THE UNIVERSITY *of York*

CENTRE FOR HEALTH ECONOMICS
YORK HEALTH ECONOMICS CONSORTIUM
NHS CENTRE FOR REVIEWS & DISSEMINATION

Concentration and Choice in the Provision of Hospital Services

The Relationship Between Volume and the Scope of Activity and Hospital Costs

CRD REPORT 8 (Part II)

Concentration and Choice in the Provision of Hospital Services

The Relationship Between Volume and the Scope of Activity and Hospital Costs

Vassilis Aletras
NHS Centre for Reviews and Dissemination
CRD Report 8 (Part II)
January 1997

ISBN 1 900640 06 6

This report can be ordered from: Publications Office, NHS Centre for Reviews and Dissemination, University of York, York YO1 5DD. Telephone 01904 433648; Facsimile: 01904 433661: email: crdpub@york.ac.uk
Price £9.50

The NHS Centre for Reviews and Dissemination is funded by the NHS Executive and the Health Departments of Scotland, Wales and Northern Ireland; a contribution to the Centre is also made by the University of York. The views expressed in this publication are those of the authors and not necessarily those of the NHS Executive or the Health Departments of Scotland, Wales or Northern Ireland.

Printed by York Publishing Services Ltd.

NHS CENTRE FOR REVIEWS AND DISSEMINATION

The NHS Centre for Reviews and Dissemination (CRD) is a facility commissioned by the NHS Research and Development Division. Its aim is to identify and review the results of good quality health research and to disseminate actively the findings to key decision makers in the NHS and to consumers of health care services. In this way health care professionals and managers can ensure their practice reflects the best available research evidence.

The reviews will cover

- the effectiveness of care for particular conditions;
- the effectiveness of health technologies; and
- evidence on efficient methods of organising and delivering particular types of health care.

Further Information

If you would like further information please contact:

General Enquiries:	01904 433634
Information Service	01904 433707
Publications:	01904 433648
Fax:	01904 433661
Email:	revdis@york.ac.uk

Publications

CRD Reports

1	Which Way Forward for the Care of Critically Ill Children? (1995)	£7.50
2	Relationship Between Volume and Quality of Health Care: A Review of the Literature (1995)	£5.00
3	Review of the Research on the Effectiveness of Health Service Interventions to Reduce Variations in Health (1995)	£12.50
4	Undertaking Systematic Reviews of Research on Effectiveness. CRD Guidelines for those Carrying Out or Commissioning Reviews (1996)	£7.50
5	Ethnicity and Health: Reviews of Literature and Guidance for Purchasers in the Areas of Cardiovascular Disease, Mental Health and Haemoglobinopathies (1996)	£12.50
6	Making Cost-Effectiveness Information Accessible: The NHS Economic Evaluation Database Project. CRD Guidance for Reporting Critical Summariesof Economic Evaluations (1996)	£7.50
7	A Pilot Study of 'Informed Choice' Leaflets on Positions in Labour and Routine Ultrasound	£7.50

Effectiveness Matters

Vol.1	(1)	Aspirin and Myocardial Infarction (1995)
	(2)	Helicobacter Pylori and Peptic Ulcer (1995)
Vol.2	(1)	Influenza Vaccination and Older People (1996)

ACKNOWLEDGEMENTS

This paper is part of a report on '*Concentration and Choice in the Provision of Hospital Services*' which was prepared by a team of researchers directed by:

Brian Ferguson, Deputy Director, Centre for Health Economics, University of York
John Posnett, Director, York Health Economics Consortium, University of York
Trevor Sheldon, Director, NHS Centre for Reviews and Dissemination,University of York

We acknowledge the contribution of the NHS CRD Information Service in developing the search strategy for this study.

We also acknowledge the useful comments and support given by Andrew Jones of the Department of Economics and Related Studies, University of York.

****

Other titles in the *Concentration and Choice in the Provision of Hospital Services* series are:

Summary Report

I. *The Relationship Between Hospital Volume and Quality of Health Outcomes*

III. *The Relationship Between Concentration, Patient Accessibility and Utilisation of services*

TABLE OF CONTENTS

1. BACKGROUND

Health care costs are sharply rising over time exacerbating the need for cost containment by means of enhanced efficiency. An important aspect of efficiency refers to the optimal scale and scope of hospital production. The existence of economies of scale - cost savings from expanding the scale of production - must be empirically assessed if the optimal size of a health care facility is to be determined. Suppose that local demand can be served by one large general public hospital or two smaller ones. If economies of scale exist a government agency may choose the former option. In reality, of course, other policy parameters - such as the costs from travelling and delayed emergency treatment - should also be taken into account. Similarly, the presence of economies of scope between hospital services - cost savings from joint production - will define the sets of services that should be bundled together and produced within the same hospital unit. If, for instance, such savings exist for inpatient and outpatient care, an agency must presumably rather direct resources towards the development of hospitals which provide both services rather than operate separate specialised entities.

The use of information on economies extends to other policy decisions. Within an antitrust context, for example, the likely anticompetitive effects of hospital mergers or acquisitions - such as supracompetitive prices - can be weighed against the theoretical gains from exploiting scale and scope economies.

The concept of economies of scale describes the behaviour of long-run costs as the scale of production changes, the long-run being a period sufficiently long to permit *all* inputs (even the number of beds) to be variable. For a firm producing a single homogeneous product, economies (diseconomies) of scale are said to exist over a range of output if long-run average cost (AC) falls (rises) throughout the relevant range following an increase in the scale of production. If AC remains invariant, neither economies nor diseconomies operate. Neoclassical theory often assumes a U-shaped relationship between AC and facility size. As facility size (and production) increases AC is thought to decrease (economies), reach a minimum, and then increase (diseconomies).

Various sources of scale economies have been suggested.[1] First, a larger scale of operation permits greater opportunities for the division of labour and hence specialisation. Second, there are technological factors giving rise to scale advantages. One such is the existence of an initial "lump" of fixed costs which implies that the unit costs can be reduced with increased production. Another potential technical economy results when a firm can double the capacity of its buildings by less than doubling the construction costs. Third, there are economies on reserves of labour or materials resulting to a larger institution facing a variable demand for its services. A larger facility is required to keep a smaller proportion of its beds in reserve to meet an unexpected demand. Finally, there are pecuniary economies, that is quantity discounts and lower interest rates on borrowed capital, enjoyed by larger hospitals. Neoclassical theory also maintains that at some expansion point diseconomies set in - due to increased managerial inefficiencies and lack of communication - which will eventually bring about an increase in the unit cost. Alchian (1959), in contrast, predicts an L-shaped AC curve with no diseconomies ever occurring.

The preceding analysis is nevertheless inadequate for hospitals which are multi-product firms. New cost concepts have therefore been introduced. Ray (overall/global) economies of scale refer to the response of total cost to a proportional change in *all* output categories, keeping other cost determinants constant. However, it may be that at some expansion point, scale advantages exist for some services whilst not for others. Product-specific economies of scale refer to cost changes when production of a particular service is increased, keeping all other service levels constant.

Finally, the concept of economies of scope relates to the effect on costs of an increase in the scope of a firm's operations to additional product lines. They exist for two hospital services (e.g. inpatient and outpatient care) if their joint production within one

---

[1] See Hefty (1969), Feldstein (1983) and Long et al. (1985).

hospital is less costly than their separate production in two independent units. Economies of scope may result, for instance, from the avoidance of duplication of medical equipment, or from the existence of related clinical specialties on site. Diseconomies may also arise, for example, from an excessive use of expensive medical equipment, unavailable to some members of the medical staff prior to the bundling of the services.

Although the theoretical arguments are strongly in favour of the existence of economies of scale at least up to a production point, this needs to be empirically validated. Research on economies of scope is also needed to determine the optimal configuration of service bundles within hospitals. This review study aims to retrieve and evaluate the relevant research. It builds on existing reviews but differs in that it adopts the methodology of a systematic review from clinical epidemiology.[2] The expectation is that a systematic review will provide more valid conclusions based on the overall - rather than piecemeal - evidence via the use of appropriate quality criteria.

---

[2] See NHS Centre for Reviews and Dissemination (CRD) Guidelines for Undertaking Systematic Reviews of Research on Effectiveness (1996) and Goodman (1993)

2. METHODOLOGY OF THE SYSTEMATIC REVIEW

2.1 The Review Protocol

A protocol was written which defined clearly the questions that the review attempted to address and the methods used. It was meant to prevent bias so that neither the questions, nor the methods would be driven by a presumption about the likely findings. The current protocol focuses on hospital based studies and is shown in Appendix I.

2.2 Literature Searching and Study Retrieval

Conclusions drawn from a review may be biased if all relevant studies are not retrieved. In order to avoid such bias this review is systematic in its search strategy.

Studies considered as relevant are those empirical papers reporting empirical estimates on economies of scale and/or scope in hospitals, hospital wards, or specific hospital services (e.g. knee replacement), including findings on returns to scale or other indirect evidence. The period the review covers is 1967-1996. Studies prior to 1967 ignored the so-called case-mix problem which implies that their results are not reliable.[3] Subsequent studies attempting to by-pass the problem will nevertheless be briefly discussed. Relevant studies are econometric, statistical, and programming methodologies of any kind, in any language but excluding psychiatric or other long-term health care facilities (e.g. nursing homes), Health Maintenance Organisations (HMOs) and home health agencies. We simultaneously looked at theoretical and empirical papers not reporting evidence on economies but which could nevertheless be of use in the evaluation of the included studies.

---

[3] The case-mix problem may be thought of as the bias in the estimated measures of economies of scale and scope caused by uncontrolled differences in case-mix across hospitals. Comprehensive discussion of the problem can be found in section 3.3.1.

Ten electronic databases were searched: (1) BIDS (Bath Information Data Service) ISI (Institute for Scientific Information, Inc.) for 1981-1996, (2) Economic Literature Index (American Economic Association Database: searched on the DIALOG System) 1969-1996, (3) Medline Express (Copyright Medline: National Library of Medicine: searched on the OVID CD-Rom) 1980-1996, (4) ABI Inform (US database with published & unpublished work: searched on the Ohio-link, Ohio-State University) 1980-1996, (5) Health Planning and Administration (US National Library of Medicine: searched on the DIALOG) 1980-1996, (6) NTIS (National Technical Information Service: searched on the DIALOG System) 1980-1996, (7) Embase (Elsevier Sciences, BV Amsterdam, Netherlands: searched on the DIALOG) 1980-1996, (8) Dissertation Abstracts Database (UMI, Ann Arbor, Michigan: searched on the DIALOG) 1968-1996, (9) Economics Working Paper Archive provided by the Economics Department of Washington University (http:\econwpa.wustl.edu), and (10) DEA WWW bibliography provided by Portland State University (http:\ www.emp.pdx.edu). Printed bibliographies were used to retrieve earlier studies: (a) Fletcher, J.(ed.) *Economic Working Papers: A Bibliography, 1978-1991*, and (b) Blades, C. et al.(eds.) *The International Bibliography of Health Economics: A Comprehensive Annotated Guide to English Language Sources since 1914*, vol. 1, Harvester, 1986.

The Health Econometrics mail-base (health-econometrics@mailbase.ac.uk) as well as some CRD contacts have been approached in search of unpublished work. Retrieved articles were themselves scrutinised to trace new relevant published and unpublished references. Finally, recent issues (1995-1996) of three key journals (Applied Economics, Journal of Health Economics, Health Economics) were handsearched to identify articles which had not yet been indexed on the electronic database.

A CRD information system expert provided assistance on accessing the databases efficiently and on the overall adoption of an optimal search strategy. The actual search strategy employed is included in Appendix II.

2.3 Search Results

The search identified approximately 100 studies providing evidence on the existence of economies of scale and scope in the hospital setting.[4] The hospital has been employed as the unit of analysis in 36 *ad hoc* econometric cost studies, 22 econometric flexible cost studies, 6 *ad hoc* and flexible econometric production functions, 8 survival-type models and 5 Data Envelopment Analyses. There are also 13 studies which focus on a specific sub-set of hospital services (e.g. heart surgery) rather than using the hospital as the unit of analysis. Summary information on all these studies can be found in Tables 1-6. Ten studies which are oriented towards the realisation of cost-efficiencies in multi-hospital arrangements (e.g. mergers) were also identified.[5] Finally, there are 3 studies (econometric and statistical) which unsuccessfully attempted to overcome the well-known case-mix problem, and 1 queuing model.

2.4 Study Validity Assessment

Each study satisfying the inclusion criteria will be assessed according to a number of criteria. Some of them might be common across the different study designs under evaluation, e.g. correction for case-mix and quality of care differences across hospitals and over time. Others will necessarily differ according to the nature of each specific methodology. For instance, Data Envelopment Analysis which is a non-stochastic frontier technique, is more likely than econometric models to face severe problems with the choice of variables to be included in the model and errors in the data. Hence a criterion must be employed to assess the success with which different studies have dealt with these problems (e.g. conduct sensitivity analysis to different output specifications). The qualitative assessment criteria are summarised at the protocol included in Appendix I.

---

[4] A study which finds evidence using more than one methodology (e.g. cost and production models) or reports evidence on more than one area of interest (e.g. evidence of economies at the hospital level and at the system level) is double-counted.

[5] Taking into account the flexible model of Sinay (1994).

Data extraction sheets were used to record the relevant information - with respect to criteria and findings - from each study (Appendix III). Criteria for judging the validity of different studies were developed and a sensitivity analysis was conducted to see whether study quality affected the results. Finally, given that results may vary *across* study designs, the <u>relative</u> strengths and weaknesses of different methodologies (e.g. DEA versus econometrics) were considered to attempt to reach the final conclusion about the strength of research evidence for economies in hospitals.

3. *AD HOC* ECONOMETRIC COST STUDIES

Studies conducted mainly prior to the mid-1980s have been characterised as "*ad hoc*" primarily due to their restrictive functional forms, the indiscriminate use of any variable that was thought to be influencing costs and the exclusion of theoretically important structural variables (e.g. input prices). Information regarding their specific characteristics and findings is provided in Table 1.

The validity assessment criteria used include the unit of measurement of hospital output, choice of functional form, adjustment for heterogeneity of output (case-mix), derivation of long-run scale estimates, inclusion of input prices, treatment of uncertainty, adjustment for quality of care and choice of model variables.[6] These are discussed in turn.

3.1 Criterion 1

3.1.1 Unit of Measurement of Hospital Inpatient Output

Hospitals are multi-product firms producing numerous inpatient and outpatient "treatments" as well as teaching and research services.[7] It is often claimed that the output from treatment should be conceptually defined in terms of final outcomes. For practical reasons, however, health economists have used intermediate outputs.

The "hospitalised case", expressed either as an admission or discharge, and the "patient day" have been proposed as candidates.[8] The former measure is more defensible since

---

[6] The R^2 cannot be used as a measure of the relative quality of studies. One reason is that an apparently excellent fit might be the result of autocorrelation or multicollinearity. More generally, it is not true that a model achieving a higher R^2 includes more appropriate or relevant regressors than other models. Theoretical arguments can better serve as indicators of relative quality.

[7] The term "treatment" is defined here broadly so as to include services provided in terms of the diagnosis and management of an illness.

[8] See Butler (1995) for a discussion of the reasons why an "episode of treatment" is not an appropriate unit of output in hospital cost analysis.

the patient day can be seen more as a time input which, combined with the intensity of treatment, produces a treated case.[9] Alternatively, the case has been seen as a more satisfactory measure of output due to the anomalous behaviour of per day costs. For example, if a hospital allocates its resources more efficiently by shortening average lengths of stay, the average cost per day will increase because much of the treatment cost is loaded early on an episode. The erroneous conclusion would be that reducing length of stay is inflationary.[10] It is apparent that the two average cost measures might hence exhibit different behaviour. Feldstein (1967) and Butler (1995) report correlation coefficients of only 0.23 and 0.20 between them. An additional corroborating argument for using cost per case is that the hospitalised case can more readily be thought of as an exogenous variable facilitating valid estimations via single-equation cost models.

In this review, which focuses on acute hospitals, studies using cases are seen as more valued than cost per day models.

3.1.2 Sensitivity Analysis and Validity Assessment

The low correlation of average cost per day with the conceptually preferred average cost per case may partly explain any conflicting evidence on scale economies from cost per day studies. Carr and Feldstein (1967), Berry (1967, 1970, 1974) and Ault and Johnson (1979) document economies of scale at least up to certain output levels. Ingbar and Taylor (1968) report an inverted U-shaped AC curve, whereas Lave and Lave (1970a,b) find constant returns to scale. Unsurprisingly, cost per case and cost per day equations yield conflicting results in some studies. Jenkins (1980) and Robinson and Luft (1985) find that both equations indicate diseconomies of scale. In contrast, Feldstein and Schuttinga (1977), Sloan and Becker (1981), Sloan, Feldman and Steinwald (1983) and Becker and Sloan (1985) find evidence of constant returns or diseconomies in cost per case specifications but economies in cost per day. In light of the earlier arguments and the typically lower R^2 found in cost per day specifications

---

[9] See Butler (1995).

[10] See Frank (1988).

we place a greater validity weight on results from cost per case equations, *ceteris paribus*, in favour of constant returns or diseconomies.[11]

These results are reinforced by the constant returns documented by the cost per case studies of Evans (1971), Lave, Lave and Silverman (1972) and Zaretski (1977) or the diseconomies found by Evans and Walker (1972). Still the criterion by itself is not powerful enough to reconcile all findings. Culyer et al. (1978) and Bays (1980), in the equation excluding the price of physicians, for instance, find economies in their cost per case specifications.

3.2 Criterion 2

3.2.1 Choice of Functional Form

Neoclassical theory assumes that a firm's long-run objective is to minimise its total cost under the constraint imposed by its production function. The solution of this optimisation problem yields the firm's cost function, C = C(Q, p), which relates the minimised cost C, to outputs (Q) and input prices(p).[12]

A major drawback of traditional cost analysis is the adoption of restrictive functional forms for the cost relation. Researchers have employed the simple additive-linear, the quadratic and the logarithmic forms. But a maintained *ad hoc* form imposes restrictions on the underlying technology. The problem is then that the assumptions made about costs or ultimately the technological relations may not be valid, resulting in biased estimates.[13] The simple linear cost function is:

$$C_i = b_o + \sum_{j=i}^{m} b_j X_{ij} + u_i$$

---

[11] Note that R^2 can in some instances be used as a useful qualitative indicator. Here it is employed to discriminate between per day and per case equations.

[12] See Gravelle and Rees (1992).

[13] Chambers (1988) gives an example: a researcher's use of a cost function linear in prices implies an assumed dual Leontief technology. If, for example, the true unknown production function is Cobb-Douglas then the econometric results are suspect.

where C_i is the unit cost in hospital i, X_{ij} the various regressors, such as the case-mix and capacity (bed size) of hospital i, u_i the error term. A significant negative estimated coefficient on the capacity variable has been seen as indicating scale economies. This form implicitly assumes that the effects of various determinants on unit costs are linear and separable. This would mean that an extra day of patient care would increase unit cost by a fixed amount, irrespective of the level of capacity and utilisation, the mix of cases treated or the wage levels. This form assumes the absence of substitution possibilities between inputs, that there are either economies or diseconomies over different hospital sizes - but not both - and that economies of scope do not exist. If the separability hypothesis is not a characteristic of the hospital cost structure then misspecification will yield biased estimates.[14]

A quadratic specification allows for economies in certain size ranges and diseconomies in other. Finally, the Cobb-Douglas cost function imposes the *ad hoc* restriction of unitary elasticities of input substitution and prejudges the existence of an L-shaped AC curve and weak cost complementarities (precursor of scope economies).[15]

3.2.2 Sensitivity Analysis and Validity Assessment

Vitaliano (1987) compared a quadratic and a logarithmic model. The former indicates the presence of a shallow U-shaped AC curve, whereas the latter an L-shaped curve and scale economies. The quadratic specification was rejected as a misspecification by Ramsey's RESET test (regression specification error test). Finch and Christiansen (1981) also found a shallow U-shaped curve with a minimum at 113 beds from their quadratic equation and economies from the respective logarithmic equation. The R^2 is higher (0.84 as opposed to 0.75) in the Cobb-Douglas model thus reinforcing the finding that no diseconomies ever set in. Yet, other studies utilising the Cobb-Douglas

---

[14] Brown, Caves and Christensen (1979) show that the assumption of separability for a multi-product firm can greatly distort estimates of marginal costs and scale economies. It also precludes the calculation of a measure of economies of scope.

[15] A logarithmic cost function rules out *a priori* the possibility of a change towards scale diseconomies in higher levels of output.

functional form, namely those by Lave and Lave (1970a,b) and Pauly (1978), document constant returns implying that neither are there economies.

For a better understanding of the hospital technology *ad hoc a priori* restrictions should not be imposed, and as will be shown later more flexible models are more valid.[16]

3.3 Criterion 3

3.3.1 Adjustment for Output Heterogeneity

Economies of scale are measured by isolating in regression analyses the effect of the independent variable "scale" (e.g. admissions or number of beds) on costs. As with all observational studies (e.g. in epidemiology), differences between studies may reflect differential success at correcting for case-mix differences. Biased estimates of economies of scale result from cost functions if case-mix is not kept constant across hospitals or over time. An erroneous finding of diseconomies of scale might be observed in so far more complicated cases are admitted in larger hospitals and raise unit costs. The artefact is known as the "case-mix effect".

3.3.2 Sensitivity Analysis and Validity Assessment

Cost per day functions typically standardise output for service-mix differences by including regressors representing the scope and range of available facilities. Carr and Feldstein (1967), Berry (1967) and Francisco (1970) adopt this approach. The problem with these studies is that they do not account for differences in the level of utilisation of the available services and that (given the variation in case-mix and its complexity across hospitals) the input "available service" - even if fully utilised - will not produce the same output across hospitals. A preferred approach controls for the number of services actually provided (e.g. number of X-rays). Cohen (1967) employs a weighted average of 13 services provided, each weight being the service's relative average cost.

---

[16] See Hellinger (1975), Cowing and Holtmann (1983), Kemere (1992), Okunade (1993) and Gaynor and Anderson (1995).

However, there is a degree of tautology and autocorrelation caused by regressing costs on costs.[17] More generally, it is assumed that hospitals providing the same number of services will produce similar outputs.

Rather than standardising in terms of inputs Feldstein (1967) introduces the case-mix approach. To cope with multicollinearity he employs 9 "medically meaningful" case-mix proportions and constant returns are revealed.[18] But the use of broad speciality groupings can hardly be considered as homogeneous with respect to resource requirements. Robinson and Luft (1985) adopt the case-mix approach and find significant diseconomies.

Progress has been facilitated by the development of case-mix classification schemes - such as the International Classification of Diseases (ICD) and the Diagnostic Related Groups (DRGs).[19] To achieve parameter parsimony and collinearity reduction Evans (1971) employs factor analysis and groups ICD diagnoses into 10 composite regressors. Correcting additionally for age-sex differences results in constant or mildly decreasing returns. Feldstein and Schuttinga (1977) include 20 variables and find constant returns in a cost per case equation. Lave, Lave and Silverman (1972) use three techniques for combining 17 broad ICD categories into 5 composite variables which, along with other supplementary variables, yield constant returns. Jenkins (1980) employs 40 diagnostic proportions and finds diseconomies with an optimum of 100-300 beds.

The case-mix has also been measured by Evans and Walker (1972) in terms of a scalar measure with the use of information theory. Moderate scale diseconomies are found. Similar application by Barer (1982) also reveals diseconomies whereas that by Culyer

---

[17] See Berki (1972).

[18] Inserting case-mix in an average cost equation in the form of proportions reduces multicollinearity. Another rationale for the use of AC rather than total cost equations is that the latter suffers from heteroscedastic disturbances.

[19] Even these disaggregated groupings, however, do not achieve ideal homogeneity in terms of resource use since illness severity is not identical within groups. They also involve a huge number of categories making empirical estimation unfeasible.

et al. (1978) indicates economies, a result that may be due to the exclusion of age-sex variables or the focus of the latter authors on teaching output.

