. The throughput has trebled during the observed period.

Diagram 7.1: Cargo turnover in the port of Norrköping 1962-1999; industrial and petroleum ports excluded (ton)



Source: Norrköping Port and Stevedoring: annual reports 1960-1999

Investments in deeper channels, new quays and improved land transport infrastructure have accompanied the increasing cargo volume.

The Lindö channel opened in 1962. It was the first of several big investments in the port during the last forty years. The Lindö channel made it possible for ships up to 16,000 dwt (dead weight ton) to call at the inner port. Before the channel was built, only ships of 4,500 dwt could call at the inner port. Complementary investments were also made in lengthening the Öhman quay to accommodate more and longer ships.

As Figure 7.2 shows, the Lindö channel cuts through the former peninsula Händelö, making the distance to the inner port both straighter and shorter. The Lindö channel has reduced the travel time by 30 minutes (one-way).

Also the petroleum quay in Pampus opened in 1962. The storage space in Gästgivarehagen in the inner port was not enough, and new storage facilities had to be built. By building the new petroleum quay at deeper water, the size of ships could be increased from 4,500 dwt to 50,000 dwt. The travel time for the petroleum ships was also reduced by 30 minutes (one-way) compared to calling at Gästgivarehagen.

But as also dry cargo ship sizes continued to increase, new investments were called for. In 1982 a dry cargo facility was added to the Pampus port, and later a new road connection passing by the city of Norrköping was built between Händelö and E4. The new road connection opened in 1990 and shortened land transports to the north by several kilometers.

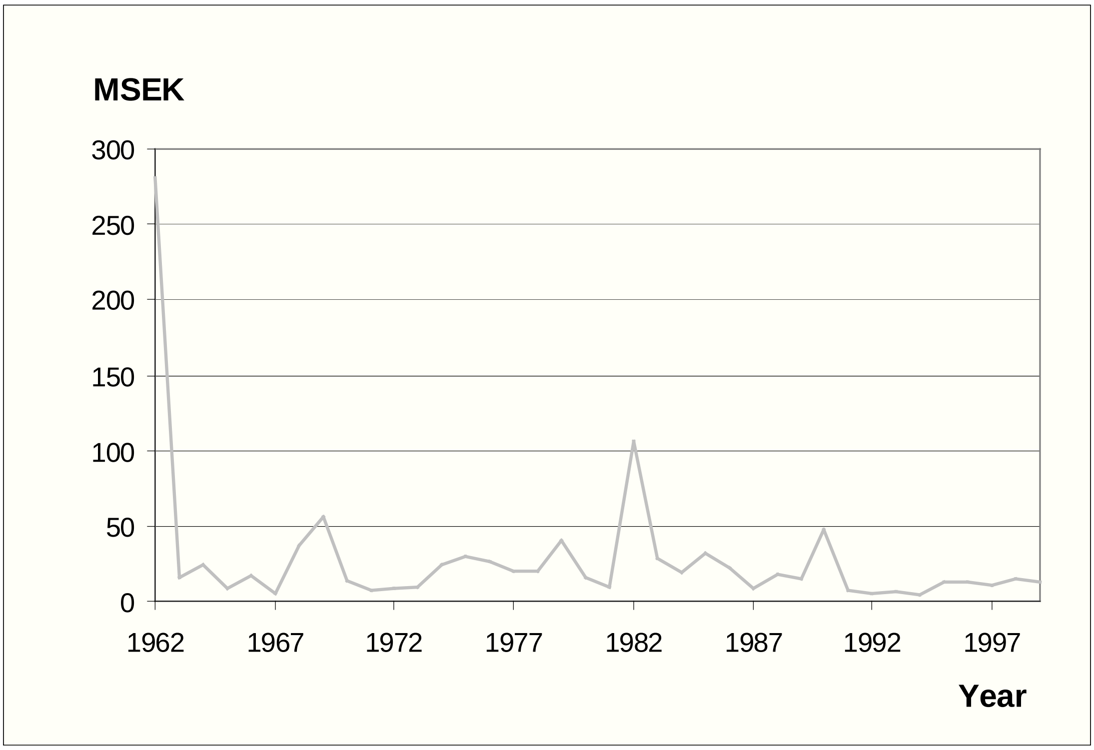
The port of Norrköping invested in several new cranes during the 1990s; both in big railbound Ndc cranes with a lift capacity of 40 ton and in five so called Essemko cranes. After scrapping a number of obsolete cranes, the total number of cranes available decreased but both total capacity and flexibility were enhanced.

Shipments of high qualitative sawn timber have been on the increase in the port of Norrköping in recent years. The timber has to be protected from rain and snow. During the 1990s two transit sheds were built on Pampus with a storage capacity of 17,000 m^3 for timber.

The investments mentioned above are not the only ones, but are the major investments made in the port of Norrköping during the 1960-2000 period. Diagram 7.2 below shows the total investment expenditure in the port of Norrköping during this period in real terms (price level year 2000) excluding investments in petroleum facilities. The peak in 1962 of 280 million SEK represents the investment in the Lindö Channel.

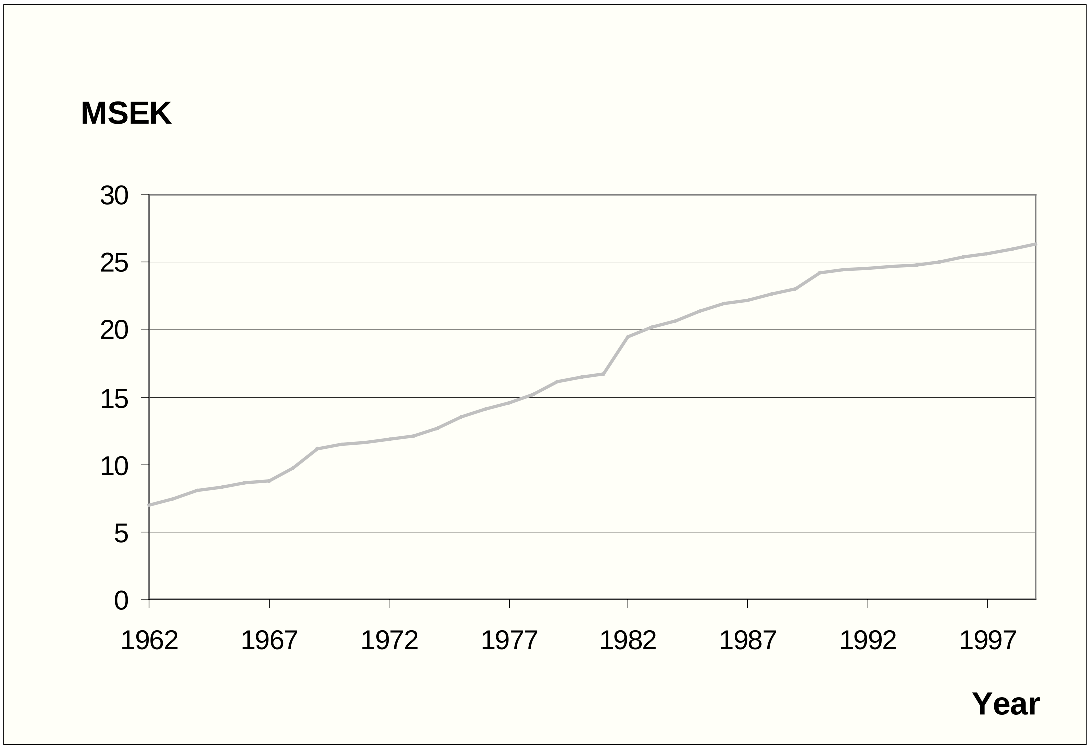
The investment expenditure has been spread over an assumed economic lifetime of 40 years for the different facilities to obtain annual capital costs. This is showed in Diagram 7.3 below.

Diagram 7.2: Investment expenditure in the port of Norrköping 1962-1999 in real terms (Million SEK)



Source: Norrköping Port and Stevedoring: annual reports 1960-1999

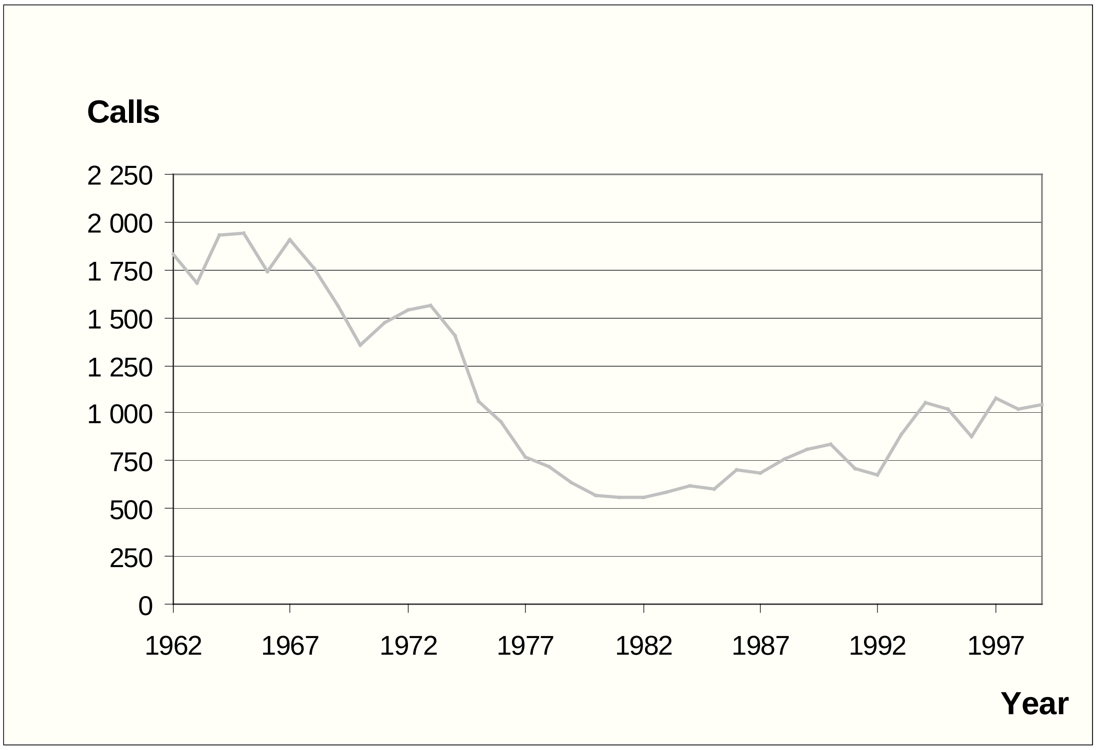
Diagram 7.3: Capital costs of the port of Norrköping 1962-1999; annual installments, 40 year; fixed price level 2000 (Million SEK)