The greater the degree of aggregation the less the within-group homogeneity and explanatory power and the less confident we can be of estimates of economies of scale and scope. In fact, empirical evidence shows that detailed variables are able to explain more interhospital cost variation than scalar indices. Moreover, it seems that, at a given level of aggregation, case-mix variables outperform service-mix proxies.[20]

Some researchers have even seen the service-mix and case-mix approaches as complementary. Zaretski (1977) finds both service and case-mix variables to be significant in the same regression suggesting that costs are both a function of what hospitals are "geared up to produce" and what "is actually produced". Constant returns are found in a cost per case equation. Fottler and Rock (1974) and Finch and Christiansen (1981) adopt this mixed approach. The former find diseconomies, the latter economies in a Cobb-Douglas model. Lave and Lave (1970a,b) assumed that case-mix is constant over short periods of time and found constant returns.

Table 1 shows that studies which better adjust for case-mix, suggest that constant returns of diseconomies may exist for average and large hospitals.

3.4 Criterion 4

3.4.1 Derivation of True Long-run Scale Estimates

In a cross-sectional sample of hospitals it may not be the case that the fixed inputs are utilised at their minimum-cost levels.[21] The use of a long-run cost function will not yield true long-run economies estimates if the assumption of long-run cost minimisation is violated. Results may be biased. In such situations, an alternative way to derive true long-run effects requires the estimation of a short-run (variable) cost

---

[20] See Watts and Klastorin (1980).

[21] Time-series models on the other hand face the problem of adjusting for technological change and variations in efficiency over time.

function, $C^{var} = C^{var}$ (Q, p , K), where C^{var} is variable costs rather than total costs, Q denotes the output produced, K the fixed input (e.g. beds) and p the prices of the variable inputs (e.g. supplies). However, the derivation of long-run scale effects from such functions also requires an accurate measure of the price of capital which is not readily available in practice.[22] The short-run structure of costs can then be related to the long-run function by the use of the so-called envelop condition of economic theory. For this review, an important criterion is that a study should provide true long-run estimates in order to be characterised reliable.

3.4.2 Sensitivity Analysis and Validity Assessment

Due to regulatory constraints, adjustment towards the long-run equilibrium may be slow and the assumption of long-run cost minimisation therefore invalid. So, the constant returns to scale finding documented by Zaretski's (1977) and Pauly's (1978) long-run cost functions might be biased.

Most early studies in contrast - for instance, Carr and Feldstein (1967), Brown (1980), Bays (1980) - include two types of output variables to deal with the problem. A "flow" variable (e.g. number of cases) is included along with a "stock" variable (e.g. number of beds) as a measure of capacity.[23] The short-run cost structure is traced by examining the impact of changes in output to costs, keeping capacity constant. It is also claimed that the long-run structure - and hence economies of scale - is depicted by changes in cost resulting from changes in capacity, with output held constant. The problem is then that keeping output constant and varying capacity levels is not equivalent to the envelop condition dictated by economic theory.[24] The accuracy of the estimates of economies in all earlier studies is therefore in doubt.

---

[22] This point is further explained in Cowing, Holtmann and Powers (1983) and Aletras (1995).

[23] As Cowing, Holtmann and Powers (1983) note, an equivalent specification is to include a measure of output along with the rate of utilisation, that is, the ratio of output to capacity. The equivalence follows from the fact that only two variables from the triptych "output-capacity-utilisation" are independent.

[24] See Cowing, Holtmann and Powers (1983).

3.5 Criterion 5

3.5.1 Omission of Input Prices

Input prices play a prominent role in the theory of cost functions. Failure to include them in econometric specifications imposes the assumption of zero input substitution. Since empirical evidence supports substitution possibilities (between, say, nursing and medical staff) the omission of input prices may result in biased estimates of economies.[25]

3.5.2 Sensitivity Analysis and Validity Assessment

Many early studies do not ignore input prices and use some proxy wage rates (e.g. Pauly (1978)) or location dummies (e.g. Zaretski (1977), Vitaliano (1987)). Pauly (1978) and to a lesser extent Hornbrook and Monheit (1985) and Friedman and Pauly (1983) provide the most reliable results according to this criterion since they used disaggregated input price proxies. Two of these find constant returns whereas the poorer (service-mix approach is used) study by Friedman and Pauly (1983) indicates mildly increasing returns. But almost all *ad hoc* studies omit the price of the physician input since data are unavailable. This is so because in the US and Canada - where most of the research has been carried out - admitting physicians are not hospital employees. Bays (1980) argues that the admitting physician is an important idiosyncratic input in hospital production which has the authority to admit patients and direct the usage of other inputs. Omission of a significant variable from the production function will yield an over- or under-estimated measure of returns to scale, the bias depending on whether the excluded variable varies greater or less than proportionately with changes in the included regressors. Assuming cost-minimisation, the bias will then be reflected in the dual cost function. The author imputes a proxy measure for the value of admitting physician services. Including it in a cost regression turns an apparent finding of mildly

---

[25] Jensen and Morrisey (1986) find a 0.547 elasticity of substitution between medical staff and nurses implying that a 10% rise in the medical wages relative to nurses wages would bring about a 5.47% drop in the medical staff / nurses ratio. Conrad and Strauss (1983) find that nursing services are fairly substitutable for ancillary and general services. Okunade (1993) finds in contrast limited scope for substitutions at hospital pharmacies.

increasing returns to scale into diseconomies. The implication of this empirical result for other studies is that the finding of diseconomies or constant returns is reinforced.

3.6 Criterion 6

3.6.1 Treatment of Uncertainty

The need for hospital care is characterised by uncertainty with respect to the timing and severity of an illness and hence a stochastic demand for services. The early literature is rather silent on this matter. One way of visualising uncertainty is to imagine the hospital as providing excess capacity in order to meet unanticipated demand. If this "standby" service - produced in addition to expected patient care - is not included in the cost model then biased estimates will result.[26] The level of "standby" service (reservation quality) can be captured by a variable K derived from queuing theory as:

$$K = (\text{number of beds - average daily census}) / (\text{average daily census})^{1/2}$$

The equation shows that an equal proportionate increase in hospital size and average daily census brings about a rise in reservation quality implying the presence of economies of scale to be exploited in the presence of a stochastic demand.

3.6.2 Sensitivity Analysis and Validity Assessment

Friedman and Pauly (1983) account for uncertainty in their cost model by including the ratio of forecasted to actual admissions as an explanatory variable. Slight economies of scale are found in their, rather unsatisfactory (service-mix approach is used), model.[27] Mulligan (1987), on the other hand, employs a queuing model and finds scale economies due to uncertain demand to be very limited.

---

[26] The point is made by Cowing, Holtmann and Powers (1983).

[27] This study reproduces the Friedman and Pauly (1981) methodology.

3.7 Criterion 7

3.7.1 Adjustment for Differences in Quality of Care

The "quality effect" has a similar interpretation and importance as the case-mix effect already discussed. A finding of constant returns or diseconomies may be caused by unmeasured higher quality of care provided by larger hospitals. A bias towards diseconomies will occur in this case if offering greater levels of quality is in fact more costly. Hotel services are definitely costly. But as for other more important quality dimensions the relationship is not simple.[28] The cost-quality relation could be positive or negative, the implication being that it may not be easy to predict the direction and magnitude of the bias on the estimates of economies of scale, unless a study encompassing all relevant elements (all relevant quality dimensions, case-mix, scale etc.) is conducted. The varying success of different studies in capturing quality differences might explain some differences in their reported estimates of scale effects.

3.7.2 Sensitivity Analysis and Validity Assessment

Ad hoc models have not incorporated direct quality measures, such as risk-adjusted mortality, readmission indices and complication rates.[29] Crude proxies have been employed since quality is extremely difficult to measure. Cohen (1970) employs a proxy dummy variable for hospital affiliation with medical schools. This treatment, though widely adopted ever since, has been criticised for its inability to really capture quality differences. His more recent work is comparable with his previous attempt in 1967 in all respects except for the inclusion of the proxy. A higher optimum of 540-555 beds is indicated when the proxy is incorporated. Fottler and Rock (1974) use instead the work-force skill level and indices of facilities and service capability. Their results seem to be suggestive of diseconomies. In any case, we cannot draw any meaningful conclusions from such evidence. The problem of quality adjustment has not been treated in any acceptable way even in the more recent models.

---

[28] See Fleming (1991).

[29] See Fleming (1991) for an empirical study analysing the relationship between quality and cost.

3.8 Criterion 8

3.8.1 Choice of Model Variables

Neoclassical economic theory asserts that profit-maximising firms pursue minimisation of the cost of producing a given output. The derived minimum cost from the optimisation problem is a function of output levels and input prices and therefore only these variables should constitute the elements of a (*structural*) cost function. Evans (1971), however, argues that viewing not-for-profit hospitals as cost-minimising entities is not appropriate since hospital management may become involved in managerial slack. It follows that a *behavioural* rather than a structural process of the determination of costs should be sought and that other variables should be considered. Estimated scale effects might depend on the set of variables chosen by the researcher.

3.8.2 Sensitivity Analysis and Validity Assessment

Most early empirical models are behavioural including any variables thought to influence costs. Pauly (1978) found that the number and composition of medical staff significantly determine costs. Presumably this inclusion may be justified given that physicians are not directly paid by the hospital so that it is their characteristics and not their cost (price) that should be included. Constant returns are found. As another example, Hornbrook and Monheit (1985) employ a profit/non-profit dummy for the hospital status and the share of total patient days used by Medicare/Medicaid patients to capture the effect of cost-based reimbursement on costs.

Unfortunately, in light of the lack of an accepted theoretical model of hospital behaviour it is not possible to conduct sensitivity analysis and differentially weight models using a particular behavioural variable.[30] The criterion is put forward since it is possible that the discrepancies in study results could be due partly to the different variables employed rather than previously identified factors.

---

[30] See Pauly and Redisch (1973), Newhouse (1970), Lee (1971) and Harris (1977) for different theoretical models of hospital behaviour.

3.9 Overview and Synthesis of the Evidence

The empirical findings on economies of scale can now be discussed in light of the assessment criteria developed. The 36 studies are presented in Table 1, which shows that ad hoc models are characterised by an adjusted R2 ranging from 0.50 to 0.99. We exclude from further consideration studies based on: a) the patient day as the unit of output (criterion 1),[31] b) the service-mix approach alone (criterion 3), or c) the "no adjustment approach" over time or across hospitals of Lave and Lave, and Wagstaff (his stochastic model).[32] The resulting studies to be examined are based on either the case-mix approach or a combined case-mix/service-mix adjustment, the service-mix presumably proxying the quality of output or technological sophistication.

Almost all of the remaining 18 studies indicate constant returns or diseconomies.[33] To these studies we can apply the finding that highly disaggregated variables perform better than less disaggregated ones or scalar indices (criterion 3). With the exception of Butler's (1995) mixed results, the 5 studies which better adjust for case-mix indicate constant returns or diseconomies.[34] Finally, Pauly (1978) corrects for case-mix differences by means of an index but also for non-physician input prices by inserting 4 variables (criterion 5) and simultaneously employs the superior Cobb-Douglas functional form (criterion 2) and reports constant returns.

To rule out the possibility that the findings are due to large-sized hospitals included in the samples we turn to Table 1 (column 2). Feldstein and Schuttinga (1977) and Pauly (1978) report a mean bed size of 180 beds and find constant returns. Evans and Walker (1972) utilise a sample of hospitals of various sizes including very small (less than 25

---

[31] If cost "per case" and "per day" models are both examined in a study the "per case" model is retained.

[32] In Wagstaff's (1989b) study only the stochastic pooled-data model makes this assumption and is rejected at this point.

[33] The remaining 18 studies are: Feldstein (1967), Evans (1971), Evans and Walker (1972), Lave, Lave and Silverman (1972), Feldstein and Schuttinga (1977), Zaretski (1977), Pauly (1978), Culyer et al. (1978), Jenkins (1980), Bays (1980), Sloan and Becker (1981), Barer (1982), Sloan, Feldman and Steinwald (1983), Becker and Sloan (1985), Hornbrook and Monheit (1985), Robinson and Luft (1985), Wagstaff (1989b: the non-frontier model) and Butler (1995).

[34] These are the studies by Feldstein and Schuttinga (1977), Zaretski (1977), Jenkins (1980), Robinson and Luft (1985) and Butler (1995).

beds) and very large (more than 1000) and find moderate diseconomies. Economies are only present for hospitals with less than 100 beds. Bays (1980) reports evidence of constant returns for rather small sample hospitals (mean bed size is 124).

The exclusion of the physician input price by most of these reinforces the view that economies of scale are absent. Such a conclusion however should be treated with caution. First, economies of scale due to stochastic demand for hospital services are not accounted for in these studies (criterion 6). Second, there might still be uncontrolled case-mix variations (criterion 3). Third, the quality effect has not been dealt with adequately (criterion 7). Fourth, it might not be the case that computed estimates are true long-run economies (criterion 4). Fifth, there may be biases from the use of overly restrictive functional forms (criterion 2). Sixth, we are not certain as to whether - and which if any - of the behavioural variables are indeed relevant (criterion 8).

Some authors have even attempted to by-pass the problem of correcting for structural cost determinants other than the scale of operation. Examples are the studies by Chernichovsky and Zmora (1986) and Schaafsma (1986). These models are mispecified and their results will not be discussed here.

4. FLEXIBLE ECONOMETRIC COST STUDIES

Flexible functional forms are an attempt to avoid the risk of mispecification of the unknown production function by specifying estimable cost relationships that impose fewer *a priori* restrictions. Being second-order Taylor-series expansions they serve as an approximation to any arbitrary differentiable function. Thus, restrictive assumptions like separability or homotheticity can become testable hypotheses.[35] The general formula of flexible functional forms is:

$$F(C) = a_0 + \sum_{i=1}^{m} a_i f_i(x_i) + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=i}^{m} b_{ij} f_i(x_i) f_j(x_j) \qquad (B_{ij}=B_{ji} \; \forall \; i,j)$$

where C is total hospital costs and $x_i, \ldots, x_n$ are explanatory variables (e.g. hospital outputs, input prices). The most common flexible form is the transcendental logarithmic (translog) with $F(C) = \ln C$ and $f_i(x_i) = \ln x_i$. Since however this form cannot deal with zero output levels a version of this, called the generalised (or hybrid) translog, is often used instead.[36]

There are advantages and disadvantages in using flexible functional forms for multiproduct analysis.[37] On the one hand, they typically satisfy - at the empirically estimated region - the regularity conditions associated with a well-behaved cost (and ultimately production) function. Moreover, they allow for flexibility in that - by not restricting first and second order derivatives - they do not prejudge the existence or

---

[35] Homotheticity implies that the cost minimising mix of inputs is not affected by the volumes or mix of outputs, a rather strong assumption. Indeed, most studies tested and rejected restrictive models (e.g. separable). See Burns (1982), Conrad and Strauss (1983), Eakin and Kniesner (1988), Kemere (1992) and Scuffham, Devlin and Jaforullah (1996).

[36] A "hybrid" translog is meant to deal with zero output levels. This is typically pursued by employing the Box-Cox transformation. Alternatively, one can use an ordinary translog and either exclude hospitals that do not produce certain outputs or use a very small positive number to replace the zero values. Finally, the actual values rather than the logs can be employed. Note also that another flexible form less often used is the quadratic with $F(C) = C$ and $f_i(x_i) = x_i$.

[37] See Lee (1987) and Ahern (1988) for a discussion.

degree of economies of scale and scope.[38] In contrast to earlier models, flexible cost functions can also accommodate important multi-product concepts, namely product-specific economies of scale and economies of scope.

On the other hand, increased flexibility of the function is obtained at the cost of greatly reduced parsimony. With just 5 inputs and outputs 55 parameters result in an ordinary translog model. This is a major limitation in health care since controlling for the extreme output heterogeneity between hospitals requires a large number of case-mix variables and other outputs (e.g. teaching).[39] Another weakness is the potential inability of these functions to provide accurate estimates away from the point at which the function is being approximated - typically the sample means.

The criteria that will be used to evaluate the validity of the empirical evidence from flexible cost functions are now presented.[40] Summary information on these models is found in Table 2. Note that flexible models have a high R^2 (greater than 0.90 in cross-sectional data).[41]

---

[38] A function is said to have "regular" properties if it is non-negative, real valued, non-decreasing in outputs, strictly positive for non-zero output, linearly homogeneous in input prices and concave in factor prices.

[39] Breyer (1987) points out the problem and suggests a number of compromising assumptions to be placed on flexible models in order to reduce the number of parameters. The validity of these assumptions however can be questioned.

[40] Note that a separate section concerning the choice of flexible functional form will not be included here. This is primarily due to the finding of Lee (1987) that the translog, generalised translog and CES translog showed a similar performance on purely statistical grounds and regarding the success with which they satisfied the regularity conditions (to be discussed shortly). In any case, the various functional forms are assessed indirectly in that we will check whether a particular model did exhibit theoretical desirable properties (regularity conditions). In addition, no separate section will be included for the choice of variables. Admittedly, some models have used behavioural variables i.e. variables other than outputs and input prices. Yet their omission has not been shown to affect estimates of economies. For instance, Custer and Willke (1991) found medical staff characteristics to be important cost determinants in hospital production. But the qualitative results from estimates of economies of scale and scope did not change when staff variables were included. Fournier and Mitchell (1992) found that competition and ownership are important cost determinants but that the inclusion of these behavioural variables did not affect the measures of economies. And in any case, there is no consensus regarding the behavioural variables that should be included in a cost function. Finally, quality of care has again not been dealt with adequately and will not be discussed in detail.

[41] Scott and Parkin (1995) attempted to estimate a flexible cost function with UK data. They found a very high R^2 (0.991) and a highly significant intercept. They thus raised the issue of data quality. It might be the case that data used in estimations reflect accounting practices (identities) rather than economic behaviour of hospitals. This problem does seem to exist in some US studies. Therefore, survival analyses - despite their limitations - might be useful since they are not based on recorded costs.

4.1 Criterion 1

4.1.1 Calculation of True Long-run Economies of Scale

There are two broad methods available for the estimation of economies of scale.[42] First, a long-run cost function can be estimated by regressing total hospital costs on outputs and input prices, with the price of capital either included or assumed constant. This approach is valid if hospitals are employing the cost-minimising levels of all inputs. However, it is generally thought that, due to regulatory restrictions, hospitals cannot quickly adjust all their input levels once a change in output levels or prices has occurred. Given that the hospitals are employing the optimal level of adjustable variable inputs but not the cost-minimising level of fixed inputs, economies of scale estimates will be biased.

Alternatively, short-run (variable) cost functions which regress variable costs on outputs, prices of the variable inputs, and the level of the fixed factors (e.g. beds) can be used. If this strategy is chosen, there are then three options.

If an accurate measure of the price of capital was available then an estimate of *true long-run* economies would be derived. Yet no such price is available. In its absence, some economists have estimated *short-run* economies from the parameters of the variable cost equation. This does not conform to an economist's definition of economies of scale, since it indicates potential savings from an adjustment in output levels, when the fixed factor is not allowed to change. A short-run estimate might indicate some short-run savings but might move the hospital further away from long-run cost-minimisation.

For simplicity, this situation is illustrated below for a firm producing a single output. A firm located at point B is operating under short-run economies (falling portion of the short-run AC curve) and can reduce its unit cost if it increases its output - via a better

---

[42] A discussion of these issues can be found in Braeutigam and Daughety (1983) and Vita (1990).

utilisation of the variable inputs - from Q_2 to Q_3 in order to reach the minimum of the short-run unit cost curve (C_3 at point C).

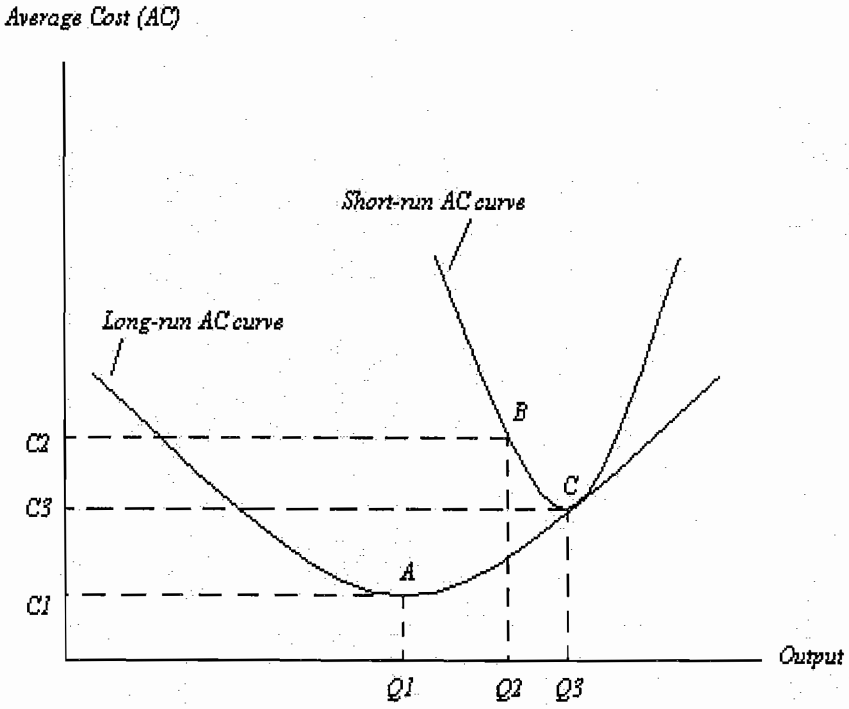

Diagram 1: Short-run vs. long-run cost minimisation in a single-product firm

In the long-run, however, the level of the fixed factor (plant size) can be altered, thus allowing a movement along the long-run average cost curve. A firm located at point C operates under long-run diseconomies (rising portion of the long-run AC curve) and can reduce its output from Q_3 to Q_1 in order to incur a unit cost of C_1 (point A).

A better pragmatic approach is to use the observed (actual) level of the fixed input instead of the optimal one (which would require knowledge of the price of the fixed input). *Long-run* economies are measured at the actual point of operation rather than the optimal one. If the current regulatory regime creates incentives to hospitals such that they will not be at the optimal point of operation anyway (i.e. adjustment towards the efficient expansion path is slow), then this measure should be used. Policy makers however, having a long-run policy perspective, might want to change the regulatory regime in order to eliminate non-optimal levels of the fixed factor. If the industry's rate of adjustment is likely to be fairly rapid, then the above measure of economies might mislead policy makers. Nevertheless, the measure can still be estimated and make inferences on true economies of scale based on *a priori* beliefs about the over- or under-employment of the fixed factor relative to its optimum.

Apparently conflicting evidence on economies may thus be partly attributable to differences in the measures used by different researchers. We think of long-run measures as more informative and theoretically consistent than short-run estimates of economies of scale.

4.1.2 Sensitivity Analysis and Validity Assessment

On inspection of Table 2 (columns 4 and 7) we check whether studies employing similar measures of economies of scale find consistent or conflicting evidence.

Short-run measures of economies of scale have been derived from estimated variable cost functions by many researchers even in recent years.[43] Most of them suggest that hospitals in the specific samples operate under short-run ray economies and should increase output - at the given level of capital investment - in order to fully exploit efficiency savings in the short-run. Some studies in contrast suggest short-run diseconomies. These results are not informative about long-run effects and the optimal hospital size.[44]

Estimates of overall *long-run* economies of scale have been obtained from the estimated parameters of long-run equations in 7, studies.[45] The studies by Conrad and

---

[43] As shown in Table 2, these are the hospital studies by Roddy (1980), Cowing and Holtmann (1983), Lee (1987), Ahern (1988), Pangilinan (1991), Kemere (1992), Fournier and Mitchell (1992), Ablett (1993), Collins (1994), Gruca and Nath (1994), Scuffham, Devlin and Jaforullah (1996) and the Okunade (1993) work in pharmacy production.

[44] The measures are presented in three equivalent ways. As the sum of the cost-output elasticities, as their inverse, or as unity minus the sum of cost output elasticities. So, one should be cautious when comparing the magnitude of measures of economies across studies.

[45] These are the studies by Burns (1982), Conrad and Strauss (1983), Banker, Conrad and Strauss (1986), Eakin and Kniesner (1988), Kemere (1992) , Banks (1993) and Wagstaff and Lopez (1996). Overall economies are computed with respect to the sum of cost-output elasticities.

Strauss (1983), Banker, Conrad and Strauss (1986) and Kemere (1992) suggest that - at the means - overall constant returns to scale operate.[46] Burns (1982), Banks (1993), and Wagstaff and Lopez (1996) document instead (unexploited) economies of scale for the average hospital. Finally, Eakin and Kniesner (1988) report significant diseconomies at the sample means.