Source: Norrköping Port and Stevedoring: annual reports 1960-1999

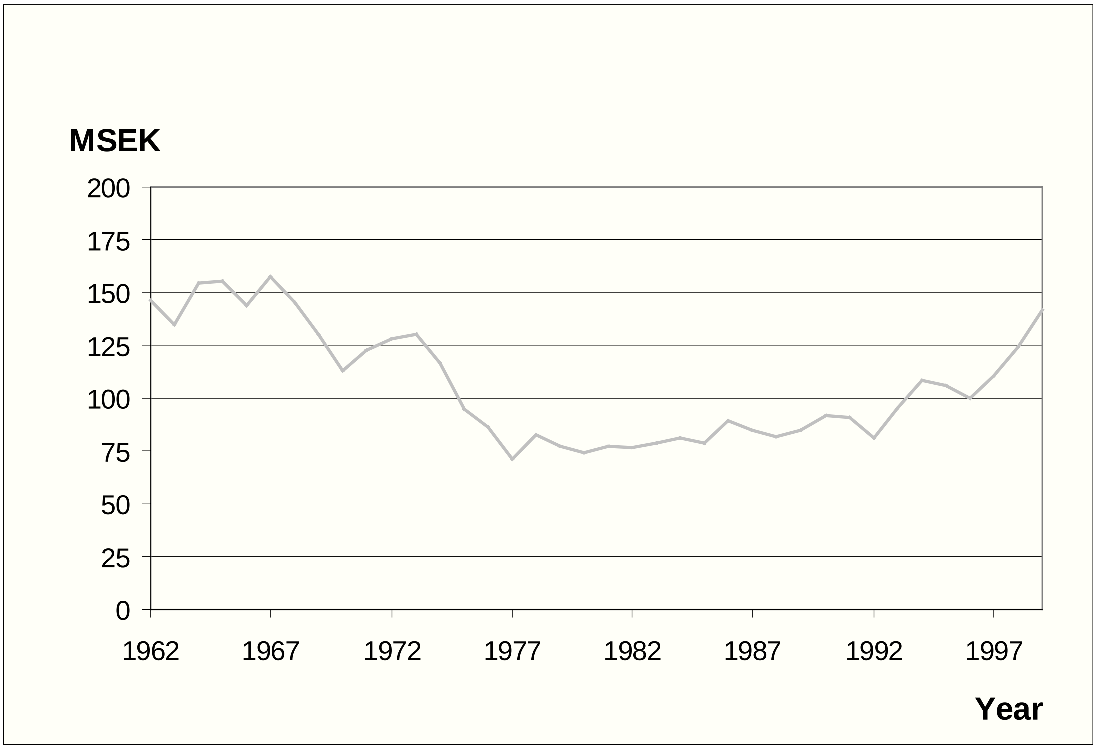
The above-mentioned investments in the port of Norrköping have had a significant effect on the ship size, and consequently on the number of calls at the port. Diagram 7.4 shows that the number of calls at the port of Norrköping have dropped markedly during 1962-1982 in spite of a doubling of total throughput. The investment in the Lindö channel made it possible to use much bigger ships (an increase in ship loading capacity from 4,500 dwt to 16,000 dwt).

Diagram 7.4: Annual number of calls at the port of Norrköping 1962-1999; industrial and petroleum ports excluded



Source: Norrköping Port and Stevedoring: annual reports 1960-1999

Diagram 7.5: Total user costs in the port of Norrköping year 1962-1999; industrial and petroleum ports excluded; fixed price level 2000 (Million SEK)



Source: Bäckelid (2001)

The total user cost is the sum of the total transport cost of all export and import through the port of Norrköping. The daily shipping time cost differs with ship size as seen in table 7.1.

The increase in day cost by size is markedly degressive, which makes cost per ton fall with ship size.

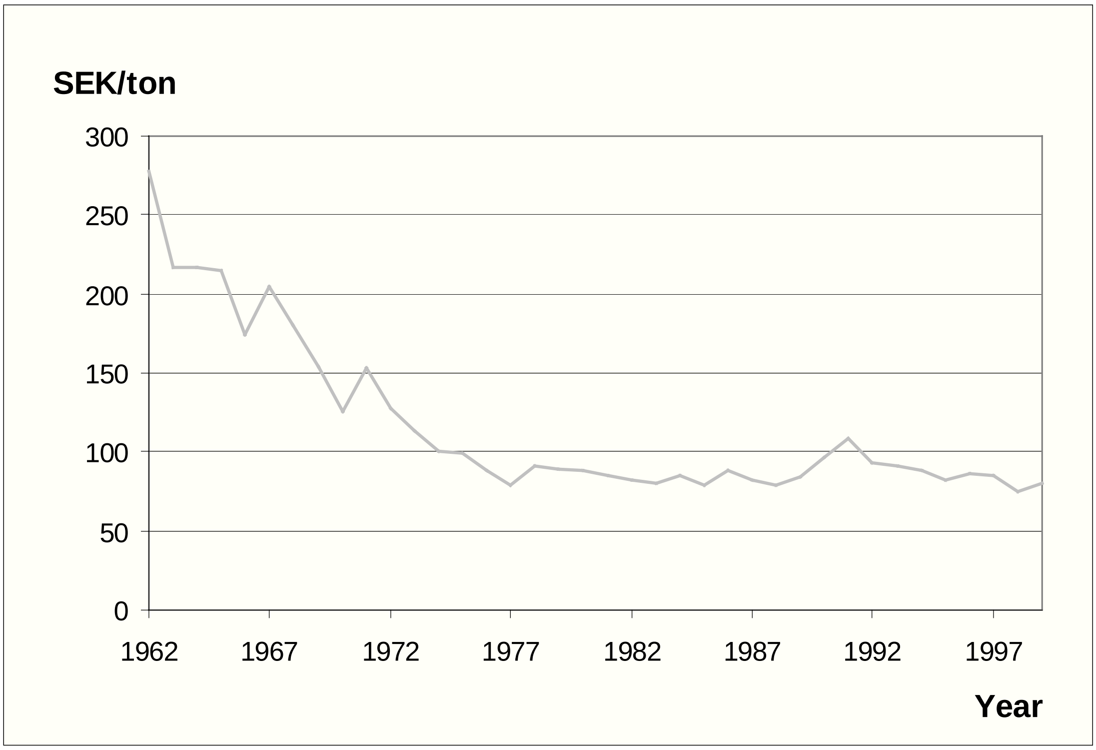
Table 7.1: Daily shipping time cost and cost per ton for different ship sizes

Class	**Ship size (dwt)**	**Average ship size (dwt)**	**Daily shipping time cost (SEK)**	**Cost per shipload ton**
I	- 4,500	2,250	45,000	20.0
II	4,500-10,000	7,250	70,000	9.7
III	10,000-16,000	13,000	85,000	6.5
IV	16,000-	30,000	120,000	4.0

Source: Bäckelid (2001)

The trebling of throughput in combination with the infrastructure investments made in the port of Norrköping have made it possible to gradually use bigger ships. Diagram 7.6 consequently shows that the user cost per ton has fallen in the observed period.

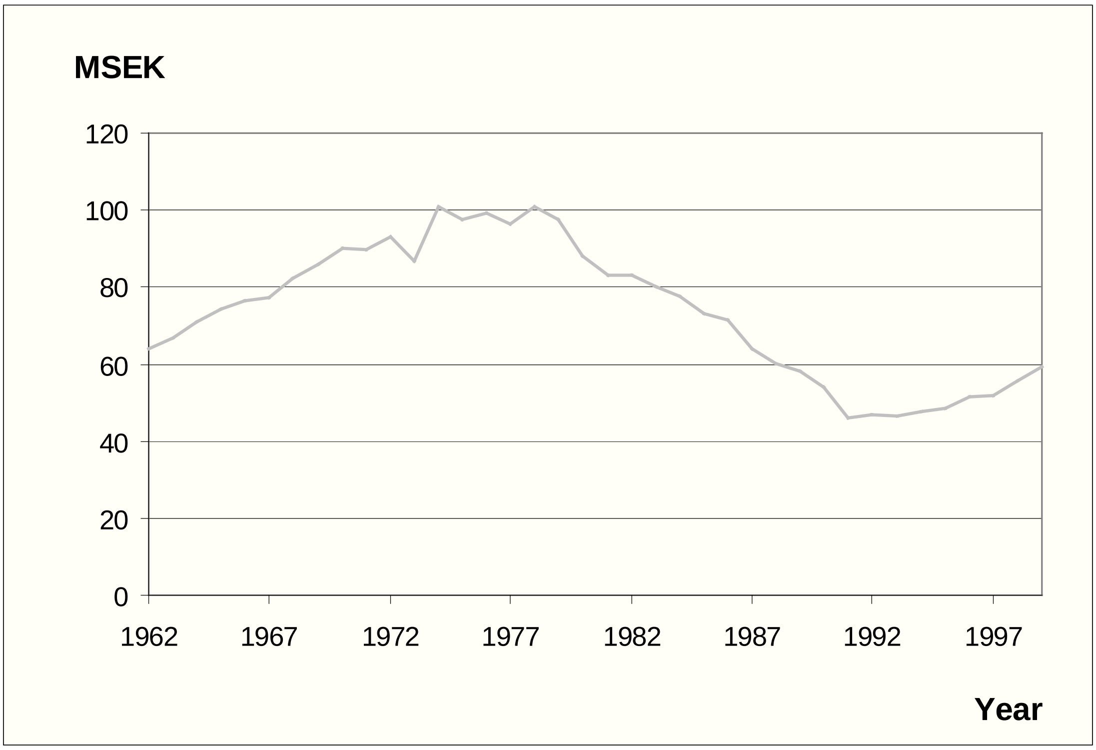
Diagram 7.6: User cost per ton in the port of Norrköping year 1962-1999; industrial and petroleum ports excluded; fixed price level 2000 (SEK/ton)



Source: Bäckelid (2001); Norrköping Port and Stevedoring: annual reports 1960-1999

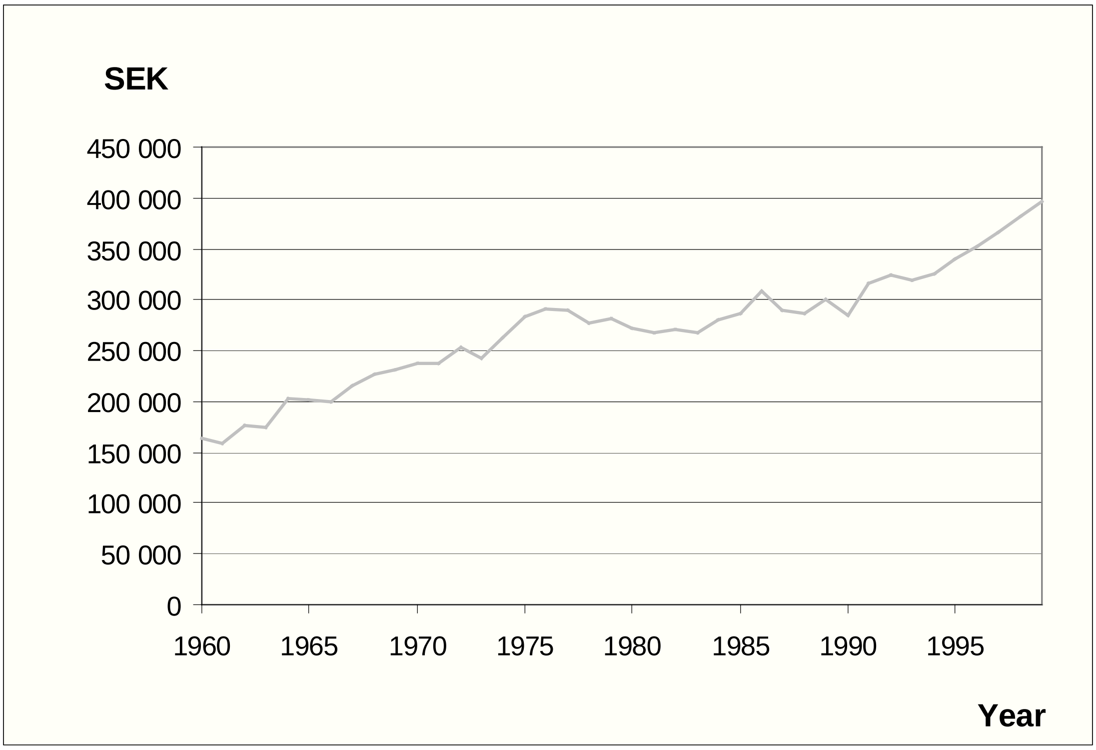
Diagram 7.7 to 7.9 below show the development of labour costs, the average labour cost and the number of employees in the port of Norrköping in 1962-1999. Investments have been markedly labour-saving, in particular in the 1970s and 1980s. In the 1960s, on the other hand, investment seemed to be more biased toward user cost savings.