There are also *long-run* estimates computed from short-run cost functions in 9 studies, by employing the actual level of the fixed input.[47] These treat the long-run equilibrium assumption as testable and show that it is rejected (excessive investment in capital).[48] Vita's (1990) long-run measure indicates diseconomies for the average hospital. He also shows that, given an over-investment in capital in the hospital industry, *true* economies of scale will be larger than the estimated measure. Slight diseconomies were also suggested when Cowing and Holtmann's (1983) short-run measure was converted to its long-run counterpart. Pangilinan (1991) and Collins (1994) also document slight diseconomies, whereas Kemere (1992) and Scuffham, Devlin and Jaforullah (1996) neither economies nor diseconomies. Sinay (1994) finds economies for one sample of hospitals prior to the merger but diseconomies for another. Gaynor and Anderson (1995) find economies. Note that Kemere (1992) finds overall long-run economies or diseconomies to be absent in both long-run and short-run equations. Roddy (1980) erroneously claims to find constant returns in the long-run by testing and accepting homogeneity of degree one within a translog short-run function.[49]

---

[46] The study by Banker, Conrad and Strauss (1986) employs the same data set and variables as the Conrad and Strauss (1983) work. It is merely its frontier version.

[47] These are the studies by Cowing and Holtmann (1983, as computed subsequently by Vita), Vita (1990), Pangilinan (1991), Kemere (1992), Collins (1994), Gaynor and Anderson (1995), Scuffham, Devlin and Jaforullah (1996) and Sinay (1994).

[48] Evidence come from Cowing and Holtmann (1983), Ahern (1988), Kemere (1992), Fournier and Mitchell (1992) and Ablett (1993).

[49] This last study might in fact be indicating constant returns in the short-run. The author employs some lemmas to show that the test of homogeneity of degree 1 in the variable cost function automatically implies constant returns in a long-run sense. However, it seems that this treatment is a misapplication of the lemmas mentioned. Specifically, the paper shows the following sequence of results: a) if a multi-output transformation function is homogeneous of degree (h.o.d.) r, then the corresponding cost function will be h.o.d. 1/r, b) if the cost function is h.o.d. 1/r, then the variable cost function will be "almost" h.o.d. (1, 1/r, 1/r). However, these results do not prove the following sequence which seems to be required to move from an estimated variable cost function to inferences about the homogeneity of the long-run function: if the variable cost function is "almost" h.o.d. (1, 1/r, 1/r), then the cost function is h.o.d. 1/r, and in turn, the transformation function is h.o.d. r.

Thus, there seem to be more evidence against the existence of unexploited economies for the typical hospital once short-run estimates are disregarded. Finally, the 6 short- and long-run studies examining product-specific economies find conflicting evidence.[50] For instance, Custer and Willke (1991) find economies for outpatient visits whereas Kemere (1992) diseconomies at the means.[51]

4.2 Criterion 2

4.2.1 Unit of Measurement of Hospital Inpatient Output

As has been already explained in section 3.1.1, the case has been seen as a superior measure. A sensitivity analysis is carried out to check for consistency in the evidence from the more reliable types of studies.

4.2.2 Sensitivity Analysis and Validity Assessment

We limit the sensitivity analysis only to research work (short- and long-run functions) providing long-run overall economies of scale estimates. Among these 14 studies, 7 employ the patient day as the unit of measurement.[52] Although results are mixed, almost all of the more reliable 7 studies employing the cost per case - often accompanied by average length of stay - did not detect economies of scale for the average hospital. This is illustrated in Figure 1 (these studies are depicted using capital letters to distinguish them from the less reliable studies).[53]

---

[50] These are the studies by Grannemann, Brown and Pauly (1986), Ahern (1988), Custer and Willke (1991), Kemere (1992), Fournier and Mitchell (1992) and Wagstaff and Lopez (1996).

[51] Note here that the former sample includes much larger hospitals and one might have expected the reverse results.

[52] The 7 "patient day" studies are those by Burns (1982), Cowing and Holtmann (1983), Conrad and Strauss (1983), Banker, Conrad and Strauss (1986), Banks (1993), Collins (1994) and Sinay (1994). Note that the study by Roddy (1980) was shown to report short-run rather than long-run estimates of economies and is thus excluded from the analysis at this point. It nevertheless also shares the limitation of using the patient day as the unit of measurement.

[53] These are the studies by Eakin and Kniesner (1988), Vita (1990), Pangilinan (1991), Kemere (1992), Gaynor and Anderson (1995), Scuffham, Devlin and Jaforullah (1996) and Wagstaff and Lopez (1996).

Economies	*Banks (1993)* *Burns (1982)* *GAYNOR & ANDERSON (1995)* *WAGSTAFF & LOPEZ (1996)*
Constant Returns	*SCUFFHAM ET AL. (1996)* *Banker et al. (1986)* *KEMERE (1992)* *Conrad & Strauss (1983)*
Diseconomies	*Cowing & Holtmann (1983)* *VITA (1990)* *Collins (1994)* *EAKIN & KNIESNER (1988)* *PANGILINAN (1991)*
Mixed Results	*Sinay (1994)*

Figure 1: The criterion of the unit of measurement and the findings on long-run economies

The exceptions are the studies by Gaynor and Anderson (1995) and Wagstaff and Lopez (1996) which both find economies. Thus, the first two criteria are not able to reconcile all the findings. And it still remains to check what the average hospital in these studies is, because it may be that differences in the evidence reported in different studies can be reconciled once the sample means are considered.

4.3 Criterion 3

4.3.1 Treatment of Uncertainty

As explained in section 3.6.1, larger hospitals might exploit additional economies of scale in light of the stochastic demand for their services. Recently, Gaynor and Anderson (1995) provided estimates of economies which include this element.

4.3.2 Sensitivity Analysis and Validity Assessment

The theory of cost and production is reformulated to account for the uncertain demand facing a hospital and the short-run translog cost function incorporates demand distribution parameters in addition to outputs and input prices. The study uses a national sample of 5000 hospitals for the years 1983-87 and the traditional model is rejected in favour of the reformulated one. The latter shows significant long-run unexploited overall economies for the average hospital. The difference in the results *vis a vis* those of other comparable studies cannot be attributed to differences in the volumes or sizes in different samples.[54] Thus it could be the case that the other studies would have found economies or at least constant returns, if they have adjusted for reservation quality (standby service). To explore this proposition further we must nevertheless look at the remaining criteria. At this point we simply note the use of national data (increasing the possibility that environmental differences are not adequately controlled for) which may have influenced the estimates.

4.4 Criterion 4

4.4.1 Regulatory Environment and Cost-minimisation [55]

The existing studies have used samples that include hospitals financed on either a retrospective or prospective basis.[56] There are reasons to believe that studies employing prospectively-paid hospitals might yield somehow more valid results. Typical econometric models of cost behaviour assume that hospitals are cost-minimising entities. In the hospital industry this is a dubious assumption.[57] In a regime where cost-based reimbursement prevails, hospitals and physicians have incentives towards excessive resource consumption. Hospitals might hence be operating on an inflated cost curve rather than on the true economic curve.[58] Estimates on economies of scale may be biased if, for example, allocative inefficiency is correlated with the scale of hospital production and independently affects costs. That is, a finding of

---

[54] Scuffham, Devlin and Jaforullah (1996) use a sample of hospitals which are smaller and produce a lower volume of output and find constant returns. In the Gaynor and Anderson (1995) study the mean number of admissions, outpatient visits and beds is 2097, 20118 and 125, whereas the respective figures in the Scuffham, Devlin and Jaforullah (1996) work are 6092, 40139 and 170.

[55] The following discussion is focused on the US environment where most empirical work has been conducted.

[56] Prospective reimbursement of costs has been selectively implemented in some States since the late 1970s. The practice has been extended since 1983 and was expected to be globally completed by 1988. See Ahern (1988).

[57] See Lee (1971), Newhouse (1970), Ellis (1993).

[58] These points are raised by Ahern (1988).

diseconomies might be an artefact if larger hospitals are more inefficient. Similar arguments hold for economies of scope. It might be reasonable to believe that samples of prospectively-reimbursed hospitals will be closer to cost-minimisation since they are reimbursed by a fixed rate per admitted case, which depends on the diagnosis.[59] They thus have an incentive to minimise costs and short as well as long-run costs are expected to shift downward toward the true cost curves.[60]

Nevertheless, inefficiency might not be completely eliminated in prospective reimbursement regimes. For instance, Certificate of Need (CON) regulation in the US created incentives for over-employment of capital since *inter alia* it allows incumbents to expand systematically in order to pre-empt entry from new hospitals. Hospitals may in fact be at a non-cost-minimising long-run equilibrium, in the pre-prospective payment system (PPS) period, due to monopolistic elements in the industry. After the implementation of PPS, some incentives for excess capacity remained since CON regulation persisted, but may have been reduced because the opportunity cost for capital accumulation seems to have been increased. If hospitals have adjusted to their new equilibrium, then the measures of scale economies derived from post-PPS cost functions will yield more valid estimates.

4.4.2 Sensitivity Analysis and Validity Assessment

Eakin (1991) has found that there is a greater percentage of allocative inefficiency in larger hospitals. Ahern (1988) tested the long-run equilibrium assumption and rejected it in favour of excess capacity in both pre-PPS and post-PPS equations. Yet over-employment of capital was less in the latter case. These results indicate that there might indeed be a lower bias in the scale estimates from post-PPS studies. Eakin and Kniesner (1988) however showed that this bias is negligible by utilising a long-run translog function which treats cost-minimisation as a testable restriction. Allocative inefficiency is 4-5% of costs but estimates on economies are almost identical for the

---

[59] Note that some studies used samples referring to the post-1983 era in which PPS might not have fully been implemented. Nevertheless, the anticipation of full implementation might have forced hospitals to adjust.

[60] See Ahern (1988) and Gruca and Nath (1994).

non-minimum and the nested minimum cost equation. The use of a classical minimum cost function is therefore justified for the study of economies, which in turn implies that the existing literature does not seem to suffer from any significant bias caused by a violation of the cost-minimisation hypothesis.[61] In favour of the hypothesis that inefficiency does not affect estimates on economies seems to be the work of Wagstaff and Lopez (1996). Their model is the frontier version of the Grannemann et al. cubic polynomial equation. Product-specific economies are computed for the average hospital at the cost frontier rather than at the hospital's actual costs. The estimates are found to be in line with those of the earlier version, indicating economies for emergency visits but diseconomies for ambulatory care.

The study by Eakin and Kniesner (1988) suggests that the importance of criterion 4 (ie. prospective payment of hospitals) for assessing the validity of individual studies cannot be over-emphasised. In any case, 6 post-PPS studies have been identified which report (overall long-run) economies of scale estimates. These involve the samples by Vita (1990), Pangilinan (1991), Kemere (1992), Collins (1994), Sinay (1994) and Gaynor and Anderson (1995).[62] Most studies indicate overall diseconomies, one constant returns. The study by Sinay (1994) finds mixed results, whereas that by Gaynor and Anderson (1995) economies.

However, the work by Collins (1994) and Sinay (1994) relies on the patient day and is therefore less reliable. So, at this point, criterion 4 only suggests that the relative superiority, in terms of validity, of the work by Vita (1990), Pangilinan (1991) and Kemere (1992), compared to these studies, is slightly reinforced. Yet once again the study by Gaynor and Anderson (1995) also meets the criterion under consideration and reports conflicting evidence that may be due (according to the criteria used so far) to the incorporation of uncertainty.

---

[61] Note that the study by Eakin and Kniesner (1988) employs a sample of US hospitals in 1975-76, thus presumably retrospectively reimbursed.

[62] The study by Vita (1990) employs California data from 1983. Kemere (1992) verifies that this sample is a post-PPS one.

4.5 Criterion 5

4.5.1 Adjustment for Differences in Input Prices

Omission of input prices can be justified if it is legitimate to assume that these are constant across hospitals or that there are no substitution possibilities between inputs in hospital production.[63] The latter assumption has been rejected by a number of researchers.[64] The former is justified - according to Ablett (1993) - in Belgium. In studies conducted in the United States however, input prices must not only be included, but also in a disaggregated manner if biases are to be avoided.

4.5.2 Sensitivity Analysis and Validity Assessment

Most studies incorporate several input price variables and pursue disaggregation. The greater the latter is, the lesser the bias may be in estimates of economies. Cowing and Holtmann (1983) use 5 variables for labour prices, namely nursing, auxiliary, professional, administrative, general.[65] They find disaggregated individual price variables to be statistically significant, the implication being that functions using a single price variable are mispecified. In this respect, the work of Custer and Willke (1991) and Wagstaff and Lopez (1996), reporting evidence on product-specific economies, as well as the studies by Gaynor and Anderson (1995) and Wagstaff and Lopez (1996), suggesting the existence of overall economies, are suspect and might have given erroneous results. The same holds to a lesser extent for the findings of Sinay (1994) on mergers. Nevertheless, the problem of aggregation bias remains even for the better studies.

Another interesting point is the inclusion of a proxy for the physician input price by Eakin and Kniesner (1988). The finding of diseconomies is in line with the evidence from Bays (1980) mentioned earlier. Yet it should be noted that the role of physicians

---

[63] This point is made inter alia by Cowing, Holtmann and Powers (1983).

[64] See for instance Cowing and Holtmann (1983).

[65] Studies employing a long-run cost function also incorporate a proxy for the price of capital or assume it constant.

in hospital production is not well understood. Some authors have thus included the level of the physician input or even medical staff characteristics into the cost function since the cost of physicians is not incurred by the hospital but rather by third-party payers or the patients.[66]

4.6 Criterion 6

4.6.1 Adjustment for Output Heterogeneity

Different hospitals provide different services. As discussed in section 3.3.1, the success with which studies controlled for case-mix is a main factor that can account for any discrepancies in the findings of the existing literature.

4.6.2 Sensitivity Analysis and Validity Assessment

Flexible functional forms are not parsimonious in parameters, implying that the adjustment for case-mix is bound to be poorer than in *ad hoc* study designs. Unsurprisingly, these models typically use about 2-5 output categories, sometimes complemented by case-mix indices or variables thought to be correlated with case-mix. The most prominent work in this respect was done by Lee (1987) who employed 8 aggregate DRG variables.[67] Unfortunately, the study provides only short-run estimates, suffers from multicollinearity and violates the regularity conditions (to be discussed shortly).

Cowing and Holtmann (1983) used 5 outputs, namely medical-surgical, maternity, paediatrics, other inpatient and emergency room, but with no case-mix index included. Their long-run estimate indicates slight overall diseconomies of scale. Diseconomies are also found by Eakin and Kniesner (1988) who instead used 4 output variables and different variables for medical and surgical conditions. Kemere (1992) employed 4 outputs but these where defined as paediatric, adult, geriatric inpatients and outpatient

---

[66] See Custer and Willke (1991), Cowing and Holtmann (1983) and Grannemann, Brown and Pauly (1986).

[67] Despite its deficiencies, the DRG classification scheme is an advancement in the sense that it has been designed in order to achieve homogeneity within groups with respect to resource requirements for treatment.

visits. He also incorporated variables thought to be correlated with case-mix, such as availability of CAT scanner and length of stay. Constant returns are found. Conrad and Strauss (1983) employed only 3 output variables (child, non-Medicare, Medicare) and no case-mix index and found constant returns.

A different school of thought maintains that there are two ways by which a hospital can produce more days of care.[68] They can either increase admissions holding average length of stay constant, or increase average length of stay keeping admissions constant. Thus, variables for both cases (e.g. admissions) *and* average length of stay should be included for each service category. Pangilinan (1991) used this approach for five types of care. Her function - which also incorporates a case-mix index - yields overall slight diseconomies. A similar earlier attempt by Vita (1990) found stronger evidence of diseconomies. Scuffham, Devlin and Jaforullah (1996) employed instead only 3 variables for each category (cases, length of stay) and no case-mix index and found constant returns.

Gaynor and Anderson (1995) document overall economies, a finding that could be attributable - according to our previous criteria - to the incorporation of uncertainty or the poor adjustment for input price differences. Another feature of this study is that it tests and finds admissions to be endogenous and hence employs instrumental variables. All other authors have assumed away any biases in the estimates from endogeneity by arguing that hospitals have limited discretion in influencing the level of admissions.[69] Nevertheless, the authors employ only 2 outputs and a case-mix index so that the study is not as successful in adjusting for output heterogeneity as is, say, the Pangilinan (1991) work which documented diseconomies.[70] Banks (1993) incorporates 4 service outputs and a facilities index proxying case-mix. The author finds economies. One distinguishing feature of his cost function is the incorporation of two socio-economic

---

[68] See for instance Vita (1990).

[69] However, the magnitude of the bias in the estimates has not been measured. Thus we only point out the fact that there is an additional source of bias not well explored yet.

[70] Kemere (1992) in fact showed that failure to capture the multi-product nature of hospital production (i.e. including a single output variable along with the case-mix proxies) gave the erroneous finding of economies. Note also that the same deficiency holds for Roddy (1980).

indices which are thought of as affecting the service provision of hospitals. The study nevertheless relies on the use of the patient day. Grannemann, Brown and Pauly (1986) include a rich set of variables describing case-mix and find product-specific economies of scale for the emergency department. However, in order to do so, they sacrifice flexibility by estimating a homothetic cost function and hence making invalid assumptions which biased their results.[71] Similar arguments about flexibility hold for the work of Wagstaff and Lopez (1996) who *de facto* assumed constancy of all input prices.

Note that the high level of aggregation in flexible studies casts doubt on the robustness of economies of scale estimates. Another effect of aggregation is that economies of scope computed between aggregate composite output variables might conceal the true production relationships.

4.7 Criterion 7

4.7.1 Regular Behaviour of Estimated Cost Functions

A cost function reflects a well-behaved production technology for medical care and rational economic behaviour if it satisfies the so-called regularity conditions.[72] That is, the reliability of estimated measures of economies of scale and scope depends on whether the fitted regression obeys some conditions defined by economic theory.[73] Typically, flexible functional forms - such as the translog, generalised translog or CES translog - guarantee *a priori* that one of these conditions (linear homogeneity in input prices) is met by imposing restrictions on the model parameters. Yet it should also be

---

[71] Kemere (1992) showed that his homothetic model indicated overall economies. However, it was strongly rejected in favour of the more general translog which instead showed constant returns.

[72] The regularity conditions are: 1) The cost function, $C(Q, p)$, is continuous and twice differentiable in outputs (Q) and input prices (p). This condition is guaranteed by the way in which all flexible cost models are defined, 2) It is also linearly homogeneous in input prices so that $C(Q, kp) = kC(Q, p)$. This condition is also typically imposed by means of restrictions placed on the model parameters, 3) The costs must be positive for positive levels of outputs and input prices, i.e. $C(Q, p) > 0$ for all $Q > 0$ and $p > 0$, 4) Marginal costs must be positive for all output levels, i.e. $C / Q \quad 0$ for all Q, 5) The shape of the cost function is such that it exhibits concavity with respect to input prices, 6) Shephard's lemma applies and the input shares are positive. See Eakin and Kniesner (1988), Gravelle and Rees (1992), Roddy (1980). Note also that Roller (1990a) refers to such a cost function as being "proper".

[73] See Lee (1987).

the case that some other properties are shown to be preserved *ex post*, i.e. by the empirical estimation. For instance, we should witness positive fitted costs in the various output levels and positive marginal costs for all outputs.[74]

As long as the cost function is well-behaved at the point it is approximated, namely the sample means, (e.g. no negative first-order output coefficients due to collinearity or aggregation) the estimates of flexible models can accurately depict hospital production technology at that point.[75] This in turn implies that they can be used to assess the degree of *overall scale economies for the average hospital.* It also means that the measurement of economies of scope can be pursued by testing for *weak cost complementarities at the means.* It should be noted however, that the absence of weak cost complementarities does not necessarily mean that economies are absent. It simply means that the given study cannot support their existence.[76] This is certainly a limitation in applied work. In any case, the point is that the above (admittedly less informative) local measures can be trusted if the regularity conditions are met at the sample means.

In contrast, the same author argues that global measures of economies will not in general be reliable because the regularity conditions are very likely to be violated away from the sample means. They are only valid for small changes in the output levels of an average hospital. Thus, global measures (such as *economies of scope* and *product-specific economies of scale* which are computed using measures of incremental costs, as well as *overall scale economies computed away from the means*) may misguide policy makers about the consequences of large incremental changes (e.g. cost impact of mergers).[77]

---

[74] Note that a researcher cannot readily impose other regularity conditions before the estimation since the additional parametric restrictions on the estimates entail a lack of flexibility in the functional form in the sense of prejudging the behaviour of the measures of economies of scale and scope.

[75] As Vita (1990) notes, "flexibility is a local property" allowing a cost function to represent the hospital's technology only at or near the mean values of the variables, without any restrictions on substitution and scale elasticities.

[76] As Butler (1995) explains, a cost function might exhibit economies of scope and yet there might be no cost complementarity i.e. the marginal cost of each output can still be independent of the output level of the other.

[77] The criticism has been primarily focused on the translog models since studies have shown that it failed to meet the regularity conditions. See Roller (1990b). Some authors have proposed the use of a flexible quadratic

This review does not rule out the possibility that some estimated translog functions may in fact be sufficiently regular away from the means.[78] Hence, in studies providing global measures of economies of scale or scope we examine whether cost functions have been shown to be regular *away from the means*. We may rely on these more highly. Moreover, a function might not be regular even at the means due to, say, collinearity problems. Therefore, studies reporting local measures are checked for the regularity of the cost function *at the means*.

4.7.2 Sensitivity Analysis and Validity Assessment

We first assess the validity of studies estimating the more informative global measures. As can be seen in Table 2 (column 7), Grannemann, Brown and Pauly (1986), Ahern (1988), Custer and Willke (1991), Kemere (1992), Fournier and Mitchell (1992) and Wagstaff and Lopez (1996) report conflicting evidence on *product-specific economies* of scale. Since none of these studies has examined whether the cost functions were globally well-behaved, their empirical results are very likely to be invalid.

Some authors have also attempted to describe the behaviour of *overall scale economies* at various output levels.[79] A study satisfying the conditions away from the means, namely Eakin and Kniesner (1988), suggests that smaller hospitals exhibit increasing and larger hospitals decreasing returns to scale. In contrast, the evidence by Conrad and Strauss (1983) and Banker, Conrad and Strauss (1986) indicating constant returns across hospitals of different sizes (or better volume of services) cannot be trusted since the functions were found to be irregular away from the means.[80]

---

cost function with properness imposed at certain regions of output levels. Nevertheless, some flexibility of the functional form must be lost. See Roller (1990a).

[78] Wales (1977) has found that translog models may in fact provide reasonable estimates if regularity conditions are satisfied for most of the sample observations.

[79] Roddy (1980) suggests that constant returns to scale operate for short-term hospitals in the long-run. The conclusion is reached by means of a global test of the restricted homogeneous of degree one model, nested within the translog. The author verifies that the function is regular for all observations, with the exception of two negative marginal costs. Hence, according to the criterion under consideration the derivation of a global measure is justified. However, as we mentioned, Roddy (1980) concludes globally constant long-run economies by testing the short-run cost function for homogeneity of degree one.

[80] Vita (1990) examined the functions estimated by Cowing and Holtmann (1983) and Conrad and Strauss (1983). He found negative marginal costs when these were evaluated near zero for an output, keeping all other regressors at their sample means. Moreover, the fitted costs were often not realistic.

Finally, measures of *economies of scope* evaluated globally are found in Grannemann, Brown and Pauly (1986), Lee (1987), Custer and Willke (1991), Fournier and Mitchell (1992) Ablett (1993) and Wagstaff and Lopez (1996). The first author examined each observation for concavity in input prices, arguing that this is the regularity condition most often violated in practice. It was not satisfied in all three flexible models (translog, hybrid translog, CES translog). The other authors did not show whether their functions were regular away from the means. In fact, some were irregular even at the sample means. The most extreme example is that by Ablett (1993) who found negative marginal costs. Thus, according to our criterion we should better turn to the less demanding local measures of economies of scope.[81] Unfortunately, the few authors who found their functions sufficiently regular for the sample range did not calculate global measures which would have been more reliable.