Diagram 7.7: Total labour costs in the port of Norrköping year 1962-1999; fixed price level 2000 (Million SEK)



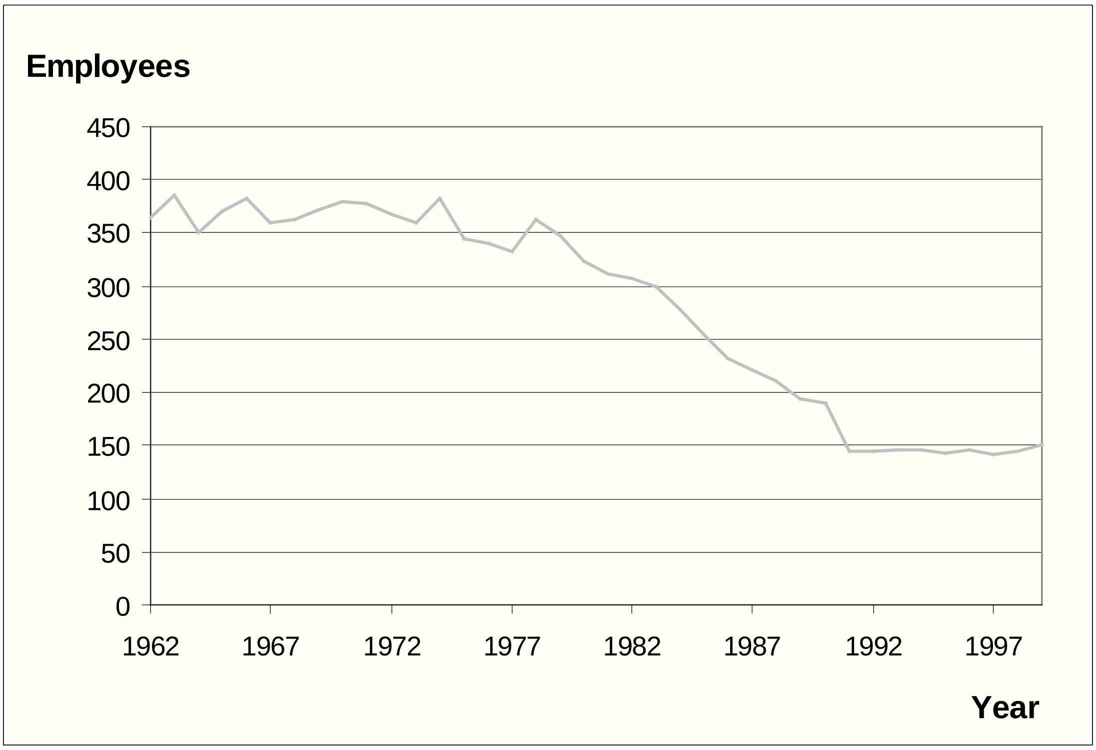
Source: Norrköping Port and Stevedoring: annual reports 1960-1999

Diagram 7.8: Average labour cost in the port of Norrköping 1962-1999; fixed price level 2000 (SEK)