We turn to the local measures of (long-run only) *overall economies of scale*. Eakin and Kniesner (1988) checked all the regularity conditions and found that these are satisfied at the means and in 86% of the observations. Significant diseconomies were documented. Kemere (1992) suggests that all regularity conditions are satisfied at the means of the variables. Some even at each data point.[82] Constant returns are found. Another study which is regular at the means and finds constant returns is that by Scuffham, Devlin and Jaforullah (1996). In contrast, Burns (1982) reports negative marginal costs for new-born output even at the means. She finds unexploited economies for the typical hospital but the measure might be overstated given that one negative cost-output elasticity (and marginal costs) is found. Banks (1993) also finds economies and suggests that the estimated parameters indicate theoretical consistency. The assumption of constancy of prices also undermines the robustness of the estimate of ray economies of scale found by Wagstaff and Lopez (1996). Yet the quadratic flexible form employed is irregular in that it does not impose linear homogeneity in

---

[81] Collins (1994) instead provides weaker evidence on the properness of her cost function. The coefficients of the squared output measures are mixed in sign. According to Roller (1990b) this indicates that a translog function *may* not be degenerate in models using 3 or more outputs so that they *may* be appropriate for the estimation of valid global measures. We find this treatment rather dubious.

[82] These are positive fitted costs and positive predicted input cost shares.

prices.[83] Finally, we mention the model by Gaynor and Anderson (1995) which incorporates uncertainty so that duality does not hold. They find positive marginal costs and significant economies at the means.

Regarding local measures of *weak cost complementarities* we note that the studies by Pangilinan (1991), Vita (1990), Banks (1993), Collins (1994) and Scuffham, Devlin and Jaforullah (1996) find no significant weak cost complementarities at the levels of aggregation used. This does not rule out the existence of long-run economies of scope. Simply we cannot prove their existence. Gruca and Nath (1994) found economies only between general medical (acute) care and obstetrics. Sinay (1994) showed that in one merger episode he examined there were economies prior to the merger between acute and subacute care, providing reasons for the consolidations. He did not find any economies in the second episode, further suggesting that economies may be more likely where there is excess capacity.

4.8 Overview and Synthesis of the Evidence

After excluding the studies which compute short-run measures (criterion 1), 14 studies remain which report evidence on overall (long-run) economies of scale.[84] The evidence come from both short-run and long-run cost functions, the characteristics of which are summarised in Table 2. Kemere (1992) found that both types of functions gave similar long-run measures of economies. Nevertheless, short-run functions are more informative in that the direction of bias with respect to true economies of scale has been assessed. Most studies found that the average hospital in the samples used operates under diseconomies or constant returns. Exceptions are the attempts by Burns (1982), Banks (1993), Sinay (1994), Gaynor and Anderson (1995) and Wagstaff and Lopez (1996) which document unexploited economies or report mixed evidence.

---

[83] Linear homogeneity means that if prices of the inputs are, say, tripled then costs will also be tripled. Lee (1987) argues that since the quadratic function does not satisfy this condition, it should not be used for multi-product analysis.

[84] Excluding of course the study by Roddy (1980) for reasons already discussed.

Only the studies by Eakin and Kniesner (1988), Vita (1990), Pangilinan (1991), Kemere (1992), Gaynor and Anderson (1995) Scuffham, Devlin and Jaforullah (1996) and Wagstaff and Lopez (1996) employ the hospitalised case as the unit of measurement of hospital output (criterion 2). Hence only the studies by Gaynor and Anderson (1995) and Wagstaff and Lopez (1996) seem to yield evidence of economies, for the average hospital, that are not suspect according to this criterion. The former incorporated demand uncertainty (criterion 3), the implication being that its exclusion in other studies may account for their inability to find economies. The finding however of unexploited economies could well be due to the poor adjustment for input prices and outputs in that study (criteria 5 and 6). The latter argument is corroborated by the absence of significant economies due to the stochastic demand witnessed by Mulligan (1987). Interestingly, the inability to control for price differences also casts doubt on the validity of the estimates derived by Wagstaff and Lopez (1996). The studies by Burns (1982) and Banks (1993) which document economies can also be criticised on other grounds. Their models fail to meet the regularity conditions (criterion 7) even at the means of the variables. In contrast these conditions are satisfied at the means and even away from the means in the studies by Eakin and Kniesner (1988) and Kemere (1992) which report large diseconomies and constant returns respectively. This work also performs satisfactory with respect to the cost-minimisation hypothesis discussed previously (criterion 4). Moreover, Eakin and Kniesner (1988) incorporate a proxy for the physician price, the implication being that Kemere's (1992) model could also be indicative of diseconomies if that price was not omitted (criterion 5).

The question then is what the average hospital is. Column 2 in table 2 reports the sample mean number of beds for each study. Scuffham, Devlin and Jaforullah (1996) find constant returns for an average hospital of only 125 beds.[85] We argued earlier that the results of this study might be relatively reliable. One point however not made so far refers to the hospitals included in the sample. This particular study incorporates very different hospital types, ranging from general hospitals to maternity and psychiatric. It may not be legitimate to do this since the underlying production structures might be

---

[85] Unfortunately, some studies do not report descriptive statistics.

different. Other studies have also lumped together diverse hospital types (e.g. for- or not-for-profit, teaching or non-teaching) but at least confined themselves to general hospitals. The study by Eakin and Kniesner (1988) does not report its means whereas that by Kemere (1992) documents constant returns for a hospital with a size of about 300 beds and a volume of some 12500 inpatient discharges and 17200 outpatient visits. This work also confines the sample to a single state in order to avoid biases from uncontrolled differences in the regulatory environment. Given an over-investment in capital in the hospital industry the *true* optimum sought in a regulatory regime consistent with long-run cost minimisation will be higher.

Other studies which also perform quite satisfactorily with respect to our evaluation criteria found diseconomies for smaller average hospitals. Vita (1990), for instance, documents slight diseconomies for an average hospital of 180 beds (about 7800 inpatient discharges and 30300 outpatient visits).[86] This might mean that economies are exhausted at even lower size levels than the study by Kemere (1992) indicates. Despite the inability of these studies to pinpoint an exact optimum hospital size there is a consensus of the more reliable studies that if any economies exist they are quickly exhausted. That is, they might be present only for small hospitals with less than 100-200 beds.[87]

Regarding the global measures of product-specific economies of scale and economies of scope we argue that these cannot be trusted and might mislead policy makers since the cost functions were not shown to be regular away from the sample means (criterion 7). Local measures of weak cost complementarities in general did not support the existence of economies of scope for the average hospital in the long-run at the levels of output aggregation used. This does not rule out the existence of long-run economies of scope. Simply we cannot prove their existence. Gruca and Nath (1994) found economies only between general medical (acute) and obstetrics care. Sinay (1994)

---

[86] It should be noted however that this study has some ad hoc elements (see Table 2).

[87] Note that it would also be interesting to investigate whether the estimates of economies differ across the various hospital types (e.g. teaching, non-teaching). This is not possible since studies have combined several types of hospitals in their samples in order to obtain the required sample sizes for the empirical estimation. Another limitation of the existing literature is thus revealed since it may not be appropriate to pool such observations into a single sample.

showed that in one merger episode there were economies prior to the merger between acute and subacute care, providing reasons for the consolidations. He did not find any economies in a second episode, further suggesting that economies may be more likely where there is excess capacity. One limitation nevertheless of these two latter studies is the use of the patient day as the unit of measurement of hospital output (criterion 2).

Note that economies of scope were also not detected by Rozec (1988) who constructed triplets of costs each consisting of the costs of a hospital without a psychiatric unit, $c(y_1,0)$, a psychiatric hospital, $c(0,y_2)$, and a hospital with a psychiatric unit, $c(y_1,y_2)$. Hospitals were matched for ownership, length of stay, location and size. The pairs $c(y_1,0) + c(0,y_2)$ and $c(y_1,y_2)$ were compared and the null of no economies or diseconomies was not rejected. However, the absence of scope effects may have been due to the poor matching (e.g. for case-mix).

5. ECONOMETRIC PRODUCTION FUNCTION STUDIES

Given that hospitals are cost-minimising entities, duality theory means that the same information on economies of scale can be retrieved from either the cost function or the underlying production function. The latter describes the relationship between inputs and outputs in physical units. The problems of definition and measurement of output (including its quality dimension) and inputs and of appropriate model specification are unfortunately dual as well.

5.1 Criteria for Validity Assessment

The choice between cost and production functions has been based on mainly statistical criteria related to endogeneity. There are only a few reported production studies presumably because it is generally believed that it is more logical to assume that hospital output and input prices are exogenous (than inputs may be).[88] Another reason was that it is less clear how the multi-product nature of hospital production can be properly modelled within a production function. A reason put forward for supporting the use of production models is that physician input might be more readily incorporated in its physical units rather than in monetary terms since in US the physicians are not paid directly by the hospital.[89] Production studies, however, have employed prices as proxies for all real inputs or at best omitted some inputs (e.g. supplies) to avoid expressing them in monetary units. There might also be a more severe problem with collinearity between inputs. For these reasons, we may cautiously rely slightly more on cost as opposed to production studies. Regarding the relative quality of individual production studies (summarised in Table 3), we employ similar

---

[88] Kemere (1992) suggests that cost functions were mainly chosen not only for computational convenience (regarding the calculation of elasticities of input substitution) but also because it has been shown that, if output is exogenous, estimating a production function is not appropriate since the fitted relationship between inputs and outputs will not describe the production. In fact, Gaynor and Anderson (1995) apply the Hausman-Wu test and show that admissions are endogenous but outpatient visits are not.

[89] Note that short-run cost functions can treat physicians as a fixed (in the short-run) input and incorporate it in its physical units. However, not all would agree that this treatment is appropriate.

criteria to the ones used for the assessment of cost models, namely the choice of functional form, unit of measurement of inpatient output, appropriate adjustment for output, quality, input heterogeneity, theoretical consistency of estimated parameters and goodness-of-fit, endogeneity of variables.

5.2 Sensitivity Analysis and Validity Assessment

All production studies have high R^2 values (0.60-0.99). Feldstein (1967) employs a Cobb-Douglas, a mixed Leontief-Cobb-Douglas (assumes potential substitution between some inputs) and a somewhat more general *ad hoc* production model. Output is either standardised in the LHS by weighting nine case-mix proportions by their average cost or in the RHS by the inclusion of nine variables. Some of the four inputs included are measured in value terms. All three equations indicate mildly decreasing returns to scale in the production of English general hospitals. Another important finding is that estimation methods accounting for possible endogeneity of variable inputs yield similar results with the OLS equations. A UK study by Lavers and Whynes (1978) on *maternity* hospitals does not adjust for output since this is regarded as homogeneous. A Cobb-Douglas and a flexible log-quadratic are used and costs - including medical pay - replace the unavailable physical input quantities. Mildly decreasing returns to scale are once again observed.

The Cobb-Douglas model by Brown (1980) found increasing returns to scale for a sample of Newfoundland's *cottage* hospitals. The authors measure inputs in physical units and adjust crudely for output by weighting services with hypothesised relative resource requirements. On the output side, however the day rather than the case is adopted as the unit of measurement. Moreover, the sample consists of very small hospitals (mean bed size is 29). A sample of 142 New York State hospitals over 1981-1987, for which a fixed-effect Cobb-Douglas models is estimated, is used by Pangilinan (1991). Significantly increasing returns to scale are said to exist. However, a single labour input is employed and the physician input is omitted. Unsurprisingly perhaps, a Ramsey test indicates mispecification.

More reliable estimates have been provided by two flexible models.[90] Van Montfort (1981) employs a translog functional form. Using weighted admissions as the dependent variable he finds constant returns to scale and positive marginal products for all inputs. Jensen and Morrisey (1986) estimate a production function and find it to be well-behaved in the sense that marginal products for all inputs are positive and decreasing. We have used the estimates of their translog model and the sample mean values of the inputs to compute the output elasticities.[91] Their sums for the non-teaching and teaching sub-samples are 0.858 and 0.952, both below unity indicating decreasing returns for the average hospital.

Despite the apparent consensus that returns to scale are at best constant, caution is warranted due to the case-mix problem being treated less satisfactorily than in cost studies and the multicollinearity between inputs that seems to be severe. Moreover, quality differences (e.g. standby capacity that can be seen as an additional output) are ignored. Finally most of the production studies do not report the mean levels of hospital size or volume. In any case the consistency of the results with those of cost analyses is encouraging.

---

[90] Hellinger (1975) found that the Cobb-Douglas model nested within a flexible production function was rejected.

[91] The output elasticity of a factor i is: $[\, f(x) / x_i \,] \times [x_i / f(x)]$, where f(x) is the production function (in our case the translog).

6. DATA ENVELOPMENT ANALYSES [92]

The mathematical programming technique - also referred to as Data Envelopment Analysis - is rooted on the Pareto optimality principle.[93] A hospital will be said to be *relatively* efficient if in comparison to other hospitals there is no proof that it utilises any of its inputs inefficiently.

The technique constructs empirically, from the observed input/output relations of existing hospitals, what has been known as a "best practice frontier", which consists of efficient hospitals having the highest total factor productivity in the sample. The relative inefficiency of the remaining ones is given by their position relative to the frontier.[94] The technique is non-parametric and does not invoke any particular functional form for the true unknown production function, thus imposing less strict assumptions.[95] DEA can readily handle multiple inputs and outputs and has therefore been seen as a useful tool for assessing pure technical or allocative inefficiencies in hospitals. Yet DEA models until 1983 imposed constant returns to scale (CRS). Subsequently however the so called Variable Returns to Scale (VRS) models emerged which allow us to identify whether a particular hospital exhibits increasing, decreasing or constant returns. Banker (1984) also proposed the calculation of the most

---

[92] The textbooks by Norman and Stoker (1991), Fare, Grosskopf and Lovell (1985), Ganley and Cubbin (1992) and Charnes et al. (1994) offer insight on DEA models.

[93] This principle states the following. A firm is said to be 100% efficient if: (a) it cannot increase the production of any particular output without either increasing the quantity of one or more of its inputs, or reducing the quantity of some other outputs, (b) it cannot reduce the quantity of the inputs used without either reducing the quantity of some of the outputs produced, or increasing the quantity of at least one of the other inputs. In the real world however, the minimum quantity of inputs required to produce a given output is not known.

[94] In its simplest form, DEA maximises the ratio of the weighted sum of outputs to the weighted sum of inputs and yields a single measure describing the overall performance of a hospital. The weights are chosen so that the efficiency of the hospital under evaluation - given by the ratio - is maximised. Two imposed constraints require that the weights should be positive and that the optimal weights should be such that no one hospital's efficiency exceeds unity. The latter constraint makes sure that efficiency will be between zero and unity. The weights are not known *a priori* but rather computed objectively by DEA itself such that it gives the benefit of doubt to each hospital while calculating its efficiency score.

[95] The main assumptions are the convexity of the production possibility set and minimal extrapolation from observed data.

productive scale size (mpss) of hospitals, i.e. the point of hospital production at which decreasing returns have not yet started to operate.[96]

6.1 Criteria for Validity Assessment

Several criteria presented previously for the evaluation of econometric studies can also be applied here. These refer to the unit of measurement of hospital output, the adjustment for differences in outputs (e.g. case-mix, teaching), inputs, quality of care, and reservation quality. However, the very nature of DEA (being a deterministic frontier approach) necessitates the construction of an additional criterion which relates to the success with which an individual study remedied problems regarding the choice of variables and errors in the data. This in turn implies that specification, measurement, and sampling errors must somehow be dealt with.

DEA avoids the problem of specifying a particular functional form for technology, yet specification error may result since the researcher has still to decide which variables to include. No goodness-of-fit tests are available in DEA. Given the different output/input aggregation schemes a researcher can employ, we wish to assess whether DEA results are robust or variable-specific. Moreover, hospital analyses are heavily based on the use of proxies and data which are far from accurate. The resulting measurement error - not captured by an error structure - may have very serious effects on the estimates of scale efficiency given that DEA is a boundary method.[97]

Several suggestions have been made in the literature in order to remedy the problems including *inter alia* stochastic DEA (Banker (1989)), or sensitivity analysis (Nunamaker (1985), Valdmanis (1992), and Grosskopf and Valdmanis (1987)). The latter solution entails subjecting the model to different sets of variables and specifications to check whether findings are robust or dependent on the variables

---

[96] A problem with Banker's (1984) model is that under some circumstances the solution of the DEA program may yield multiple mpss. Zhu and Shen (1995) nevertheless suggest that these circumstances rarely appear in practice and, in any case, offer a solution to the problem.

[97] If the error occurs for a hospital found erroneously by DEA to be efficient the very frontier - on which all computations of efficiency scores are based - is affected. See Banker, Das and Datar (1989).

chosen. If results do change significantly then outliers should be sought and eliminated. However, sensitivity analysis is not a panacea since not all possible specifications can readily be considered.

6.2 Sensitivity Analysis and Validity Assessment

The synthesis of the existing evidence is based on Table 4. There is a general agreement between studies that hospitals with less than 200 and more than 620 beds are scale inefficient, the reasons for inefficiency being, respectively, increasing and decreasing returns. However, as illustrated in Figure 2 the evidence, at least on appearance, is conflicting regarding the precise position of the optimum. The horizontal axis indicates the possible bed size ranges, whereas the other horizontal lines and the asterisk (*) show the ranges of optimal sizes documented by each study.

Banker, Conrad and Strauss (1986) and Byrnes and Valdmanis (1994) calculate a mean mpss of 220-260 beds, which does not seem to be in wide disagreement with the position of the optimum that may be inferred from the work of Valdmanis (1992). In the latter, the mean scale efficiency is found to be 0.97-1.00 for public hospitals which have a mean bed size of 350, whereas 0.92-0.97 for private not-for-profit ones whose mean size is 428.

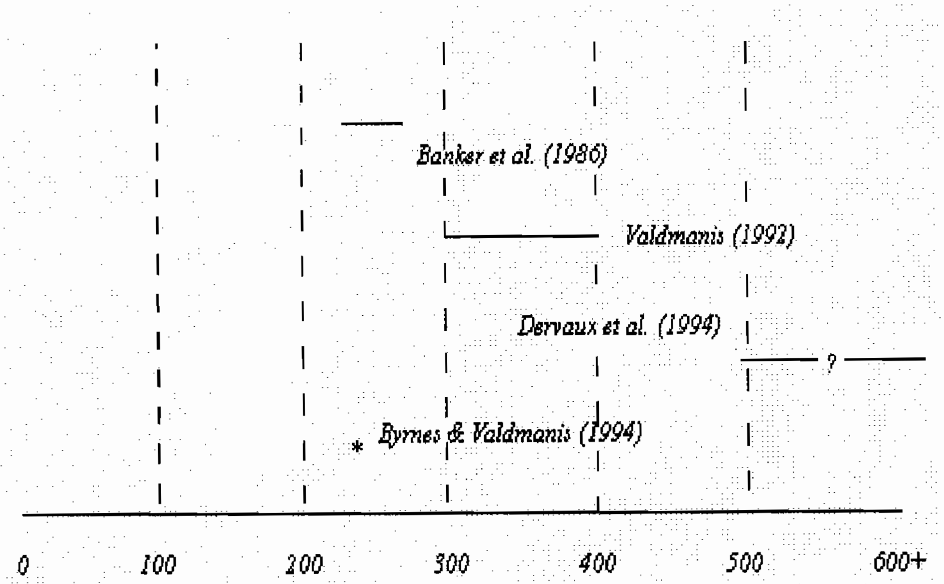

Figure 2: Most productive scale size (mpss) of hospitals

Maindiratta (1990) examines not only whether input savings might be realised by a hospital, given its observed task, but also whether additional savings could result if the task itself was to be optimally apportioned to a number of smaller hospitals. The findings suggest that decreasing returns set in very gradually so that a hospital must be a lot bigger than its mpss before it pays to apportion its task to smaller units. The largest-size efficient scale exceeds the mpss in some selected hospitals by a factor of 1.55-1.81. The consensus however breaks down once the French study by Dervaux et al. (1994) is considered, suggesting that scale efficiency is achieved at 500-620 beds.

Regarding the unit of measurement, the hospitalised case is more defensible than the patient day. We nevertheless find that this choice does not seem to affect the results. Quality of care has not been controlled for adequately by means of outcome indicators and the measurement of hotel-type services that may be valued by patients. Nevertheless, some adjustment has been made in three of them. These studies have limited their samples to a particular ownership type, (Dervaux et al., Byrnes and Valdmanis) or have run different DEA programs for different ownership types (Valdmanis).[98] In fact, Valdmanis (1992), and Byrnes and Valdmanis (1994) also excluded small and rural, or teaching hospitals respectively in order to control further for environmental differences. The other two studies (Banker, Conrad and Strauss, Maindiratta) do not report the types of hospitals included. In addition, all US studies employ data from a single State to control for regulatory differences.

Some of the above restrictions in the selected samples might have provided some adjustment for differences in case-mix. But the studies have also directly controlled - to varying degrees - for output heterogeneity. Two research attempts (Banker, Conrad and Strauss, Byrnes and Valdmanis) relied on the use of 2 to 3 output variables defined either in terms of patients' age or type of treatment. Their findings are similar. A little higher optimum might be suggested by Valdmanis (1992) who defines outputs according to patient age or type of treatment in different models. The researcher adopts somewhat richer specifications consisting of 3 to 5 output categories including

---

[98] Sherman (1984) found that hospitals with a similar organisational form produce similar types of care.

outpatient variables in some of them. The optimum is well below 400 beds and is insensitive to slight changes in the input/output variables used.

The French study by Dervaux et al. (1994) also experiments with different models. Two of these involve 6 outputs, some of which are meant to capture average- and long-term services (e.g. psychiatric). The levels of aggregation in short- versus average- and long-term services are slightly altered in these two models, placing greater emphasis on the former or the latter accordingly. These two specifications yield an optimum of 500-520. When using a third specification which incorporates additional activity variables (case-mix complexity index, services consumed per day of care, average length of stay), the optimum size increases to 620.

In any case, we might place a lower validity weight on this study since its findings are not followed by a thorough sensitivity analysis as is the case in the Valdmanis (1992) work. Thus, its results could be due to errors in the data.

Summing up, there is consistent evidence that the optimum is located in the 220-620 bed region. Evidence from DEA simply indicates that small and very large hospitals may be suboptimal. Moreover, we solve the apparent discrepancy in the determination of the optimum itself in favour of a 220-400 bed size range, keeping in mind the Maindiratta (1990) argument that decreasing returns may set in very gradually so that a hospital may have to be a lot bigger than its mpss before it pays to apportion its task to smaller units.

Nevertheless, the results of DEA studies should be interpreted with caution. First, differences in quality of care (and presumably case-mix) are inadequately controlled for. Second, reservation quality services provided by hospitals in response to demand uncertainty have not been taken into account. Finally, the treatment of specification, measurement and sampling errors is non-existent in most studies and not very satisfactory in the work of Valdmanis (1992).

7. SURVIVAL ANALYSES

The preceding review revealed the persistence of seemingly intractable problems - such as the treatment of the case-mix and quality effects within a flexible cost or production function. An alternative approach, called survival analysis, is thought of as less data-demanding. The technique shifts attention away from cost and its determinants towards the study of the intertemporal changes in hospital size distribution. Its essence lies in the assertion that competition among hospitals of different sizes will bring about the disappearance of hospitals with inefficient sizes. Thus, rather than trying to directly estimate production functions, it infers the optimal size from the results of the operation of the health care market.

Hospitals are classified into size classes and the market share of each class over time is computed. If the market share of a given size class increases over time its relative efficiency is inferred. Moreover, hospitals within a class are more efficient (inefficient) the more sharp is the rise (decline) of its share. A range of optimum sizes rather than a single optimum size will emerge, one reason being the lack of access to identical resources by hospitals. The technique is to some extent valid in markets other than purely competitive since oligopolistic firms also have an incentive to adjust to more efficient sizes in search of larger profits.

A heroic further step is to attribute an intertemporal expansion (shrinking) of a size class to the presence of economies (diseconomies) of scale.[99] "Those size categories which grow relative to the rest....are presumed to have *some* [emphasis added] advantage over other sizes".[100] The size advantage, however, may not be scale economies implying that growth may be partly or wholly ascribed to other factors. Simple survival analysis can be used to estimate optimal sizes but only in the sense of

---

[99] See Stigler (1958).

[100] See Bays (1986).

"competitively most effective".[101] This notion of optimum may differ from that of "efficient size" determined solely by scale effects.

The main problem of interpreting changes in market share is that several factors influence hospital survival (closure) or growth:

i) Observed survival and growth of larger hospitals may be due to the exploitation of suppliers or predatory policies.

ii) A documented fall in their market share may be caused by the fear of anti-trust legislation. And the growth of small hospitals may be an attempt to escape anti-trust laws. So, the survival approach does not disentangle the portion of growth that can be ascribed to scale economies.[102]

iii) Evidence that small hospitals are more likely to fail does not support the existence of economies since smaller hospitals are typically rural hospitals so that location could equally explain the findings. The list of potential confounding factors is very long.[103] In principle, nevertheless, progress can be made if some of these other determinants of growth are controlled for.

iv) It implicitly assumes that the magnitude of economies of scale will be large enough to significantly affect the survival of the hospital. This may not be true so that a finding of no significant association between size and hospital survival cannot rule out their existence.[104]

---

[101] See Weiss (1964).

[102] See for instance Weiss (1964).

[103] For some additional factors affecting survival, see Mullner, Rydman and Whiteis (1990).

[104] This point is raised by Gruca and Nath (1994). The authors acknowledge this limitation and do not use survival analysis to explore economies of scale or scope. Instead they run a cost regression to check whether scale or scope phenomena exist and then examine whether their possible presence affects survival.

7.1 Criteria for Validity Assessment

A number of factors may confound estimates of economies of scale, beyond the ones referring to the cost or production structure (e.g. profitability). Economies and returns to scale are cost and production concepts and as such can be studied directly via cost or production studies with the guidance of economic theory as to which minimum set of variables is relevant (e.g. inputs, outputs, quality, uncertainty, efficiency). In the context of survival-type analysis, which is an indirect method of studying economies, determining a detailed set of criteria to assess biases from the omission of relevant individual variables is more difficult. However, multi-variate survival analyses are more reliable than descriptive uni-variate studies since the former at least control for some potential confounding factors. And some of the biases in the results of multi-variate models can be highlighted. More interestingly, the potential biases may indeed be expected to be severe if a model does not perform satisfactorily in light of goodness-of-fit measures.

7.2 Sensitivity Analysis and Validity Assessment

Traditional (uni-variate) survival methodology has been applied in the hospital setting by Bays (1986), Mobley (1990), Vita et al. (1991), and Mobley and Frech (1994). Some of these studies do not attempt to draw any inferences on the presence of economies of scale, acknowledging that entry, exit, price and output decisions are partly determined by regulatory agencies for which efficiency may not be the main objective. Figure 3 illustrates their findings on the range of optimal sizes, as reported in Table 5.

The horizontal axis indicates the possible bed size ranges, whereas the other horizontal lines show the ranges of optimal sizes documented by each study. It is apparent that there is no consensus apart from the fact that very small hospitals (with less than 100 beds) are characterised inefficient, a finding which - as we will shortly see - has been questioned by Simpson (1995). In any case we choose to assign a zero validity weight on these uni-variate descriptive models on the basis of extremely poor quality.

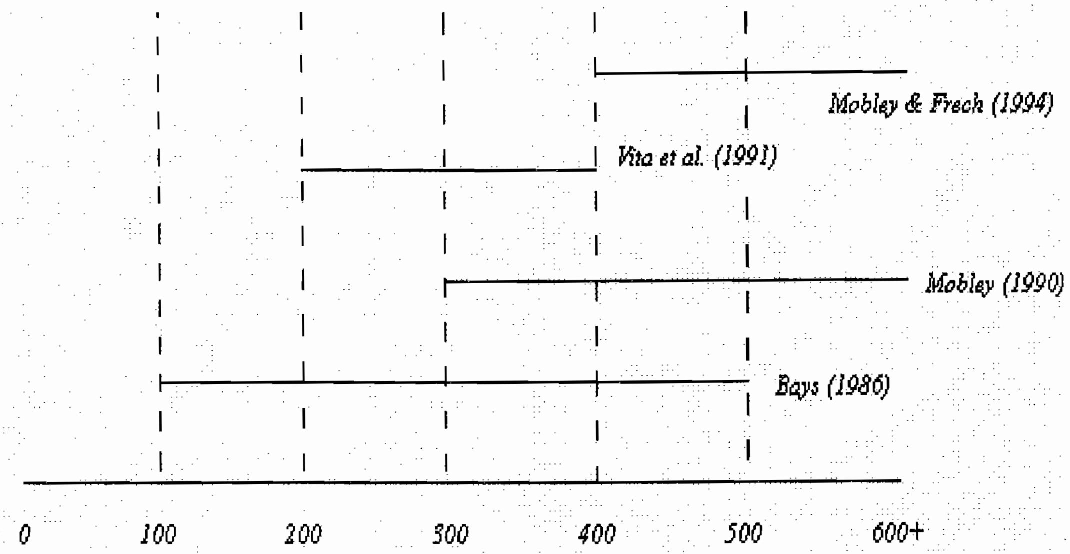

Figure 3: Findings on optimal hospital size from uni-variate survival studies.

A more sophisticated version is employed by Mobley and Frech (1994). It is a model of hospital growth and failure that controls for other demand-related determinants of survival and growth in order to isolate the size effect. Specifically, a break-even expected growth rate is one at which expected benefits (EB, a function of current size, ownership, etc.) minus expected costs (EC, a function of market competition, market share, etc.) of operation equal zero. The hospital will close if its expected growth rate (G^*) is less than this minimum (G_{min}) which is unobservable. In turn, the probability that the hospital will close, that is,

$$\text{Prob}\ (G^* > G_{min}) \Leftrightarrow \text{Prob}\ (EB - EC) > 0$$

is reduced as G^* declines. Hence increases in EC or decreases in EB will make exit more probable for a given G_{min}. So:

$$\text{Prob (Exit)} = \text{Prob}\ (G^* < G_{min}) = F(EC, EB)$$

with $dF/dEC > 0$ and $dF/dEB < 0$.

Equivalently, the probability of survival is:

$$\text{Prob}\ (G^* > G_{\text{min}}) = F\ (\text{current size, ownership, market competition etc.})$$

The model of exit decision is applied to 1980-1989 California hospital data, a period of increased competition and deregulation. The expected growth and exit decision equations are estimated simultaneously. Scale economies in quantity and quality are jointly found to exist up to a point of 325 beds.[105] Similar multi-variate continuous and binary growth/survival models have been used by Mobley (1990) and Frech and Mobley (1995) in an attempt to isolate the net impact of size. Continuous models estimated via OLS seem to perform better. In the earlier study, economies of scale are found to exist up to about 300 beds, with no diseconomies ever occurring. The more recent one suggests an optimum of 200 beds, with a 95% confidence interval extending the range to 370. R^2 in the three models takes values between 0.136 and 0.206.

Two other studies examine the factors affecting the risk of closure by the use of logistic regressions but do not claim to have adequately adjusted for the effects of confounding factors. Both use regressors that have been known to affect the risk of closure, such as utilisation and market characteristics. Among their differences is the inclusion of a case-mix index by Lillie-Blanton et al. (1992) and a DEA efficiency measure by Lynch and Ozcan (1994) as independent variables. The former study finds that hospitals larger than 200 beds in size are 2.5 times less likely to fail than those in the 100-199 capacity category, and 5 times less likely to close than hospitals with less than 100 beds. Yet it combines all hospitals with more than 200 beds into a single category (variable) which in turn hinders the identification of the precise location of the optimum. It can only be inferred that this lies somewhere in the 200+ region. The study by Lynch and Ozcan (1994) is even less informative in this respect since it uses a single variable for size, the sign and significance of which simply indicate that there are economies of scale for larger hospitals.

---

[105] This basically means that the authors failed to keep case-mix constant.

The findings from multi-variate survival-type studies are illustrated in Figure 4 below. More details on survival studies can be found in Table 5. The study by Lynch and Ozcan (1994) is not depicted for reasons already explained.

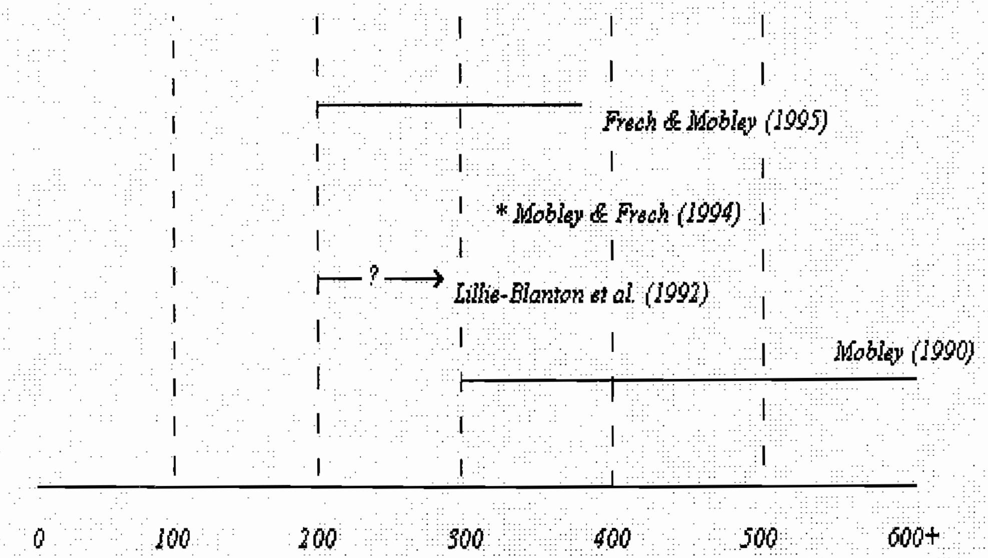

Figure 4: Findings on optimal hospital size from multi-variate survival studies

There is a wide agreement among these studies that hospitals with less than 200 beds are scale-inefficient. The studies by Mobley and Frech (1994) and Frech and Mobley (1995) suggest an optimum of 325 and 200-370 beds respectively, whereas Mobley (1990) in addition argues that no diseconomies ever set in for larger hospitals.

The low adjusted R^2 in the regressions indicates a large proportion of unexplained variance so that many uncontrolled variables possibly correlated with the size variables may exist. Simpson (1995) explores one such factor that may have accounted for the minimum efficient size computed in exit (mortality) and growth (survival) models. He criticises previous studies claiming the inefficiency of small, sub-100 bed, hospitals because they utilise California data from periods prior to 1987 when Certificate of Need (CON) legislation was still in effect. Post-1987 hospital data show a large number of sub-100 bed hospitals entering the market. This implies that small hospitals may also have a relatively higher rate of entry than other hospitals. Hence the documented decline in market share may be reflecting entry restrictions preventing the replacement of the exiting sub-100 bed hospitals by new ones of the same size.

Another potential confounding factor may be the existence of economies of scope exploited by some hospital sizes but not by others. The role of economies of scope within the survival framework has not been explored. The observed decline in the market share of smaller hospitals or their higher probability of exit may be the result of a combined effect of both scale and scope disadvantages. In addition to the problems mentioned, case-mix and quality adjustments are non-existent or at best unsatisfactory.

More generally, there are financial and other possible risk factors - in addition to those related to the cost or production structure - that make the isolation of the scale effect difficult. Hence, the results from survival studies cannot readily be interpreted as due to economies of scale and should be interpreted with extreme caution. This conclusion is corroborated by the unsatisfactory values of the measures of goodness-of-fit.

Finally, one should be reminded of the argument that a finding of no association between size and survival cannot preclude the existence of economies. Mullner, Rydman and Whiteis (1990) did not find a significant association between either beds or admissions and the risk of closure in their initial multivariate analysis. Indeed, there can be no implications drawn about the absence of economies whatsoever and hence the study was not presented here.

8. STUDIES EXAMINING A HOSPITAL SERVICE IN ISOLATION

The search strategy has also revealed 13 studies (presented in Table 6) which focus on a particular hospital ward or service. In essence, these assume that hospital production can be broken down to many independent production procedures, each referring to a particular ward or service. It follows that a separate cost or production function can be employed to study the existence and magnitude of scale effects of each service or ward. Since the output produced in these cases is considerably less heterogeneous than the whole range of hospital product-lines, this argument in turn means that the seemingly intractable case-mix problem is *ipso facto* solved. The studies have used various statistical or econometric methodologies, the validity of which is now assessed.

8.1 Criteria for Validity Assessment

One criterion is used to assess the validity of the approach itself. That is, we will examine whether it is indeed appropriate to examine production processes of sub-sets of hospital services in isolation. Moreover, it is crucial to check whether a particular study design does in fact produces estimates of economies of scale rather than misinterpreting their notion. Finally, some of the criteria already developed can still be applied here. For instance, the *ad hoc* or flexible nature and the derivation of long- rather than short-run estimates form the criteria to evaluate an estimated cost function for a specific sub-set of hospital services.

8.2 Sensitivity Analysis and Validity Assessment

Finkler (1979) examined the presence of economies of scale in heart surgery. The aim was to explore evidence for or against regionalisation of heart surgery.[106] A method of

---

[106] The concept of regionalisation refers to potential savings that may accrue by using centralised facilities at fewer hospitals with high patient volumes. Savings are the difference between economies of scale - if any - and the increased cost of travelling, patient inconvenience, etc.

cost-finding is implemented which excludes joint costs and identifies the costs of the hospital service that would not be incurred if the service were not provided. Fixed costs and variable costs are identified. However, variable costs are said to be the same per patient at any volume. An average cost curve is derived which falls with the number of open-heart surgery patients per year.[107] The author thus argues that there are considerable economies and that the least cost size of heart surgery units is at about 500 procedures per year. However, note that this is an average fixed cost curve thus only capturing the obvious scale economy from spreading fixed costs over a larger number of output units.[108] The same mistake of not considering potential savings from increasing the efficiency of variable inputs via an increase in the volume is also made by Finkler (1981) and McGregor and Pelletier (1978). These studies are thus assigned a zero validity weight.

Munoz et al. (1990c) applied a simple statistical methodology to study the existence of economies of scale in the treatment of urology patients. Urologists at a large medical centre are arbitrarily divided into low volume (treating 5 patients or less within a DRG over a 3 year period) and high volume (8 patients or more). Cost per patient for low- and high-volume urologists is adjusted for case-mix and severity of illness (proxy) and subsequently compared. High-volume urologists are found to have significantly lower unit costs. The same methodology is applied by Munoz et al. (1990a,b) for orthopaedic and neurosurgical patients with similar results documented. The authors themselves stress the crude proxy for severity of illness which implies that study findings might be due to the fact that low volume surgeons might treat more complex - and hence costly - cases. Moreover, some of the physicians under study were full-time staff members whilst others were private practice part-time employees. There is therefore the possibility that the study classified a physician as "low volume" when in fact he was a "high volume" one in aggregate. These two problems were better treated within the linear regressions of average charges on surgeons' volumes of cholecystectomy, prostatectomy, hysterectomy and intervertebral estimated by Arndt Bradbury and

---

[107] There is a kink in the curve since some costs are fixed only for certain ranges of volume.

[108] Well stated by Long et al. (1985), the study "adds nothing to the understanding of economies of scale within a specific procedure".

Golec (1995). A dummy variable was incorporated for each hospital in order to account for the different charges to costs ratios of different institutions.

Yet even if these limitations are assumed away, what do these results really tell us about economies of scale and the optimal configuration of hospital services? If such economies exist they might be due to say a learning curve for the surgeon. Other important sources of economies or diseconomies of scale that might arise through an increased concentration of services within few large hospitals are ignored. For example, economies due to indivisibilities in medical equipment or uncertain demand are not measured in such study designs. Nor are the diseconomies related to management that could arise in larger hospitals. Further, economies of scope (if they exist) might be lost by not offering, say, neurosurgical services in relatively small hospitals. These studies cannot deal with these crucial issues.

Other studies have employed cost or production functions to examine in isolation a particular sub-set of relatively homogeneous services. The *ad hoc* models claimed to have found economies of scale for maternity, clinical laboratory, nuclear medicine, dental care, neonatal care and knee replacement surgery (see Table 6).[109] A study that could deserve more attention is the flexible cost model by Okunade (1993) exploring the cost structure of hospital pharmacies. In the one-output, multiple-input translog model, slight but statistically significant short-run diseconomies of scale are found at the sample means. The most efficient operating size in the short-run is the median bed size category i.e. 200-299 beds.

However, this along with all the other studies mentioned in this section share a common limitation. They assume that hospital production is separable. That is, hospitals do not use their inputs to produce joint products (e.g. one piece of medical equipment to be used from two different hospital wards). Stated in terms of costs, there are no economies or diseconomies of scope in hospitals.[110] However it might be

---

[109] These are: Hu (1971), Anderson (1974), Wilson and Jadlow (1982), Wan et al. (1987), Fordham et al. (1992) and Culler, Holmes and Gutierrez (1995).

[110] See Cowing, Holtmann and Powers (1983) and Gravelle and Rees (1992).

expected that the cost of providing - for instance - maternity care might increase if a hospital eliminates its main paediatric service. It is apparent that these are very restrictive assumptions that are likely to be violated and bias the results on economies of scale.

9. EVIDENCE FROM MULTI-HOSPITAL ARRANGEMENTS [111]

A separate and voluminous literature has been concerned with the impact of various forms of consolidation on hospital , financial performance as well as on effectiveness (improved quality). A main question was whether the motive for the observed mergers or other multi-hospital arrangements in the US was increased operating efficiencies or instead more aggressive pricing policies and the exploitation of monopoly power. Operating efficiency was primarily examined in terms of changes in cost per case or per day and productivity indices following an arrangement. Cost per case, adjusted for case-mix, was expected to be lower in hospitals participating in a multi-hospital arrangement, one reason being the exploitation of economies of scale (or even scope). With the exception of the study by Sinay (1994), this literature at best fails to identify the portion of the change in average cost that is attributable to economies of scale as opposed to scope effects, and *vice versa*. At worse, other uncontrolled factors (e.g. change in pure technical efficiency) due to the radical restructuring of hospitals following the merger might be partly or wholly responsible for any change in the behaviour of cost per case. Mobley (1990) for instance, mentions the bias that might characterise efficiency studies given the change in accounting practices in the post-merger period. Or it could be, as Manheim, Shortell and McFall (1989) note, that efficiencies are gained by improving the management of a hospital acquired by an investor-owned chain. Although a detailed review of the mergers literature falls beyond the scope of this study, it is still useful to examine some of the findings.

Empirical studies have used basically two methodologies to examine this important issue of hospital performance. The first is the statistical technique which constructs pairs of observations by matching independent and system-affiliated (or merged) hospitals for various factors (e.g. available services, ownership, location) other than

---

[111] The term "multihospital arrangements" is meant to describe formal inter-hospital collaborations. The degree of autonomy of individual hospitals varies. A merger is an extreme form of such a collaboration in which hospital boards of previously independent units dissolve and a new organisation with common governance and management emerges. See Markham (1995).

the ones of interest. Performance indicators, such as adjusted average cost per case, can then be compared across the matched samples. Levitz and Brook (1985) adjusted both cost per case and per day for case-mix intensity and found the former to be significantly higher for system-affiliated hospitals, whereas the latter measure insignificantly so.[112]

A paired comparison, before-after type analysis was also employed by Treat (1976) for the assessment of mergers. The study compared average cost per case using multiple time frames to allow an evaluation of the short-, intermediate- and long-run impact of mergers. It was higher for urban but lower for rural hospitals, suggesting that efficiency can only be improved through mergers of small rural units. However, the findings cannot be solely attributed to the presence of economies or diseconomies of scale (or better to a combination of scope and scale effects). For instance, a higher cost could well be due to a merged hospital making capital investments in order to improve quality. Moreover, time / travel costs are not considered and it could well be that even small rural mergers are not justified on efficiency grounds. Pattison and Katz (1983) also found investor-owned chain hospitals to have a higher cost per case than free-standing voluntary institutions, attributing the enhanced profitability of the former to the aggressive marketing and pricing strategies used rather than to cost savings. Similar results for cost per day were obtained by Lewin, Derzon and Margulies (1981).

This methodology has been criticised for its inability to control for the myriad confounding factors affecting hospital performance. Multiple regression has thus been suggested as a more appropriate design. A sample of hospitals is typically employed including independent hospitals as well as members of multi-hospital systems. Average cost per case is regressed on control variables and a dummy variable taking the value 1 if the hospital is system-affiliated, 0 otherwise. The significance and sign of its estimated parameter then indicate whether system hospitals indeed fulfilled the expectations of lowering unit costs.

---

[112] Cost per case is the preferred measure since it accounts for length of stay.

Coyne (1982) runs separate cost per case regressions for different hospital ownership types (religious, other non-profit, investor-owned, county) and controls for differences in case-mix, demography, competition and regulation. System hospitals of all ownership types are shown to incur higher unit costs except for county-owned system hospitals, which instead have lower cost per case (R^2 0.79-0.98). Becker and Sloan (1985) did not also find any conclusive evidence that system hospitals are more efficient than independents, after controlling for case-mix and other factors. Mobley (1990) suggested that past research suffered *inter alia* from failing to adequately control for a number of factors affecting average cost (e.g. case-mix, quality, increased insurance market competition). His model of cost inflation however, did not find any significant evidence that economies of multi-plant operation (economies of scope included) existed for system hospitals.

The failure of multihospital arrangements in achieving lower unit costs is also supported by the reviews of Ermann and Gabel (1984) and Markham (1995). Thus, the merger literature might suggest that in practice economies of scale and scope existent at the hospital level are not realised in such arrangements. The evidence however is also in line with our earlier review results that economies are in fact absent for medium and large-sized hospitals. Reinforcing evidence comes from the studies finding average cost to be lower only in small rural systems.

The translog study by Sinay (1994) tested the hypothesis that hospital mergers in the 1980s reduced production costs by achieving economies of scale and scope. It is a before-after study with cost functions estimated for an experimental group (merged hospitals) and a control group (non-merged), matching the characteristics (location, ownership and system status, size, services provided) of the merged hospitals. Economies of scale and scope are computed from the cost functions which are estimated one year prior, and one and two years after the mergers. Several scenarios are considered. Suppose that the experimental and control groups prior to the merger experience, on average, (unexploited) economies of scale and scope. Following the merger, the control group is expected to continue to operate inefficiently (e.g. average non-merged hospital experiences the same unexploited economies), whereas the merged hospitals are expected to become more efficient (e.g. evidence of no unexploited economies for the average hospital). The analysis is repeated separately for two merger episodes to test the robustness of results.

The author claims to have found that merged hospitals eventually manage to exploit economies. In one merger episode unexploited ray scale economies were found prior to the merger and neither economies nor diseconomies two years after the merger. Control hospitals at the same period were and remained efficient (constant returns). However, the results do not seem to be conclusive: in the other episode studied, merged hospitals experienced diseconomies prior to the merger but (unexploited) economies two years after when the average hospital size in the sample increased from 229 to 429 beds. Moreover, the model employs the day as the unit of output and adjusts crudely for input prices, factors which may have biased the results.

In short, the evidence is in line with our finding that economies of scale may be more likely to exist only for small hospitals. It is also worth while noting that economies of scale, even if they are present for small hospitals, could presumably be achieved through internal expansion of hospitals rather than merger activities. There might then be alternative motives for multi-hospital arrangements. Dranove and Shanley (1995), for instance, found that local multi-hospital systems do not have lower costs but rather appear to enjoy reputation benefits over non-system hospitals.

Finally, Lynk (1995) claims that evidence on economies from studies using the hospital as the unit of observation would only be relevant for the evaluation of mergers in which all the activities of the merged hospital are subsequently transferred to a single fully-integrated site. Typically, administrative and overhead activities are indeed centralised and clinical departments operated at a single site. Yet inpatient care is still provided at separate locations. The authors therefore focus on mergers and their empirical analysis shows that there exist economies of scale due to stochastic demand from clinical consolidation. Nevertheless, excessive consolidation might not be justified since eventually these economies are exhausted and patient travel costs might increase. We note however, that the points made by the author imply that if economies do exist at the hospital level they might not be realised via mergers. The argument hence applies for small mergers for which there is some evidence that economies are in fact present. But our finding that they do not exist even at the individual facility level for average and larger hospitals means that they will not exist at a larger system level. The only real problem is that in traditional cost analyses all factors (especially quality and case-mix) cannot be easily kept constant, thus possibly under-estimating measures of economies.