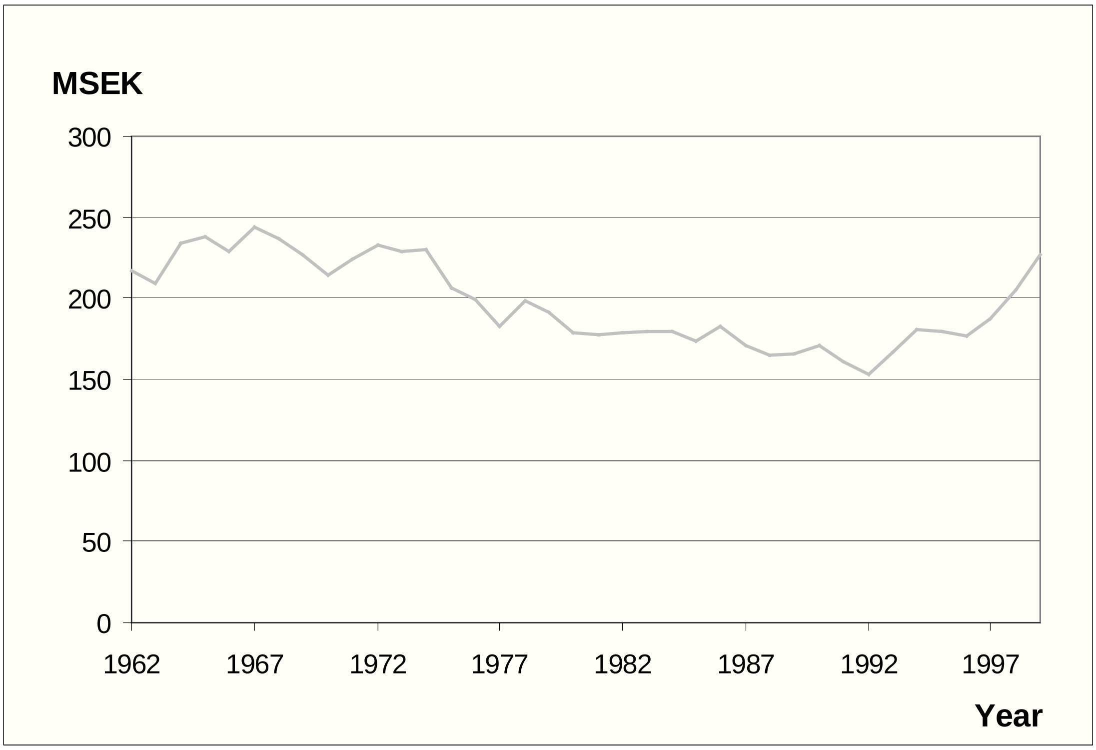
Source: Norrköping Port and Stevedoring: annual reports 1960-1999

Diagram 7.9: Number of employees in the port of Norrköping 1962-1999.



Source: Norrköping Port and Stevedoring: annual reports 1960-1999

Diagram 7.10: Total costs in the port of Norrköping 1962-1999 (capital + labour + user costs); fixed price level 2000 (Million SEK)



The empirical approach is now to study by time series analysis how the total transport system costs have developed as traffic has been growing, including port capital and labour as well as user inputs in the form of ships time.

7.5 Empirical estimation of the long-run total cost function

As mentioned in chapter 7.2, we want to empirically estimate the long-run total cost function:

$$TC = f\left(Q, t, \sum Q\right) \qquad (7.3)$$

By including time and cumulative throughput in the total cost function, we control both for technical progress over time and “learning by doing”. It is important to control for both factors in order to estimate the true economies of scale by time-series analysis.

“Learning by doing” reflects the importance of learning from experience. The literature (e.g. Griffiths and Wall 2000; Pindyck and Rubinfeld 2001) sometimes refers to experience curve, learning curve or progress cost curve, when discussing firms cost savings over time as cumulative output increases. These cost savings differ from those arising from economies of scale.

Economies of scale arise when all factor inputs are variable and when the average cost of a firm has declined as a result of factor inputs being adjusted. When the optimum plant size has been achieved, microeconomic theory implies that the firm will produce the same output every year.

But workers tend to improve their performance as their skill increases due to repetition. They also tend to find more effective working methods as labour experience grows over time. In other words: total production experience accumulates with cumulative throughput, making the average cost fall even further due to “learning by doing”.

Taking these two factors into account, we now rewrite equation (7.3) on logarithmic form as

$$\log TC = \log a + b \cdot \log Q + c \cdot \log Q_{cum} + d \cdot year \qquad (7.4)$$

where Q_{cum} is defined as cumulative throughput.

Running a regression analysis on equation (7.4), we will get significant t-values and appropriate signs for all three coefficients. The adjusted R^2 is 0.86.

Table 7.2: Result of regression analysis of the total costs on port throughput, cumulative port throughput and time (38 observations; 1962-1999).

Model	**B**	**Std. Error**	**t-value**	**Sig.**
Constant	32.883	3.291	9.991	0.000
LN_Q	0.590	0.060	9.905	0.000
LN_QCUM	-0.09234	0.021	-4.390	0.000
YEAR	-0.01031	0.002	-5.572	0.000

Dependent variable = total costs; independent variables = port throughput, cumulative throughput and time

There is high correlation (0.912) between time and the logarithm of cumulative throughput. But this is expected as cumulative throughput by its nature is highly linked with time. However, the two variables represent two sides of the same coin. The former is trying to explain the influence of technical progress on the total cost function over time. The latter is meant to explain that, given the technology, total experience accumulates with cumulative throughput, which will affect the total cost function downwards.

Dropping one or the other of the variables from the equation does not change the signs of the coefficients. The parameter estimate of LN_Q is only slightly affected. All t-values are significant in both cases and all standard errors are still small. The adjusted R^2 is 0.79 when dropping the cumulative throughput and 0.74 when dropping time from the equation.

Table 7.3: Result of regression analysis of the total costs on port throughput and time (38 observations; 1962-1999).

Model	**B**	**Std. Error**	**t-value**	**Sig.**
Constant	45.015	2.206	20.406	0.000
LN_Q	0.504	0.069	7.260	0.000
YEAR	-0.01659	0.001	-11.485	0.000

Dependent variable = total costs; independent variables = port throughput and time; excluded variable = cumulative throughput

Table 7.4: Result of regression analysis of the total costs on port throughput and cumulative port throughput (38 observations; 1962-1999).

Model	**B**	**Std. Error**	**t-value**	**Sig.**
Constant	14.899	0.879	16.959	0.000
LN_Q	0.522	0.079	6.563	0.000
LN_QCUM	-0.183	0.018	-10.084	0.000

Dependent variable = total costs; independent variables = port throughput and cumulative throughput; excluded variable = time

By not including both variables in the model, we will not be able to estimate the true economies of scale. By controlling for both variables we will get a more reliable estimation of the parameter b, i.e. the elasticity of total cost with respect to throughput.

Plotting the estimated residuals against the total cost suggests that we might have a problem of heteroskedasticity in the model. However, using White's test for heteroskedasticity[2] we conclude that the hypothesis of homoskedasticity can not be rejected at the 10 percent critical value[3]. The White statistic is 5.168 while the 10 percent critical value for the chi-squared distribution with four degrees of freedom is 7.78.

Testing for first order autocorrelation (between the current and lagged residual) by using the Durbin-Watson test fails as the DW-statistic falls in the range where the DW-test is inconclusive. The Breusch-Godfrey test[4], a Lagrange multiplier test, is an alternative test that is less restrictive. Using the Breusch-Godfrey test for testing for autocorrelation of the first, second, third and forth order, we conclude that the hypothesis of no autocorrelation can not be rejected at the 10 percent critical value.

---

2 For more information on White's test for heteroskedasticity, consult a book on econometrics, e.g. Greene (2000) or Maddala (2001).

[3] With three variables in the equation, the White test implies that the degrees of freedom should be 9. However, five of the variables entered were highly correlated and removed by the program when running the regression. The reason for this is that Q and Q_{cum} are on logarithmic form and that the changes in values are very small for both variables. For instance, the values of log Q lie in the range 13.18-14.39 while the values of log Q_{cum} lie in the range 13.18-17.43. With four variables entered the White statistic is 5.168, calculated from nR^2 with n=38 and R^2=0.136.

[4] For more information on the Breusch-Godfrey test, consult a book on econometrics, e.g. Greene (2000).

Table 7.5: Results of the Breusch-Godfrey test of no autocorrelation

Autocorrelation	**R^2**	**N**	**$(n-1)R^2$**	**d.f.**	**χ^2 (90%)**
1st order	0.062	38	2.294	1	2.71
2nd order	0.068	38	2.516	2	4.61
3rd order	0.085	38	3.145	3	6.25
4th order	0.086	38	3.182	4	7.78

Our tests indicate that there is no presence of heteroskedasticity or autocorrelation in our estimated model. Thus equation (7.4) seems to be a good empirical estimation of the long-run total cost function. The elasticity of total cost with respect to throughput, b, is estimated to 0.59.

7.6 The final result of the development cost calculation

Going back to formula (7.2) for the ratio of the price-relevant long-run marginal cost to the port service producer average cost, MC/AC_{prod}, the result of the regression analysis is now used to calculate this ratio. It was found that $E = 0.59$. The ratio of user to producer cost can is also obtainable from the cost data of the empirical study. In an appendix a complete sheet of all the data used in the estimation is presented, containing i.a. total user and producer cost series from 1962 to 1999. Since the ratio of these two costs has varied during the observation period, the question is, whether to take an average of all the years, or a value for a particular year? An average value for all 38 years of the ratio AC_{user}/AC_{prod} comes to 1.17, whereas taking just the ten years of the last decade, the average value of this ratio is 1.39. Applying the former value to (7.2) gives as a result that the price-relevant cost is only a fraction = 0.11 of the producer average cost, and with the latter value this fraction becomes virtually zero (= 0.02).

8. CONCLUSIONS

In the case of the port of Norrköping, which is a typical example of a category B seaport on the move out of town towards the open sea, the level of the price-relevant long-run marginal cost has been found to be close to zero. This is consistent with the reporting of practically no queuing of ships waiting for a vacant berth, as well as the absence of congestion on quay aprons, access roads, or in transit sheds, which might have a negative effect on the cargo handling productivity. It appears that port capacity has been expanded well ahead of demand, and at the same time the quality of service seems to have been improved, which is manifest by falling costs of the port users. Perhaps this could be put the other way round: substantial port investments have been made in order to accommodate bigger ships (with lower shipping costs per ton) and to reduce the access time to the port, and the service time for the ships at the berths. And these quality improvements have also led to increasing capacity.

In principle the same conclusion has consequently been arrived at, as that which was drawn for the first time in Walters (1968) concerning non-urban roads: barring transport-system externalities, the price-relevant marginal cost of most non-urban roads is practically zero. The main reason for this is that capacity and quality of service are markedly joint attributes of non-urban road investments or “joint products” as Walters put it.

A similar distinction as that between urban and non-urban roads is the present division of seaports into category A and category B. In the former case, where a seaport remains in its original location in the old town/city, where land is very scarce, and/or expansion of the back-up area is next to impossible, the price-relevant marginal cost can rise to very high levels, if the demand for port services is increasing. The contrast with category B ports is as sharp as that between a central city road and an interurban motorway, so it should not come as a surprice that the pricing policy recommendations can be completely different for category A and category B ports.

For category A ports marginal cost pricing in the form of peak-load pricing can be an important means of improving the resource allocation. The charges could be differentiated also in other dimensions in accordance with varying capacity requirements of different ships and cargoes.

In category B ports it can be expected that only the system-external marginal cost should have a more appreciable influence on port charges. The third term of the general expression for the price-relevant marginal cost, MC_{ext} has not been discussed very much in this study. This is not an unimportant issue. For example, the practice of generating the required electric power while in port by running the bunker-fuelled ship's engine at half speed, rather than relying on a mains connection is a malpractice that could be stopped by proper price incentives. Polluter payments should not be a source of revenue for the port authority – that could create a perverse incentive – but go to the state purse, so the present conclusion as to the financial result for category B ports of optimal pricing is not changed by taking MC_{ext} into account.