10. OVERVIEW AND CONCLUSIONS

Economies of scale for general hospitals have been examined extensively by econometric flexible cost models. The more reliable studies find constant returns or even diseconomies for the average hospital, the latter being defined as one with roughly 200-300 beds. Although it is impossible to pinpoint the exact optimum size, it is apparent that if any economies exist they are quickly exhausted or outweighed by diseconomies. Large hospitals of 400 beds or more might be too large with respect to cost minimisation or at best no more beneficial than smaller units. The absence of cost savings from expanding the scale of production is corroborated by *ad hoc* cost models. These are in general less reliable than flexible cost functions but have used a larger number of case-mix variables. The production function approach confirms these findings.

DEA reinforces the view that economies can be exploited only up to a hospital size of about 200 beds. It also suggests that hospitals larger than some 650 beds are scale-inefficient. Results are conflicting regarding the exact position of the optimum. It seems however that the optimum may (ie between about 200-400 beds. It is encouraging that these findings are not very different from those in flexible econometric models, since both techniques have their relative strengths and weaknesses. The econometric methodology is based on more restrictive assumptions about a hospital's technology, whilst DEA might face more problems with the choice of variables and errors in the data. In fact, a Monte Carlo simulation study has shown that DEA outperforms econometric techniques only if errors are small.[113]

Survival analyses are less reliable mainly because it is more difficult to identify and control for factors, other than size, that affect survival. This proposition is corroborated by the low R^2 found in such studies. Yet they have been seen as a complementary tool in the analysis of economies, in light of the potential biases that

---

[113] See Banker, Gadh and Gorr (1989).

arise from the fact that hospital data may be reflecting accounting costs and conventions rather than true economic costs and hospital behaviour. Unexploited economies are again reported for hospitals with less than 200 beds. It is less clear however, whether diseconomies ever set in and whether hospitals within the 200-300 range can still exploit some further cost savings.

The findings of the literature focusing on the hospital level are also broadly in line with the evidence provided by studies examining the impact of mergers and other multi-hospital arrangements on costs.

Econometric hospital cost studies have also been used to examine the existence of economies of scope. Their existence was not in general empirically validated for the average hospital, but this does not necessarily mean that they are indeed absent. Some questionable evidence exists that there might be some economies between obstetrics and medical (acute) care and between acute and sub-acute care.

Global measures of product-specific economies of scale and economies of scope were found to be of dubious validity in the specific studies and hence are not discussed. Finally, we pointed out the limitations associated with research attempts to study a particular sub-set of hospital services in isolation. This does not of course mean that all other hospital-based studies are unproblematic. Highly reliable estimates of economies can only be obtained, in principle, if all other factors are adequately controlled for. It seems, however, extremely difficult to adjust for differences in case-mix and quality of care across hospitals. Other limitations of the existing literature have also been identified (e.g. derivation of true long-run estimates of scale effects, uncertainty). Thus, all evidence reported in this review should be interpreted with a degree of caution.

TABLES AND APPENDICES

Table 1: *Ad hoc* econometric cost studies

Author Year	Sample Characteristics	Study Objectives	Specification of Cost Function(s)	Output Measurement (adjustment)	Findings on Economies of Scale
Feldstein, M. (1967)	Cross-section of 177 NHS (UK) large short-term acute care, non-teaching hospitals in 1960-61. Mean bed size is 303.	First comprehensive attempt to explore efficiency in health care	Quadratic specification incorporating size (beds), case-flow and case-mix.	9 case-mix proportions by medical speciality group.	R^2 is 0.74. U-shaped curve with implied minimum at about 900 beds but coefficients not significant. Constant returns or slight diseconomies are concluded. So, the optimum could well be at 300 beds.
Carr & Feldstein, P. (1967)	Cross-section of 3147 US voluntary short-term general hospitals.	Estimation of the effect of size on costs for the whole sample (number of services provided in RHS) and for groups of hospitals with similar number of available services.	Quadratic specification of a total cost function from which AC (per patient day) is retrieved. Size (patient days), number of provided services, outpatient visits and 5 variables capturing teaching services in RHS.	Teaching services measured by existence of nursing school, number of student nurses and residents etc. Variable measuring the number of available services (maximum 28) in a hospital.	Regression indicates economies of scale up to a point of 190 (Average Daily Census) after which diseconomies operate. Group analysis shows similar patterns.
Berry (1967)	Grouping of 5293 non-federal, short-term general and special US hospitals in 1963. 40 resulting samples each with the same facilities-services.	Adjustment for product heterogeneity to isolate the effect of scale of production on costs. Grouping of hospitals according to the availability of services / facilities and separate regressions for each "homogeneous" group.	For each of the 40 groups, linear regression of AC per patient day on total patient days. So, output "patient days" is used as the scale variable.	Service-mix approach used for grouping without adjustment for service intensity.	AC declines as output increases in 36 cases. 26 negative coefficients are said significant having a t-statistic greater than 1. Economies of scale concluded but constant returns are in fact present at a 5% level.
Cohen (1967)	Initial sample: 82 short-term general hospitals from six US States in 1963-64. Subsamples: one from New York (35) and one from other five states (47).	Study of the significance of adjusted output and wage differentials on hospital costs.	Quadratic regression of adjusted (for wage differentials) total cost on the constructed composite measure of output.	Adjusted measure of output is used as the measure of size to compute economies. Output measured by a weighted sum of 13 service outputs (operations, X-rays, outpatient visits etc.). The weight is the service's relative average cost.	New York subsample: an approximately U-shaped Average Cost curve with a minimum at 290-295 beds. R^2 is 0.99. Five-state subsample: Same shape of AC curve but optimum size is at 160-170 beds.
Ro (1968)	68 US hospitals over a 10-year period.	Study of the determinants of hospital costs.	AC per day specification with 25 independent variables used in separate analyses.	Service-mix adjustment with no measure of case-mix.	Economies over sample range. The minimum cost point is well above 794 beds (sample's max.).
Ingbar & Taylor (1968)	Pooled data from 72 short-term general Massachusetts hospitals during 1958-59.	Study of the cost determinants of Massachusetts hospitals.	Quadratic AC per patient day specification. Principal component method reduces number of regressors to 14 including length of stay, utilisation and service activity.	Measures of service activity.	Inverted U-shaped AC with a maximum cost at 150 beds. But cost reductions for larger or smaller hospitals were very small and hence a constant AC was implied.
Berry (1970)	Some 6000 short-term general US. hospitals in 1965.	Adjustment for product heterogeneity in cost functions for hospitals.	AC per patient day regressed on Average Daily Census, its squared term (output level) and on other service-mix dummies & other controls.	Service-mix captured by 27 dummy variables measuring the availability of facilities and services.	The best equation's R^2 is 0.50. AC is U-shaped.

Table 1: *Ad hoc* econometric cost studies (Cont'd)

Author Year	Sample Characteristics	Study Objectives	Specification of Cost Function(s)	Output Measurement (adjustment)	Findings on Economies of Scale
Cohen (1970)	46 New York short-term general hospitals in 1965.	Refinement of Cohen (1967) work.	Quadratic regression.	15 service outputs used. Quality proxied by a dummy capturing the degree of medical affiliation.	Similar results as in Cohen (1967). R^2 again high. But optimum size at 540-555 beds when quality is included. Teaching hospitals have a higher optimum than non-teaching.
Francisco (1970)	1328 US short-term general hospitals in 1966.	Examination of cost per patient day variations by means of regression analysis.	Quadratic regression of AC per patient day on output (patient days).	25 "homogeneous" groups of hospitals with respect to the availability of facilities/services. Separate regressions within each group. Alternatively, an unweighted facility index and a single regression for the whole sample.	Significant economies of scale for small hospitals (with less than 70 beds and 5 facilities). Constant returns to scale for large hospitals. R^2 lower than 0.50 in all equations.
Lave & Lave (1970a)	74 Western Pennsylvania hospitals for the period 1961-67. 14 semi-annual observations obtained on each hospital.	Estimation of time-series models for individual hospitals and of a model based on pooled data. Models assume constant output mix within a hospital over "short" periods. Several explanatory variables included in search of significance.	Generalised Cobb-Douglas (linear in logs) functional form. AC per patient day is the regressand. Hospital size, utilisation, time the main regressors.	No adjustment: output-mix excluded by assumption.	Constant returns are concluded.
Lave & Lave (1970b)	Pooled annual data from 74 western and 35 eastern Pennsylvania hospitals for the period 1961-67.	Estimation of a cost equation applied to time-series, cross-section data which assumes that output-mix is constant over time.	Similar to Lave and Lave (1970a).	No adjustment for case-mix: it is assumed that over time case-mix is constant and that across hospital variation is captured in size, location and teaching status variables.	Constant returns for Pittsburgh hospitals and modest economies for Philadelphia ones (R^2 is 0.90).
Evans (1971)	185 Ontario hospitals in 1967.	Analysis of cost relations reflecting behavioural rather than merely structural factors. Also development of cost functions adjusted for output-mix.	Cost per case in a quadratic specification. RHS includes case-mix variables, beds, beds-squared, length of stay, occupancy rate, case-flow rate.	10 case-mix variables result from factor analysis.	R^2 is about 0.60. Weak diseconomies or constant returns concluded.
Lave, Lave & Silverman (1972)	Two cross-sections of 65 and 47 US hospitals in 1968 and 1967 respectively. Mean number of beds is 246.	Control for output heterogeneity by combining diagnostic variables together to reduce multicollinearity.	Linear-additive (and quadratic term subsequently tested) AC (per patient) equation with many control variables.	17 broad ICD diagnostic categories reduced to five composite variables using 3 techniques. Also age and 5 other supplementary variables.	R^2 between 0.93 and 0.97. Coefficients indicate economies but are insignificant. So, constant returns are concluded.

Table 1: *Ad hoc* econometric cost studies (Cont'd)

Author Year	Sample Characteristics	Study Objectives	Specification of Cost Function(s)	Output Measurement (adjustment)	Findings on Economies of Scale
Evans & Walker (1972)	90 British Columbia (Canadian) hospitals. Hospitals of various sizes with a few very small (<25 beds) and very large (>1000) ones.	Development of case-mix measures based on information theory which are included as variables in cost analysis.	AC per case in a quadratic specification. RHS includes beds, beds-squared, case-mix indices developed, age-sex variables, length of stay, occupancy rate and case-flow rate.	Case-mix indices developed using the degree of concentration of case-types across hospitals to proxy case complexity. Also age-sex variables.	R^2 is 0.93. Moderate diseconomies of scale suggested. Economies only for hospitals with less than 100 beds. Economies implied in per day equation.
Lee & Wallace (1973)	Two cross-sections of US data from 52 hospitals in 1966 and 73 in 1967.	Discussion of the problems in estimating multi-product cost functions for hospitals. Also, development of two output classification schemes.	Unit cost per day as a linear function of case-mix proportions.	Two case-mix classification schemes tested: a) 5 groups based on duration and extent of disability. b) 16 groups based on ICD.	R^2 ranges from 0.14 (1st specification) to 0.52 (2nd). Considerable economies concluded for all types of care.
Berry (1974)	6000 US short-term general hospitals in 1966.	Empirical investigation of the determinants of hospital costs.	Unit cost quadratic equation. RHS includes variables for the level and quality of output, services provided, factor prices (two proxies) and efficiency.	Many service-mix variables for the services provided by the hospital. Quality measured by accreditation status dummy.	R^2 ranges from 0.57 to 0.69. A shallow U-shaped AC curve is found. Economies are not large in absolute terms to justify expansion of scale.
Fottler & Rock (1974)	Two cross-sections of 36 voluntary and 14 municipal New York hospitals in 1965 and 1970.	Measurement of the impact of several variables on costs in two types of hospitals. RHS variables include work-force skill level, hospital size (beds), technological sophistication, work-load, case-mix and case-severity.	AC per patient day is the regressand and the additive-linear functional form is adopted.	Case-mix measured as the proportion of medical-surgical cases. Case-severity as the average length of stay adjusted for occupancy rate. Quality proxied by the workforce skill level and an index of facilities and services.	Voluntary hospitals: diseconomies in the 1965 and 1970 equations. Municipal hospitals: economies in 1965 but slight diseconomies in 1970. The R^2 is high in all equations ranging from 0.67 to 0.88.
Feldstein, M. & Schuttinga (1977)	55 Short-term, non-teaching Massachusetts hospitals in 1972. Means of beds and cases are 188 and 6652.	Development of costliness measures for valid comparisons of costs across hospitals.	AC per case and AC per day quadratic functions.	10 diagnostic and 10 surgical case-mix proportions derived from principal component analysis.	R^2 is 0.68. Constant returns to scale are found in the cost per case equation whereas economies of scale in the cost per day one.
Zaretski (1977)	1971 data from 176 acute care California hospitals. Mean number of admissions is 7400.	Modelling of the determinants of hospital costs with emphasis on output standardisation.	Linear cost per case equation including admissions, variables for short-run excess capacity, service- and case-mix, and location dummy in RHS.	12 principal components reflecting case-mix result from 45 original ICDA-8 aggregated groups. Also two methods for service-mix adjustment.	The best equation has an R^2 of 0.75 Constant returns to scale are evidenced.
Culyer et al. (1978)	268 acute hospitals of more than 100 beds in England in cost year 1969-70.	Modelling of the determinants of hospital costs in teaching hospitals.	Quadratic AC per case equation similar to Evans and Walker with the number of students per case in the RHS.	The Evans-Walker case-mix complexity index.	R^2 = 0.78. Economies of scale are present.

Table 1: *Ad hoc* econometric cost studies (Cont'd)

Author Year	Sample Characteristics	Study Objectives	Specification of Cost Function(s)	Output Measurement (adjustment)	Findings on Economies of Scale
Pauly (1978)	50 non-major-teaching, short-term hospitals in California for a three-month period in 1975. Mean bed size is 181.	Examination of the impact of medical staff characteristics - such as speciality mix and concentration of hospital output among attending physicians - on hospital costs.	Cobb-Douglas total cost function. The RHS includes adjusted admissions/discharges, case-mix index, 4 input prices, % of discharges for normal delivery, medical staff characteristics.	Outpatient visits converted into inpatient equivalents by relative visit revenues. Case-mix index is the weighted sum of the proportion of cases in each diagnostic category. Each weight is the mean charge for the category to that for all diagnoses.	The R^2 is 0.982 and constant returns to scale are evidenced.
Ault & Johnson (1979)	Cross-sections for types of US hospitals (short-stay Veterans Administration, voluntary etc.) with observations ranging from 40 to 350 (1972-74).	Comparative estimation of cost functions for various types of hospitals.	AC per bed or per patient day the dependent variables. Quadratic specification with beds as the size variable. Other variables are the occupancy rate, medical affiliation, location and service mix dummies.	Quality of output proxied by affiliation. A vector of 46 dummy variables expressing the services provided by each hospital.	In the observed bed range (20-2000 beds), short-stay hospitals have a U-shaped AC curve. R^2 between 0.12 and 0.88 in different equations estimated.
Jenkins (1980)	Cross-section of 101 Ontario general hospitals in 1971.	Derivation of comparative estimates from equations defining output in terms of service-mix and case-mix. Also use of a more plausible specification of the cost function remedying collinearity.	Linear AC equations (per day, per admission, per rated day) RHS includes patient-days, admissions, capacity, output variables and variables reflecting returns to scale etc. Case-type, service-type equations estimated. Size measured by rated patient-days.	Service mix measured by bundling services into indices to reduce the number of variables. Case-mix involves 40 O.H.S.C. diagnostic categories. Principal component estimation along with backward stepwise regression is applied to reduce multicollinearity.	R^2 for service- and case-type cost equations are 0.82 and 0.91. Both equations have significant positive coefficients on the squared rated patient-days term indicating diseconomies of scale. Optimum at 100-300 beds.
Brown (1980)	Pooled data from 12 Newfoundland's (Canada) cottage hospitals, 1952-58, 1971, 1976. 24 observations in the cost equation. Mean bed size is 27-30.	Estimation of cost and production functions for Newfoundland hospitals which have salaried physicians so that physician time input can be included. Product mix is also fairly constant so that a simple output measure is said to suffice.	Quadratic AC (per unit of composite output) function. Independent variable is basically size/scale measured as output and output^2 or as beds and beds^2. Idle bed capacity, patient stay and a trend variable also considered.	Composite output computed by weighting outpatient, laboratory and X-ray services by hypothesised / assumed relative resource requirements	Using output as a measure of scale results in economies of scale. The minimum of the AC curve is at 13298 patient days or 57 beds but lies outside the sample range. R^2 is 0.93. Using instead beds similar results are obtained
Bays (1980)	Pooled data for 41 short-term general non-teaching California hospitals during 1971 and 1972 yielding 64 observations. Mean bed size is 124 (min 33, max 285).	Comparison of estimates from a cost function that ignores the physician input price with those from a function that incorporates a proxy to assess the bias on the estimate of scale economies.	AC per case in linear and quadratic specifications. Regressors are the number of beds (size), case-flow (rate of utilisation) and a vector of case-mix proportions. In one equation the physician price is included.	Aggregate case-mix measures derived from initial 19 broad ICDA categories by combining case-mix proportions with similar estimated average cost. 4 and 5 composite regressors result for the two equations to be estimated.	Inclusion of physician input price turns an apparent finding of economies to one of constant returns to scale. R^2 is 0.50.

Table 1: *Ad hoc* econometric cost studies (Cont'd)

Author Year	Sample Characteristics	Study Objectives	Specification of Cost Function(s)	Output Measurement (adjustment)	Findings on Economies of Scale
Finch & Christianson (1981)	Pooled data from 116 rural hospitals in 5 Rocky Mountain States (US) in the period 1971 to 1977.	Estimation of cost functions for rural hospitals and computation of cost-savings from a policy towards optimal size (long-run) and occupancy rate (short-run).	AC per patient day in behavioural quadratic and logarithmic specifications. Bed is the unit of scale. Other regressors involve both service and case-mix variables, occupancy rate, average salary, ownership, time, accreditation.	18 dummies reflecting the presence or absence of a service or facility. Case-mix proxied by birth per 1000 population, the % of an area's population over 65 and the % of long-term care beds. Quality of care proxied by accreditation status.	Quadratic equation: shallow U-shaped long-run AC found with a minimum at 113 beds. $R^2 = 0.75$. Logarithmic specification implies significant scale economies. $R^2 = 0.84$.
Sloan & Becker (1981)	Cross-section of 1228 non-federal short-term general US hospitals in 1972-73. Hospitals are of various sizes including small (<100) and large (>400).	Analysis of the impact of selected factors of internal organisation on hospital costs.	AC per case and AC per day specifications with many controls including organisational factors. 3 dummies for hospital size class.	Resource Need Index and other case-mix controls.	AC per case equation indicates diseconomies whereas AC per day economies. R^2 is 0.65 and 0.49 respectively.
Barer (1982)	Pooled time-series / cross-sectional analysis of 87 British Columbia (Canada) acute care hospitals during 1966-73.	Further testing of the ability of the Evans-Walker case-mix adjustment (based on information theory) in explaining interhospital cost variations.	AC per case equation.	Evans-Walker methodology to adjust for case-mix heterogeneity.	Findings in line with Evans-Walker ones, indicating weak diseconomies of scale. R^2 between 0.77 and 0.87.
Friedman & Pauly (1983)	870 US hospitals over 1973-78.	Econometric modelling that explicitly allows for the impact of unexpected and expected demand changes on hospital costs. Two cost models treat length of stay as endogeneous and as exogeneous.	Deflated AC per admitted case in linear specification. RHS includes the ratio of forecasted to actual admissions, beds, average length of stay, 2 input prices, and service-mix variables.	Service-mix approach used. 5 variables included.	Very slight economies of scale indicated by the "beds" coefficient. R^2 is 0.50.
Sloan, Feldman & Steinwald (1983)	Pooled data from 367 US community teaching and non-teaching hospitals observed in 1974 and 1977. Sample sizes include units with less than 100 and more than 400 beds.	Estimation of the effect of teaching on hospital costs. 6 estimated cost functions of cost per admission and cost per case as the dependent variable using the overall, teaching and non-teaching samples.	Linear specification. RHS includes three dummies for size (bed) classes taking value 1 if a hospital lies within it. Three variables to capture teaching effect. Also location dummies, regulatory environment dummies etc.	Teaching services captured by approved residency programme, medical school affiliation, member of Council of Teaching Hospitals. Resource Need Index (RNI) case-mix index also employed.	Size variables imply economies of scale with cost per adjusted patient day ($R^2 = 0.63$) but diseconomies with cost per adjusted admission used instead ($R^2 = 0.67$).
Becker & Sloan (1985)	Cross-section of about 1650 US hospitals in 1979. Sample includes hospitals in bed size ranges 1-99,......., 400+.	Comparison of the performance of different ownership forms of hospitals.	Linear cost per case and per day equations including three size dummies, ownership variables, and other controls.	Resource Need Index complemented by 5 other case-mix proxies.	Cost per case equation ($R^2 = 0.68$): scale diseconomies. Cost per day equation ($R^2 = 0.58$): scale economies.

Table 1: *Ad hoc* econometric cost studies (Cont'd)

Author Year	Sample Characteristics	Study Objectives	Specification of Cost Function(s)	Output Measurement (adjustment)	Findings on Economies of Scale
Hornbrook & Monheit (1985)	380 US short-term, general, non-federal hospitals in 1977. Sample mean bed size is 268, admissions 9302.	Removal of the bias from differences in utilisation and case-mix across hospitals through the inclusion of both the service- and case-mix measures of output.	AC per case in a quadratic specification. RHS includes volume and composition of output, 2 factor prices, profit status etc. Admissions is the unit of scale, not beds since this would give rise to a short-run cost function.	Case mix: a Laspeyres index of case-mix proportions using average length of stay (proxies resource requirements) weights. Service mix: number of "complexity-expanding" services possessed by a hospital.	Admissions variable insignificant: constant returns to scale evidenced. Omission of case-mix and/or service mix introduces upward bias in estimates of returns to scale. R^2 = 0.73.
Robinson & Luft (1985)	1084 US community hospitals in 1972 is the sub-sample for which available info for the construction of case-mix variables exists. Sample includes hospitals in bed size ranges 1-99,......., 400+.	Analysis of the impact of market structure on average hospital costs, measured both in terms of cost per case and cost per day.	AC per admission and per day equations. 5 bed size classes to measure scale. Capacity utilisation expressed as admissions/year and average length of stay. Many variables to capture case-mix, input cost (crude labour wages) and market conditions. Dependent variables and admissions, length of stay in logs.	Case-mix approach adopted. Variables correlated with case-mix are used (e.g. bed categories, dummies for profit and non-profit status, medical school affiliation). Also 17 diagnosis-specific case-mix proportions, outpatient visits and a variable to capture county-wide case-mix differences.	Significant diseconomies of scale are indicated by the bed-size variables in both per case and per day specifications. R^2 is 0.75 and 0.71 respectively. Diseconomies are found even when the 17 case-mix proportions are excluded from the model.
Vitaliano (1987)	Cross-section of 166 general hospitals in New York State in 1981.	Statistical comparison of logarithmic and quadratic cost functions and assessment of the validity of their estimates of scale economies.	Logarithmic and quadratic total cost equations. The number of hospital beds is used as the size-related "output". Also in RHS: market share, dummy for affiliation, location dummies to capture input price differences.	Service mix unweighted index is used.	R^2 is 0.92 in both regressions. Log model indicates scale economies. Quadratic indicates a U-shaped AC curve but is rejected by a Ramsey test.
Wagstaff (1989b)	49 acute Spanish public hospitals.	Comparative estimation of three statistical cost frontier models and one non-frontier model.	AC per case in quadratic specification (beds, beds^2)	5 case-mix categories included with case-flow proxying case complexity.	OLS estimation indicates neither economies nor diseconomies. Frontier panel data estimation indicates U-shaped with a turning point at about 540-590 beds.
Butler (1995)	4 cross-sections of 121 public hospitals in Queensland (Australia) in 1977-81.	Hospital cost analysis including both theoretical and empirical sections.	AC per case with linear and quadratic terms on the "beds" (size) variable. Measures of scale and utilisation included to correct for non-minimising behaviour in the long-run.	Three adjustments explored: a) 18 ICD diagnostic categories b) 47 ICD categories and c) scalar case-mix index. Also additional complementary case-mix dimensions (e.g. age) considered.	R^2 between 0.58 and 0.92.. Results sensitive to specification either showing a U-shaped curve with a minimum at a size of 469 beds or neither economies / diseconomies.