8.1 Ramsey pricing of seaport services

Self-financing ports are desirable for other reasons than first-best allocative efficiency, so the port industry sector consisting of category B ports should resort to Ramsey pricing in order to achieve second-best allocative efficiency under a budget constraint. This is what many ports actually do, and have done for a long time under the motto of "charging what the traffic can bear". This old pricing principle takes two main forms in practice:

i) to the extent that an explicit tariff of port charges is issued (and adhered to) high-value goods are charged much higher port prices than low-value goods, irrespective of the relative ease of handling the cargo,

ii) the pricing is done by direct (secret) negotiations with each more important customer in terms of cargo volume.

Both methods aim at a differention of charges in accordance with the inverse of the price-elasticity of demand – in the former case by a time-honoured rule-of-thumb, which, however, is rather blunt, whereas the latter approach can take many more factors into account.

A last word of caution is appropriate before concluding that theory and practice for once are reasonably well in accord in the field of transport infrastructure pricing. The Ramsey pricing should be worked out from the point of view of the whole port industry. If inter-port competition were negligible, independent (decentralized) Ramsey pricing in each individual port would come to the same, but if the competition between adjacent ports is strong for some commodities, the resulting price structure of each individual port "charging what the traffic can bear” could be appreciably subotimal from the port industry point of view. High charges would apply to commodities for which a particular port enjoys a natural monopoly, and low charges to commodities for which adjacent ports are competing. The resulting pattern of output of the industry, i.e. the import and export volumes of different commodities, could substantially differ from the second-best pattern according to the Ramsey Rule, which prescribes that the same relative difference between first-best and actual output quantities should apply.

REFERENCES

Adams, G (1973). Organisation of the British Port Transport Industry. National Ports Council.

Bennathan, E and Walters, A A (1983). Port Pricing and Investment in Developing Countries. A World Bank research publication. Oxford University Press. Oxford.

Bird, J (1971). Seaports and Seaport Terminals. Hutchinson University Library. London.

Bäckelid, D (2001). CBA i hamnväsendet – hur och varför, fallstudie Norrköpings hamn. Working paper, EKI, Linköping University.

Cowiconsult (July 1995). Strategic Study of Ports and Maritime Transport in the Baltic Sea.

European Commission (1995). COM (95) 317. The development of short sea shipping in Europe: Prospects and challenges. Brussels.

European Commission (1997). COM (97) 678 final. Green Paper on Sea Ports and Maritime Infrastructure. Brussels.

European Commission (1998). COM (1998) 466 final.White Paper on Fair Payment for Infrastructure Use: A phased approach to a common transport infrastructure charging framework in the EU. Brussels.

European Commission (1999). COM (1999) 317 final. The Development of Short Sea Shipping in Europe: A Dynamic Alternative in a Sustainable Transport chain – Second Two-yearly Progress Report. Brussels.

European Commission (2001a). COM (2001) 35 final. Reinforcing quality service in seaports: a key for European transport. Brussels.

European Commission (2001b). COM (2001) 370. White Paper on European transport policy for 2010: time to decide. Brussels.

European Commission (2001c). SEC (2001) 234. On public financing and charging practices in the Community sea port sector (on the basis of information provided by the Member States). Commission Staff Working Document. Brussels.

Heggie, I (1974). Charging for port facilities. Journal of Transport Economics and Policy. January.

Greene, W H (2000). Econometric analysis. Prentice Hall, London.

Griffiths, A and Wall, S (2000). Intermediate microeconomics – theory and applications. Prentice Hall, London.

Jansson, J O (1974). Intratariff cross-subsidization in liner shipping. Journal of Transport Economics and Policy. September.

Jansson, J O, Rydén, I (1979). Samhällsekonomisk analys för hamnar. EFI, Stockholm.

Jansson, J O, Shneerson, D (1982). Economics of scale of general cargo ships. Review of Economics and Statistics. May.

Jansson, J O, Shneerson, D (1982). Port economics. MIT Press, Cambridge Mass.

Jansson, J O, Shneerson, D (1987). Liner Shipping Economics. Chapman and Hall, London.

Jansson, J O, Lindberg, G (1997). Transport pricing principles in detail. Deliverable 2 – Appendix. **In** PETS – Pricing European Transport Systems. European Commission, Brussels.

Lieberman, M B (1994). Learning curves in the chemical processing industry. Rand Journal, 15.

Maddala, G S (2001). Introduction to Econometrics. John Wiley & Sons, West Sussex.

Norrköping Port and Stevedoring. Annual reports 1960-1999. Norrköping.

Phare (Final June 1997). Strategic Study of Ports and Maritime Transport in the Baltic Sea.

Pindyck, R S and Rubinfeld, D L (2001). Microeconomics. Prentice Hall, New Jersey.

Saaty, T L (1961). Elements of Queuing Theory. McGraw-Hill. New York.

Shneerson, D (1976). The structure of liner freight rates – a comparative route study. Journal of Transport Economics and Policy. January.

SOU 1995:112 (1995). Svensk sjöfart - Näring för framtiden. Fritzes, Stockholm.

SOU 1998:129 (1998). Svensk sjöfartsnäring – hot och möjligheter. Fritzes, Stockholm

Stopford, M (1997). Maritime Economics. Routledge, London.

Swedish Maritime Administration (1998). Environmental differentiated fairway and port dues. Sjöfartsverket, Norrköping.

UNCTAD (1973). Port pricing. Report by the UNCTAD Secretariat TD/B/C. United Nations. New York.

Walters, A A (1968). The Economics of Road User Charges. World Bank Occasional Paper. No 5.

Walters, A A (1976a). Marginal Cost Pricing in Ports. The Logistic and Transportation Review. Vol 12, No 3.

Walters, A A (1976b). The economics of road user charges. World Bank Occasional Papers. No 5.

Walters, A A (1977). Marginal cost pricing in ports. The Logistics and Transportation review. **12**, 3.