Table 2: Flexible econometric cost studies

Author Year	Sample Characteristics	Study Objectives and Methodology	Specification of Cost Function(s)	Output, Quality, Uncertainty, Efficiency Adjustment	Model Diagnosis	Findings on Economies of Scale	Findings on Economies of Scope
Roddy (1980)	52 observations from Wisconsin short-term general hospitals in 1976. Non-teaching (non-profit, religious, government) hospitals. Mean bed size is 217.	One objective is to investigate economies in the industry. Another is to assess whether outpatients pay or not their fair share of provided services. It uses a flexible cost function to pursue them.	Total short-run structural (only case-mix additionally included) translog cost function. Beds are the fixed factor. 2 outputs - 3 input prices. Case-mix index included.	Units of measurement: a) inpatient day b) outpatient visit. The two outputs are of course the above. Case-mix index included. Inefficiency might be implied for theoretical reasons (retrospective regime).	Fit: No R^2. All regularity conditions satisfied at the means and for all other observations, except for 2 observations for which the marginal costs are negative.	The most restrictive (homogeneous and separable) model is not rejected. Global long-run economies are absent (it is claimed). Short-run diseconomies exist (computed as the sum of output - cost elasticities, not actual level of fixed input: 1.28 st.error 0.12).	None.
Burns (1982)	140 observations by pooling 35 Connecticut hospitals over 1972 to 1979. Non-federal acute general hospitals. All are non-profit, 1 municipal, 1 State-owned, 3 have religious affiliations, the rest private. Mean bed size is 296.9.	The study uses the structure - conduct - performance framework to analyse the hospital industry. A performance analysis studies the effects of structure on utilisation and costs.	Total long-run structural translog cost function. 5 outputs - 3 input prices. Technological change is not measured.	Units of measurement: a) inpatient day b) outpatient visit. Outputs are: adult patient days, paediatric days, intensive care days, new-born days, outpatient visits. Efficiency cannot be claimed on theoretical grounds.	Fit: R^2 is greater than 0.96 in the unrestricted model. Negative cost-output elasticity and marginal costs at the means of the variables for one output.	Economies are measured as one minus the cost-output elasticity. At the means, this equals 0.360 indicating unexploited economies. No statistical tests.	None.
Cowing & Holtmann (1983)	138 observations for 107 parameters. Short-term, general care New York hospitals in 1975. Teaching and non-teaching, proprietary and non-profit. Mean emergency visits 16563, med. / surg. days 62078, maternity days 4064, paediatric 4347.	Use of a short-run theoretically consistent cost function with which the long-run equilibrium assumption can be tested.	Structural short-run translog cost function. 5 outputs - 5 input prices. Fixed inputs are the number of admitting physicians and capital stock. Also dummies for the type of ownership and teaching status.	Unit of measurement: The patient day and the visit. Outputs: medical-surgical, maternity, paediatrics, other-inpatient and emergency room care. Quality of output proxied by teaching and profit-status dummies. Efficiency: Cost-based payment regime.	Fit: No measure. Regularity conditions not satisfied away from the means. There are negative marginal costs at low levels of output and implausible fitted costs. At the means, one cost-output elasticity is negative but insignificantly so.	Ray economies computed as one minus cost-output elasticities. At the means, short-run economies of scale are present (0.14). For higher levels of output diseconomies are implied, supporting a U-shaped short-run AC curve. No statistical test is performed.	Computed as the difference between sum of separate costs and joint costs as a proportion of joint costs. Economies suggested for [paediatric care, other services] possible diseconomies for [emergency room, other]. No tests.
Conrad & Strauss (1983)	114 North Carolina hospitals in 1978. 36 parameters to be estimated. Mean bed size is about 181.	Estimation of a theoretically consistent multi-product cost function.	Structural long-run translog cost function. 3 outputs - 4 input prices (including price of capital).	Unit of measurement: Patient day. Outputs: child days, non-Medicare days, Medicare days. Efficiency: claim that in cost-based regimes, auditing suffices for cost-minimisation.	Fit: No R^2. Regularity conditions not satisfied away from the means. There are negative marginal costs and implausible fitted costs at low output levels.	Test for homogeneity of degree one (restricted model). Constant returns to scale, neither economies nor diseconomies.	None.

Table 2: Flexible econometric cost studies (Cont'd)

Author Year	Sample Characteristics	Study Objectives and Methodology	Specification of Cost Function(s)	Output, Quality, Uncertainty, Efficiency Adjustment	Model Diagnosis	Findings on Economies of Scale	Findings on Economies of Scope
Banker, Conrad & Strauss (1986)	The same sample as in Conrad and Strauss (1983). Cross-section of 114 North Carolina hospitals in 1978.	Comparative application of an econometric model using the translog cost function and a Data Envelopment Analysis using the same data set.	Structural (frontier) long-run translog joint cost function. 3 outputs - 4 aggregated input prices (including price of capital).	Unit of measurement: Patient day. Outputs: constructed by classifying patient days according to (patients') age.	The same fit as in Conrad and Strauss (1983).	The Translog cannot reject the null of constant returns to scale. In DEA the "most productive scale size" indicates both increasing and decreasing returns to scale in different segments of the production correspondence.	None.
Grannemann, Brown & Pauly (1986)	867 observations for 64 parameters. Non-federal, short-term US hospitals (profit, non-profit, affiliated or not, in metro- or non-metropolitan areas) in 1981. Means: Total inpatient days 71260, acute days 63180, intensive care days 4420, outpatient visits 24910, emergency visits 19940.	Development and estimation of a behavioural model with ad hoc and flexible elements. Not only the cost function has a hybrid functional form but also includes behavioural variables which characterised ad hoc models. Economies are assessed.	Cubic long-run cost function combining features of flexible (e.g. cross-output terms) and ad hoc (e.g. no output-price interaction terms). Many outputs - 4 input prices (price of capital excluded). Behavioural variables include profit status & revenue sources, per capita income.	3 categories for inpatient days and 3 for inpatient discharges (acute, intensive care, subacute & other). Outpatient, emergency, physical therapy, home care, family planning visits, ambulatory procedures. 11 case-mix proportions. Two teaching dummies. Quality: Variables for use of high tech equipment, teaching dummies, for-profit dummy.	Fit: R^2 is 0.9267. The regularity conditions are not checked away from the means despite the computation of global measures of economies of scale and scope.	Product-specific economies of scale are found - at the means - for the emergency department but constant returns to scale for other outpatient departments. No formal statistical tests of significance.	Diseconomies of scope between emergency department and inpatient care.
Lee (1987)	New York, New Jersey and Maryland teaching and non-teaching hospitals. Pooled data over 1979-80, 1982-84. Sample size not reported but must be large enough (initially used to estimate a flexible model with 23 outputs). Mean bed size is 290.95.	Comparative analysis of the theoretical properties and empirical estimates of three flexible cost functions. Models used to derive measures of short-run economies of scale and scope and check robustness to model specification.	Three flexible short-run cost functions (translog, hybrid translog, CES translog). 8 outputs - 4 labour input prices. Control variables for States and years of observation, for outliers, for teaching status. Fixed input is «beds».	Units of measurement: case and visit. Outputs: aggregate categories defined by the DRG classification scheme. 2 of these represent outpatient services. Quality: Teaching dummies. Efficiency: Prospective Payment System.	Fit: Error sum of squares is 3990 (translog), 3999 (hybrid translog), 3970 (CES translog). The condition most often violated in practice (concavity in input prices) is not met in all 3 models at each sample observation. There are also three negative cost-output elasticities and marginal costs for 3 out of 8 outputs.	Hybrid translog and CES translog models slightly superior on stat/cal grounds. Hybrid model indicates diseconomies which may be due to the Box-Cox metric. The CES and translog estimates reveal significant short-run (computed without use of fixed input) scale economies at the means (0.22 in the translog).	Some cross-output terms significant at 10% (none at 5%). Significant economies for [diseases - disorders of eyes, other inpatient care], [mental disorders, substance use disorders], [outpatient visits, burns]. Diseconomies for [substance use disorders, other inpatient care], [outpatient visit, emergency room visit].

Table 2: Flexible econometric cost studies (Cont'd)

Author Year	Sample Characteristics	Study Objectives and Methodology	Specification of Cost Function(s)	Output, Quality, Uncertainty, Efficiency Adjustment	Model Diagnosis	Findings on Economies of Scale	Findings on Economies of Scope
Ahern (1988)	Financial years 1981-82 and 1985-86 cross-sections of 290 and 274 California acute hospitals. These are non-profit and investor-owned. All produce medical / surgical and outpatient services. Outliers (w.r.t. input - output levels) excluded.	Use of the 1981-82 cost data (when cost-based reimbursement was in effect) and 1985-86 data (when Medicare DRG reimbursement was phased) to assess the impact of DRG reimbursement on hospital costs.	Translog short-run (includes capital and number of physicians) cost function. 4 outputs - 4 input prices. Other controls: ownership dummy, Herfindahl index, case-mix, dummies relating to payer distribution.	Unit of measurement: Discharges for inpatient care, visits for outpatient. Outputs: medical-surgical acute discharges, outpatient visits, total laboratory units, other ancillary procedures and treatments. Medicare case-mix index. Quality: ownership dummy. Efficiency: Derive economies from pre-, post-DRG equations.	Fit: No R^2. In general estimation results were plausible. Positive significant first-order terms.	Computed without the use of the fixed factors. Overall significant scale economies in the 1981-82 (1.18) and 1985-86 (1.22) equations at sample means. Results are inconclusive with respect to product-specific economies of scale.	None.
Eakin & Kniesner (1988)	Cross-section of 331 hospitals in US in 1975-76. These are teaching and non-teaching hospitals. Sample means are not reported.	Estimation of a non-minimum cost function which incorporates the minimum cost function as a limited case. Economies of scale are estimated in both models and are compared.	Non-minimum, long-run, hybrid translog cost function which allows for systematic allocative inefficiency. 4 outputs - 4 input prices, with proxies for physician price and the cost of capital. Teaching dummy, average length of stay and occupancy rate.	Unit of measurement: The hospital case and the visit. Outputs: general medicine, obstetrics and gynaecology, outpatient visits and weighted surgery (cases weighted by time under anaesthesia). Average stay to proxy case-severity and a teaching dummy. Quality: teaching dummy. Uncertainty: occupancy rate. Efficiency: Model allows for inefficiency.	Fit: Generalised $R^2 = 0.956$. All regularity conditions are met at the sample means and in 86% of the observations.	Allocative inefficiency is 5% of total observed costs but does not influence estimates of scale effects. Significant ray diseconomies are found at the sample means. In the non-minimum cost function they equal 1.358 (t-stat 8.70) and in the minimum cost function 1.341 (8.59). Half of the observations experience economies, half diseconomies.	None.
Vita (1990)	296 observations for 98 parameters to be estimated. California short-term acute care general hospitals in 1983. These are profit and not-for-profit. Mean bed size is 178.17 beds.	Assessment of the validity of estimates from a flexible functional form outside the neighbourhood of the approximation point. Derivation of long-run economies of scale from short-run cost functions.	Behavioural generalised (Box-Cox) translog short-run cost function with a few ad hoc elements. 5 input and 5 output categories, average length of stay (in the 5 categories), case-mix index and dummies for system membership & profit status.	Unit of measurement: discharged case and average length of stay for inpatient care and the visit for outpatient care. Case-mix index. Outputs: medical/surgical, obstetric, outpatient and emergency room, all other. Quality: profit status dummy.	Fit: No R^2. Positive cost-output elasticities and marginal costs at the means. Factor shares are positive as indicated by the first-order price coefficients. Own price elasticities of the factor demands are all negative.	Absence of ray scale economies at various output vectors. At the means, scale diseconomies exist, computed with the use of the actual level of the fixed input (0.79). No. tests provided. And the true long-run economies may be higher than 0.79 if we believe that the actual number of beds exceeds its optimum.	Insignificant weak cost complementarities at the sample means. Statistical tests are provided.

Table 2: Flexible econometric cost studies (Cont'd)

Author Year	Sample Characteristics	Study Objectives and Methodology	Specification of Cost Function(s)	Output, Quality, Uncertainty, Efficiency Adjustment	Model Diagnosis	Findings on Economies of Scale	Findings on Economies of Scope
Pangilinan (1991)	924 observations for 56 model parameters. Pooled data from 132 New York State hospitals over 1980-86. These are general hospitals, teaching and non-teaching, proprietary and public, urban and rural. Mean bed size is 210.68.	Study of the production and cost inefficiency in hospitals. Different aspects of efficiency examined. Technical efficiency measured by fixed and random effects models. Scale and scope efficiency via the estimation of a flexible cost function.	Short-run translog cost function employing the Vita (1990) two dimensional definition of output and its ad hoc elements (e.g. no interaction between input prices and outputs). Uses Box-Cox transformation for outputs. 5 discharges variables (+ 5 for length of stay) - 3 input prices. Variables for ownership type, location and year of observation.	Unit of measurement: discharged case and length of stay for inpatient care. Outpatient services ignored. Outputs: medical / surgical, paediatric, maternity, nursery, psychiatric & rehabilitative services. Case-mix index, teaching dummy. Quality: Some dummies and a case-mix index. Efficiency: Hospitals may adjust their behaviour in anticipation of PPS fully implemented in 1988.	Fit: System weighted R^2 is 0.909. Positive marginal costs, cost shares of factor inputs and negative own-price elasticities for factor demands.	Discharges used in calculations. Short-run economies computed as the inverse of the sum of cost-output elasticities. At the means they equal 1.649 suggesting large scale economies. Also computed using the actual level of beds. They equal at the means 0.855 indicating slight diseconomies. True long-run economies may be higher if the actual number of beds exceeds the optimal one.	Measured by weak cost complementarities at the means. Analysis suggests that it may be more likely for economies to occur for pairs [paediatric, maternity], [paediatric, nursery] than for [med. / surgical, paediatric]. Yet findings cannot validate the presence of any economies.
Custer & Willke (1991)	564 observations for 76 parameters. US short-term, acute-care teaching hospitals in 1986. Most are private non-profit, a few for-profit and the other public. Mean bed size is 443.38.	Assessment of the impact of medical staff characteristics on the cost of inpatient care and medical education.	Behavioural Grannemann et al. cubic total cost model. 5 - output, 1 - input price long-run model assuming constant cost of capital. Many behavioural variables including medical staff characteristics.	Units of measurement: as below. Outputs: discharges, patient days, outpatient visits, excess capacity and number of residents. Case-mix index. Quality: Medical staff characteristics proxy it.	Fit: Adj. R^2 ranges from 0.938 to 0.958 (unrestricted model with staff variables). Plausible first-order (positive) parameters on outputs and price.	Coefficients show a U-shaped AC curve. Product-specific economies implied for all outputs. No statistical tests.	Global measure. Economies of scope between each output and all others but no statistical tests performed.
Kemere (1992)	250 observations from pooling 50 Maryland acute care short-term hospitals (most non-profit) over 1981 to 1985. Mean bed size is 298.	Comparative analysis of several cost determinants via both long- and short-run flexible cost functions.	Long- and short-run translog cost functions. 4 outputs - 6 input prices. Hospital characteristics: medical school affiliation, length of stay, urban / rural dummy, capacity utilisation, availability of CAT scanner, indigent care provided.	Units of measurement: cases and visits. Outputs: paediatric, adult, geriatric inpatients, outpatient visits. Quality: Some of the hospital characteristics included may proxy it. Efficiency: Prospective Payment System.	Fit: No R^2. All regularity conditions are satisfied at the means. Some were satisfied even away from the means. Yet concavity of input prices and positive marginal costs not examined away from the means.	Long-run function: Economies (one minus sum of output-cost elasticities) equal 0.01694 indicating increasing returns which are insignificant. Product-specific diseconomies for all outputs. Short-run function: Computed as above they equal 0.09182 indicating significant short-run economies. Computed with actual level of fixed factor they are absent (0.9835).	Measured by weak cost complementarities at the means. Long-run function: significant economies for pairs [outpatient visits, paediatric output], [adult output, geriatric output]. Short-run: economies for [adult, geriatric].

Table 2: Flexible econometric cost studies (Cont'd)

Author Year	Sample Characteristics	Study Objectives and Methodology	Specification of Cost Function(s)	Output, Quality, Uncertainty, Efficiency Adjustment	Model Diagnosis	Findings on Economies of Scale	Findings on Economies of Scope
Fournier & Mitchell (1992)	534 pooled observations from 179 short-term general teaching Florida hospitals over the period 1984-86. Speciality hospitals excluded. Four ownership types: not-for profit, government, investor-owned independent, chains. F-statistic validates pooling.	Estimation of the effect of market competition on hospital costs, after controlling for other cost determinants. Also attempt to calculate t-statistics for the scale and scope economies estimates.	Generalised (i.e. Box-Cox) translog short-run cost function. Estimation of a structural-oriented translog and a behavioural one which includes market structure and ownership status variables. 5 outputs - 3 input prices. Capital stock and number of physicians are the fixed inputs.	Units of measurement: see below. Outputs: acute & intensive care admissions, outpatient visits, emergency room visits, no of maternity procedures, surgery minutes. Teaching output by no of residents. Case-mix index and % of patients in intensive care. Quality: Teaching, urban dummies, no of physicians, case-mix index. Efficiency: Prospective Payment System in effect.	Fit: Small Root Mean Square Errors for both models (0.150-0.169). Cost-output elasticities positive. Other coefficients plausible.	Ray economies computed without the use of the fixed input. They equal 1.23 (t-statistic 2.91) indicating significant economies at the means. Product-specific economies of scale exist for outpatients (1.18), maternity (4.91), emergency (1.49) and surgery (1.79). Although these are insignificant, this may be due to the approximate nature of the t-statistics.	Global measure of economies. There are significant economies for [outpatient, others], [maternity, others], [emergency, others], [surgery, others], [outpatient & maternity, others], [outpatient & emergency, others], [maternity & emergency, others].
Ablett (1993)	140 observations from public and private, teaching and non-teaching Belgian hospitals for 1988 and subsample of 64 hospitals having positive all their outputs (but which are larger in size). Mean bed size is not reported.	Application of a flexible cost function to assess the presence of short-run economies of scale and economies of scope.	The Cowing and Holtmann (1983) specification with input prices omitted (constant across hospitals in Belgium). Dummies: availability of diagnostic imaging, teaching, private / public status.	Unit of measurement: the patient day. Outputs: surgical, medical, paediatric, maternity and other services (all in days). Quality: dummies may roughly proxy it. Efficiency: Prospective budget.	Adjusted R^2 is 0.915 for the sample of 140 hospitals and 0.920 for the subsample. An F-statistic indicates that the regression is significant. The function is irregular even at the sample means (negative marginal costs for one output).	Overall short-run economies of scale are found (0.63: computed as one minus the cost-output elasticities i.e. without the use of the actual level of the fixed factor) at the means. No significance test is performed.	Evidence of economies for all combinations except for the bundling of surgical services with any of the remaining services (here diseconomies exist). Yet no statistical tests are performed.
Banks (1993)	Two samples employed since an F-test shows differences in estimates. 193 non-profit and 159 proprietary short-run acute care California hospitals in 1981. 41 model parameters to be estimated. No sample means are reported.	Incorporation of socio-economic variables into the analysis of hospital cost behaviour. These along with other variables are used to measure the ultimate hospital output, namely health status improvement.	Long-run flexible quadratic approximation as the cost equation. Improved health status is unobserved. Its determinants - case-mix, insurance cover, teaching responsibility, socio-economic environment - are included as regressors. 3 labour input prices.	Unit of measurement: Patient day and visit. Outputs: Acute, intensive care, emergency room, clinic visits the 4 service outputs. Facilities index proxies case-mix. Two socio-economic indices. Insurance cover and teaching index also affect outcome and are included. Quality: Some of the above capture crude differences. Efficiency: Cost-based reimbursement in effect.	Fit: Adj. R^2 is 0.959 (non-profits), 0.949 (proprietary). The model parameters indicate theoretical consistency. Yet the quadratic form used does not impose one condition (linear homogeneity in input prices).	Computed from a long-run function. Significant (at the 1% level) economies of scale at the sample mean levels for both non-profit (1.10) and proprietary (1.23) hospitals. Further, if non-profit hospitals were to adopt proprietary hospitals' output-mix (while maintaining their own marginal cost estimates) they would on average exhibit enhanced (1.41) economies.	Measured by weak cost complementarities. No significant complementarities found suggesting that we cannot validate the existence of economies of scope.

Table 2: Flexible econometric cost studies (Cont'd)

Author Year	Sample Characteristics	Study Objectives and Methodology	Specification of Cost Function(s)	Output, Quality, Uncertainty, Efficiency Adjustment	Model Diagnosis	Findings on Economies of Scale	Findings on Economies of Scope
Okunade (1993)	126 observations for 34 parameters. US 1981/9 time-series data, consisting of seven cross-sectional bed class sizes for two types of hospitals. One sample for general and specialised hospitals since pooling is justified by an F-test. Mean patient days are 7061791.	Estimation of the structure of hospital pharmacy production, using a dual cost function approach. Economies of scale, input substitution elasticities, and total factor productivity growth are assessed.	Structural single-output, multiple-input short-run translog cost function. A time trend included to capture technical change along with seven bed size variables. 4 input prices. Omission of fixed capital justified since pharmacies are not capital-intensive.	Output: hospital patient days proxy the volume of hospital pharmacy output, since the latter is known to be correlated with. the former.	Fit: system R^2 is 0.99. D-W statistic is 1.90. The regularity conditions are met at the means and for most of the observations.	Short-run economies computed as unity minus the cost-output elasticity. Slight but statistically significant short-run diseconomies of scale are found at the sample means (-0.01: t-value -2.50). The most efficient operating size in the short-run is the median bed size category i.e. 200-299 beds.	None.
Collins (1994)	49 observations from acute care Maryland hospitals in 1990. These are profit, non-profit and teaching, non-teaching. The mean bed size is 303.	Estimation of a cost function with ancillary and overhead services included. Comparison with a model using conventional output aggregation. Also comparative estimation of a structural and an augmented model.	A purely structural and an augmented short-run translog cost function. 4 outputs - 3 input prices. Augmented model includes also Medicare case-mix and Herfindahl index of market concentration.	Unit of measurement: The inpatient day. Outputs: inpatient days, outpatient, ancillary, overhead services. Medicare index to capture differences in illness severity. Efficiency: Maryland's all payer rate regulation may increase probability of cost minimisation.	Fit: No R^2. Collinearity problems indicated by insignificant t-ratios and a negative first-order output co-efficient (may also be due to aggregation). Bivariate correlations 0.921 - 0.983. Partial correl. coeff. 0.885 - 0.976. Negative cost-output elasticities (and marginal costs) for 3 outputs. Residual variance shows homoscedasticity. Endogeneity: length of stay may be dictated by health plan so that patient days are exogeneous.	The preferred augmented model reveals slight diseconomies, at the sample means, when these are computed as the inverse of cost - output elasticities (-0.022199) and when the actual level of the fixed input is used (-0.02007).	Measured by weak cost complementarities. There are no economies of scope in hospital production.
Sinay (1994)	Pooled data for two years used in equations. Functions estimated for merged and non-merged hospitals one year prior, and one - two years after the merger. Two merger episodes: 1987-88, 1989-90. Mean bed size is 200-225 for non-merged, 183-229 for merged (prior to merger), 313-436 (after).	Test of the hypothesis that hospital mergers in the 1980s reduced production costs by achieving economies of scale and scope. Cost functions are estimated before and after merger episodes for merged and control hospitals.	Short-run translog cost function including beds as the fixed factor. 4 outputs - 2 input prices. Dummies for membership in a system and for-profit status.	Unit of measurement: a) Inpatient day b) Outpatient visit. Outputs: acute, intensive, subacute care days, outpatient visits. Quality: very rough differences may be captured by dummies. Efficiency: Increased price competition and Prospective Payment Scheme since 1983.	Fit: R^2 ranges from 0.982 to 0.992. At the means, positive cost-output elasticities and cost shares of the factor inputs.	Ray economies computed using the actual level of the fixed factor. No statistical tests. Merged hospitals eventually achieve scale efficiency two years after the merger, whereas control hospitals exhibit no change in efficiency over these years. Prior to the first (second) merger episode, merged (to be) hospitals exhibited diseconomies (economies).	Weak cost complementarities with stat. tests. Economies exist prior to first merger episode for merged hospitals between acute and subacute care services, providing reason for consolidation.No economies for merged and controls in second episode.Economies may be more likely where there is excess capacity (e.g. ray diseconomies in first episode).

Table 2: Flexible econometric cost studies (Cont'd)

Author Year	Sample Characteristics	Study Objectives and Methodology	Specification of Cost Function(s)	Output, Quality, Uncertainty, Efficiency Adjustment	Model Diagnosis	Findings on Economies of Scale	Findings on Economies of Scope
Gruca & Nath (1994)	96 observations for 56 parameters to be estimated. Short-term, acute care Chicago hospitals in 1986. Federal govnt, speciality and rural hospitals excluded.	Test relationship between adaptation constraints (including those from scale-scope disadvantages) and hospital failure in the post-PPS environment. Use cost and survival analysis.	Structural short-run translog cost function. 5 outputs (plus teaching dummy) - 3 input prices (including price of drugs and supplies). Book value of fixed assets is the capital variable.	Unit of measurement: Patient day and the visit. Outputs: general medical (acute), surgical (intensive care), paediatric, obstetric, emergency room and outpatient. Teaching output by dummy. Quality: Teaching dummy. Efficiency: Prospective Payment System in effect.	Fit: System weighted R^2 is 0.934. All first-order output and price terms significantly positive at 10%.	Computed as the inverse of cost-output elasticities without the use of the actual level of fixed input (capital). At the sample means they equal 1.19 indicating short-run ray economies of scale. No statistical tests performed. The finding is reinforced by a form of survival analysis.	At the sample means they can be measured by cost complementarities. Significant at the 5% exist only between medical and obstetrics care (-0.0741385). Finding reinforced by survival analysis. Insignificant economies for [medical, paediatric], [medical, emergency], [paediatric, emergency].
Gaynor & Anderson (1995)	26323 observations from pooling over 5000 US hospitals over 1983-87. These are teaching and non-teaching, public and private, for- and non-profit. Mean bed size is 170.16.	Theory of cost reformulated to account for uncertain demand of hospital services. Comparative estimation of cost functions including and omitting uncertainty.	Short-run translog cost function incorporating "beds". 2 outputs - 1 input price (proxy). Augmented not only for uncertainty variables but also for occupancy rate, case-mix, dummies for hospital type, hospital and time fixed effects.	Unit of measurement: a) admitted case, b) visit. Outputs: the above and a case-mix index. Uncertainty: demand distribution parameters enter the cost function. Quality: dummies crudely adjust. Efficiency: Prospective Payment System.	Fit: R^2 is 0.706. Variables have the signs and significance expected. Positive marginal costs. Heteroscedasticity: Present, so consistent standard errors are derived. Serial correlation: Fixed effect estimation. Endogeneity: instrumental variables employed.	Computed using the actual level of fixed factor (beds) with no statistical test performed. At the means they equal 1.11 indicating economies.	None.
Scuffham, Devlin & Jaforullah (1996)	Cross-section of 67 New Zealand public hospitals of various types (teaching, maternity, psychiatric, geriatric, base, small general, country) in 1987. Mean bed size is 125.54.	Computation of short- and long-run estimates of economies of scale, economies of scope, input substitution possibilities and marginal costs.	Structural short-run translog variable cost function (beds included). 3 outputs and 5 input prices.	Two - dimensional definition of inpatient output: cases (admitted) and average length of stay. Outpatient output measured as visits.	Fit: R^2 equals 0.986. Estimated function is well behaved satisfying the regularity conditions at the means. Evaluated at different levels of output shows that cost increases with output.	Computed using the actual level of capital they equal 1.090 indicating an absence of economies or diseconomies. Short-run estimate is 0.349 indicating substantial economies.	Measured by weak cost complementarities at the means. Economies are absent at this level of aggregation. No stat tests performed.

Table 2: Flexible econometric cost studies (Cont'd)

Author Year	Sample Characteristics	Study Objectives and Methodology	Specification of Cost Function(s)	Output, Quality, Uncertainty, Efficiency Adjustment	Model Diagnosis	Findings on Economies of Scale	Findings on Economies of Scope
Wagstaff and Lopez (1996)	Panel data for 43 Catalan hospitals over the period 1988-91. These are public and private hospitals contracted to provide care to the public sector.	Estimation of inefficiency, economies of scale and economies of scope in Catalan hospitals.	Frontier version of the Granneman et al. long-run cubic cost function. Basic output categories include inpatient cases, ambulatory and emergency visits. Other outputs variables are also present (eg teaching, rehabilitation programmes, operating theatres) and a case-mix index is included along with the availability of CAT scanners. No input price variables.	Units of measurement are mainly the case and the visit. Outputs: 3 main outputs along with a case-mix index which is based on the US Patient Management System that groups discharges according to their resource intensity. Quality: Teaching dummy, availability of CAT scanners can be seen as very rough proxies.	R^2 is 0.898. Linear homogeneity in prices is not imposed since price data were not available. Almost all of the estimated parameters are of the expected sign.	Evaluated for the average hospital at the cost frontier. There are ray economies of scale, product-specific economies for emergency visits and inpatient care and diseconomies for ambulatory care.	Economies of scope computed by comparing predicted costs if no service was produced with those a hospital at the frontier would incur if it offered all outputs. Mild economies found. Complementarities between ambulatory and emergency care can offset the anti-complementarities between inpatient and outpatient care.

Table 3: Econometric production function studies

Author Year	Sample Characteristics	Study Objectives and Methodology	Specification of Production Function(s)	Model Diagnosis	Findings on Returns to Scale
Feldstein (1967)	Cross section of 177 NHS short-term acute care hospitals.	Estimation of a Cobb-Douglas, a mixed Leontief-Cobb-Douglas (assumes potential substitution only between some inputs) and a more general ad hoc model.	Output is either standardised in the LHS by weighting nine case-mix proportions by their average cost or in the RHS by the inclusion of nine variables. As is typical in production studies some of the four inputs included are measured in value rather than physical terms. Inputs: beds, doctors, nurses and other supplies.	Cobb-Douglas: R^2 is 0.86-0.91 depending on the output specification. The parameter estimates for some inputs are unacceptably low indicating mispecification. Mixed Cobb-Douglas-Leontief: more plausible input parameters.	Cobb-Douglas: Constant or slightly decreasing returns. Similar results using instrumental variables rather than OLS estimation. Mixed Cobb-Douglas Leontief: Decreasing returns. More general model confirms the findings.
Lavers & Whynes (1978)	193 English maternity hospitals in financial year 1971-72. The mean bed size is just under 40.	Empirical assessment of the structure of English maternity hospitals using a Cobb-Douglas and a flexible Log-Quadratic approximation models. Focus on maternity hospitals is meant to avoid adjustment for output heterogeneity.	Output is thought to be homogeneous in maternity hospitals so that the dependent variable is the number of cases discharged. Sharing the limitation of Feldstein's work, the authors use costs as proxies for actual inputs. Inputs: occupied beds, annual expenses on nurses, doctors, drugs and dressings.	R^2 in Cobb-Douglas is 0.91. All marginal products positive at the means. R^2 in flexible quadratic function is 0.94. Negative marginal product only for doctors (in one equation medical staff is excluded).	Cobb-Douglas model: Sum of output elasticities is 0.97 and constant returns are found. Log-Quadratic model: Slightly (but statistically significant) decreasing returns to scale.
Brown (1980)	Pooled data from 12 Newfoundland's (Canada) cottage hospitals, 1952-1958, 1971, 1976. 24 observations in the cost equation. Mean bed size is about 27-30.	Estimation of production functions for Newfoundland hospitals which have salaried physicians so that physician time input can be included. Product- mix is also fairly constant so that adjustment for output is simpler.	Cobb-Douglas functional form. The dependent variable is either patient days or a composite output measure derived by weighting patient days, outpatient services, laboratory and X-ray services by hypothesised time input requirements. Inputs are beds, number of doctors, number of nurses and technicians, and number of other workers. A shift variable capturing productivity change over time and a time trend variable also incorporated.	R^2 is 0.75 for the pooled data.	Significant increasing returns to scale over the sample range.
van Montfort (1981)	100 Dutch general hospitals in 1971. No means reported.	Estimation of a Translog function which better fits the data than the Cobb-Douglas or CES translog.	Translog estimated by OLS. Dependent variable is weighted admissions. The estimated parameters from past cost functions make up the weights. Inputs include the number of specialists, registered nurses, student nurses, other nursing staff, paramedical staff, the number of beds, a facility index, and the monetary value of drugs-dressings and of other medical-paramedical devices.	Adjusted R^2 is 0.93. Positive output elasticities, which (given positive mean output and input levels) imply positive marginal product at the means.	Increasing returns are implied by the sum of output elasticities (1.04), which are nevertheless insignificant. The same result is obtained in a specification which uses intermediate outputs as the dependent variable.

Table 3: Econometric production function studies (Cont'd)

Author Year	**Sample Characteristics**	**Study Objectives and Methodology**	**Specification of Production Function(s)**	**Model Diagnosis**	**Findings on Returns to Scale**
Jensen & Morrisey (1986)	National data from short-term general US hospitals in 1983. The two samples consist of 3540 non-teaching and 628 teaching units. Separate regressions are run for the two samples. Mean bed size is given in logs.	Estimation of a translog production function which incorporates the physician input. The role of the latter in allocating resources and influencing the productivity of other medical resources is thus accounted for.	Translog estimated by OLS since studies have shown that OLS bias is negligible. Dependent variable is adjusted admissions (by means of a case-mix index). Inputs are medical staff, nursing staff, other non-physician staff and beds. In the teaching specification medical residents (and membership in the Council of Teaching Hospitals) is included. Also: an index of presence of competing hospitals, ownership and location dummies.	All production functions are fairly well-behaved in that marginal products for all inputs are positive and decreasing. Cross-sectional R^2 is 0.95.	The authors do not provide estimates of returns to scale. Nevertheless we calculated them. The sum of output elasticities for the non-teaching and teaching sub-samples are 0.858 and 0.952, both below unity indicating decreasing returns for the average hospital.
Pangilinan (1991)	Panel data from 142 New York State hospitals over 1981-1987.	Fixed-effect Cobb-Douglas production model.	Dependent variable is total inpatient days. Inputs: buildings and fixtures, utilisation rate, full-time equivalent labour. Physician input time is not included. Case-mix index.	R^2 is 0.60. A RAMSEY test rejects the null hypothesis of no misspecification	Significant increasing returns to scale but they might be due to the omission of physician time.

Table 4: Data envelopment analysis

Author Year	Sample	Model description	Adjustment for outputs, inputs, quality, uncertainty, environment	Treatment of specification, measurement, sampling error	Findings on returns to scale and optimal size
Banker, Conrad & Strauss (1986)	114 North Carolina hospitals in FY 1978.	Introduction of the Variable Returns to Scale (VRS) DEA model which does not impose constant returns, and which permits the calculation of the most productive scale size.	One State chosen to control for regulatory differences. Unit of measurement: The inpatient day. 3 output variables according to the patients' age. 4 inputs including capital but excluding physicians.	No.	Different behaviour in different segments of the production correspondence. For hospitals having a proportion of patients over 65 that is not very high (<44%) the optimum is in the 223-260 bed size range. For geriatric care it is 110 beds and for paediatric 99.
Maindiratta (1990)	55 hospitals randomly chosen from the sample used by Banker, Conrad and Strauss (1986).	An extended VRS DEA model which not only checks whether input savings could have been effected by a hospital given its observed task, but also whether even greater savings would be possible if the task was to be optimally apportioned to smaller units. The notion of size efficiency is introduced to measure this potential for further savings.	Data from one State are used to control for regulatory differences. Unit of measurement: The patient day. 2 outputs - surgery days and other days. 3 variable inputs excluding capital.	No.	Largest-size efficient scale is higher than the mpss. Decreasing returns may set in very gradually (justifying the Banker et al. (1986) translog results) so that a hospital must be a lot bigger than its mpss before it pays to apportion its task to smaller units.
Valdmanis (1992)	Large (over 200 beds), urban , public and private not-for-profit hospitals in Michigan during 1982.	A VRS DEA model similar to the Fare, Groskopf and Lovell (1985) program which decomposes overall technical efficiency into pure technical and scale components.	One state chosen to control for regulatory differences. Large, urban units to control for environmental differences (e.g. case-mix). Unit of measurement: The inpatient day, surgery minutes, outpatient visits. 3-5 outputs defined according to age or type of treatment (including outpatient care). Inputs include physicians and capital.	Yes. Sensitivity analysis is performed to check the robustness of findings to morel specification. Ten of them used employing various input / output variables. Results are robust but not all feasible specifications are checked.	Mean scale efficiency for public hospitals is 0.97-1.00 and mean bed size is 350. For private not-for-profit it is lower, 0.92-0.97 with a mean bed size of 428. This could imply an optimum of 300-400 beds.
Dervaux et al. (1994)	193 French public hospitals with short- and long-term care facilities. Some must be teaching since one input refers to the teaching staff.	The paper is ultimately concerned with the construction of a Malmquist productivity index. Both VRS and CRS DEA programs are estimated.	Unit of measurement: Admission for short-term care, patient day for longer-stay. Three output specifications. The first two have 6 output variables but differ on the aggregation level for short- vs. long-term care. The third includes 3 more variables (e.g. complexity of case-mix, intensity of treatment). 4 labour and capital inputs.	Only crude sensitivity analysis. Three output specifications used. Results are unchanged with respect to the first two specifications. The third, which is preferred shows a different optimum.	Three specifications used. First two show a mean optimum at 496-522 beds. The third, capturing more hospital activity aspects indicates a mean optimum of 623 beds.

Table 4: Data envelopment analysis (Cont'd)

Author Year	Sample	Model description	Adjustment for outputs, inputs, quality, uncertainty, environment	Treatment of specification, measurement, sampling error	Findings on returns to scale and optimal size
Byrnes & Valdmanis (1994)	123 community (non-teaching) not-for-profit California hospitals in FY 1983.	Four DEA-type linear programming programs for each hospital are used. Six measures of relative efficiency are computed, namely allocative, pure technical, input congestion, scale, and overall efficiency measures.	Control for organisational differences (e.g. case-mix) by omitting teaching hospitals and for differences in amenities (quality) by excluding public ones. Unit of measurement: discharge to avoid endogeneity. 3 outputs (medical / surgical, intensive care, maternity) - 6 inputs (5 labour, and 1 capital).	No.	33 hospitals with a mean size of 202 beds have increasing returns, 49 with 230 beds have constant, and 41 with 366 mean bed size have decreasing returns. Optimum is at about 230 beds.

Table 5: Survival analysis

Author Year	Study Methodology	Description of multi-variate models	Findings on economies of scale and optimal hospital size
Bays (1986)	Traditional survival methodology classifying US hospitals into 8 bed size classes and computing the market share each of these constituted in 1971, 1974, and 1977.	Uni-variate analysis.	Hospitals with less than 100 and more than 500 beds experienced a significant and insignificant, respectively, decline in market shares. All other classes increases. The range of surviving sizes is 100-499 beds.
Mobley (1990)	Use of both traditional survival analysis and multi-variate (binary and continuous) growth models. A sample of US short-term hospitals is used. Among the traditional models the one using bed-days for output is preferred and compares market shares in 1980 and 1986.	The continuous growth model (dependent variable: change in market share in average daily census by size group) is preferred to the binary model (dependent variable equals unity if a bed class shrinks) since it yields jointly significant regressors. Independent variables are average daily census (with a squared term), market share, market extent and market concentration.	Traditional statistical model indicates economies up to about 300 beds, since smaller sized hospitals experienced a decline in their market shares. Continuous model suggests that economies of scale exist up to about 300 beds, and no scale diseconomies for larger hospitals. Yet the confidence interval around this predicted size is wide. And adjusted R^2 in the OLS growth equation is only 0.202.
Vita et al. (1991)	Application of the traditional survival methodology in US hospitals for years 1976, 1981, and 1987.	Uni-variate analysis.	The 1976-87 time period saw significant increases in the percentage of hospitals with 200-399 beds, which is seen as the optimum hospital size.
Lillie - Blanton et al. (1992)	Logistic regression is used to examine ceteris paribus individual factors that affect the risk of hospital closure. The data come from US hospitals in 1985-88. The characteristic of interest is bed size, defined in terms of four categories (6-49, 50-99, 100-199, and 200+). N = 5320.	The selected characteristics - in addition to size - included in the regression are the Medicare case-mix index, % occupancy, % Medicare days, % Medicaid days, availability of long-term units, the number of hospitals in the county, location variables, area wage index, % change in population and unemployment, and per capita income.	Results suggest that smaller hospitals are at greater risk of closing. Hospitals with less than 50 and 100 beds were 13 and 5 times as likely to close as the ones having more than 200 beds. Hospitals in the 100-199 size category are 2.5 times as likely to fail as the 200+. Hence the implication is that the optimal size is greater than 200 beds. Log-likelihood ratio 1471, C = 0.876, Somer Dyx = 0.751.
Mobley & Frech (1994)	Development and estimation of models of growth and failure that control for other demand-related determinants of hospital growth and failure. The aim is to isolate the size effect. Sample from short-term general California hospitals in 1980-89 (a period of increased competition).	Dependent variable is either binary or continuous. Independent variables: (log of) average daily census, six case-mix variables, a teaching dummy, various variables for market extent and competition (e.g. Herfindahl index), and to control for firm-specific factors (e.g. proportion of bad debt and indigent expense to net patient revenue).	The probit survival model s suggests that scale economies are a statistically significant determinant of survival. The growth models (tobit II and continuous OLS model) also imply economies but the coefficient on size is marginally insignificant at 10%. The authors suggest that the optimum is at 325 beds, yet after dropping out case-mix and teaching dummy. Economies are interpreted as of the "size-quality kind". Adjusted R^2 is only 0.206 in the continuous growth model.

Table 5: Survival analysis (Cont'd)

Author Year	Study Methodology	Description of multi-variate models	Findings on economies of scale and optimal hospital size
Lynch & Ozcan (1994)	Logistic regression analysis used to test the hypothesis that inefficient and under-utilised US hospitals in competitive markets were at greater risk for closure in 1988.	Rather than employing crude measures of efficiency, it employs a Data Envelopment Analysis measure as a predictor of failure. The index plus measures of hospital utilisation, competition, size (beds), and utilisation by Medicaid and Medicaid patients enter the regression.	Coefficient on size variable significant and negative indicating that larger hospitals are less likely to close. Model chi-square = 80.99 with 7 d.f., $p = 0.0001$.
Frech & Mobley (1995)	Application of both traditional survival analysis and multi-variate (binary and continuous) growth models to California short-term general hospitals in the period 1983-89.	Independent variables are output (average daily census), dummy for hospital affiliation (1989), Herfindahl index (1983), and change in local market-level proportion of hospital revenues under discount contracts (1983-89).	Traditional (Stigler-type) statistical model suggests economies up to 400 beds. Both multi-variate models have jointly significant regressors (at 5%), but the continuous one is slightly more powerful (even at 2.5%). Optimal size is between 200 and 370 beds. Adjusted R^2 in the continuous model is 0.136.
Simpson (1995)	Assessment of the bias in survival studies using data from the period prior to 1987 resulting from Certificate of Need (CON) regulation in that era.	The study aims to cast doubt on the finding that smaller (sub-100 beds) hospitals are inefficient. Use of post-1987 data to examine whether small hospitals have a higher rate of entry so that a decline in market share will be partly due to entry restrictions, not solely economies.	After 1987 a larger number of sub-100 bed hospitals entered the market. So, the previously documented decline in market share may be due to entry restrictions preventing the replacement of existing sub-100 bed hospitals with new ones of the same size.

Table 6: Studies examining a hospital service in isolation

Author Year	Hospital Service(s) Examined	Study Methodology	Finding on Economies of Scale
Hu (1971)	Maternity Ward	Maternity wards of 30 non-profit hospitals in Pennsylvania in financial year 1967-68. Quadratic AC function including the number of maternity admissions and the number of maternity patient-days per maternity admission (quality of output measure). No adjustment for output since it is assumed homogeneous within the maternity ward.	R^2 is 0.40. A U-shaped short-run AC curve is implied with economies for the average maternity ward. The optimal number of maternity admissions is 1616 per year.
Anderson (1974)	Clinical Laboratory Hospital Service	44 clinical laboratories in Saskatchewan public general hospitals are viewed as independent production units which provide test results to the medical staff at no charge. The average cost model is a function of scale (labour time), two complexity variables, and utilisation. Prices are dropped out since they are said to be constant across hospitals. The log-log functional form is used in the cost equation.	R^2 is about 0.60. Economies of scale are found in the provision of clinical laboratory services.
McGregor & Pelletier (1978)	Heart Surgery	The study examines whether many small or few large institutions should provide heart surgery services. Attempts this by focusing on a large specialised heart surgery institute with a turnover of 600 operations per year and a small university hospital with 180 operations per year. Identifies and adds the fixed costs with the variable costs to define the cost of operation. The latter is then related to the number of operations per year. Note that variable costs are assumed to be constant i.e. $733 per operation for all levels of output.	The least cost size is about 100 operations per year.
Finkler (1979)	Heart Surgery	A methodology of hospital-cost finding called Component Enumeration (CE) is applied to heart surgery within a single hospital. It identifies the costs of a program that would not be incurred if the program was not offered (i.e. excludes joint costs). The costs of open heart surgery equal fixed costs plus variable costs. The latter are said to be the same per patient at any volume. From this the average cost is related to the number of surgery patients per year.	The least cost size is reached at about 500 procedures per year.
Finkler (1981)	Heart Surgery	This paper replicates the previous CE methodology to a different geographic setting, type of hospital and excludes physician costs from the calculations. The least cost size is again defined to determine the point at which economies of scale are exhausted. Note that again variable costs are the same per patient at any volume.	The least cost size is now 300 procedures per year.
Wilson & Jadlow (1982)	Nuclear Medicine Services	The paper examines the relative efficiency of proprietary and non-profit hospitals in the provision of nuclear medicine services. It is argued that in order to avoid the case-mix problem separate specialised production functions should be defined for each hospital service, one of which is nuclear medicine. The model is a frontier version of the Cobb-Douglas specification. Data come from 922 short-term general hospitals. A composite output variable is derived by weighting each nuclear medicine procedure by an index price. The right hand side of the equation includes labour and capital inputs.	Nuclear medicine services are characterised by increasing returns to scale.
Wan et al. (1987)	Hospital Dental Care Programs	A cost model is applied to (2 year) quarterly data from 25 hospital-based ambulatory dental care programs in the US. Average cost is a function of facility size, ownership type of the clinic, service-mix, input prices, payment-mix, provider-mix and dummy variables capturing quarterly effects.	In this linear model, there are economies of scale for public dental clinics (R^2 is 0.91).

Table 6: Studies examining a hospital service in isolation (Cont'd)

Author Year	Hospital Service(s) Examined	Study Methodology	Finding on Economies of Scale
Munoz et al. (1990c)	Urology Patients	Analysis of the volume of urology procedures performed by an urologist for patients treated over a 3-year period at a large New York academic medical centre. Surgeons were arbitrarily divided into low (treating 5 patients or less within a DRG) or high (eight patients or more during the 3-year period) volume ones. Comparison of hospital cost per patient (adjusted for case-mix and severity-of-illness) in low and high volume urologists for emergency and non-emergency urology patients.	Adjusted cost per patient was significantly higher for low-volume urologists than for high-volume urologists in both cases, i.e. emergency and non-emergency.
Munoz et al. (1990a)	Orthopaedic Surgical Patients	Analysis of the volume of orthopaedic surgical procedures performed by a surgeon for patients treated over a 3-year period at a large New York academic medical centre. Surgeons were arbitrarily divided into low (treating 5 patients or less within a DRG) or high (eight patients or more during the 3-year period) volume ones. Comparison of hospital cost per patient (adjusted for case- mix and severity-of-illness) in low and high volume surgeons for emergency and non-emergency orthopaedic surgical admissions.	Adjusted cost per patient was significantly higher for low-volume surgeons than for high-volume surgeons in both cases, i.e. emergency and non-emergency.
Munoz et al. (1990b)	Neurosurgical Patients	Analysis of the volume of neurosurgical procedures performed by a surgeon for patients treated over a 3-year period at a large New York academic medical centre. Surgeons were arbitrarily divided into low (treating 5 patients or less within a DRG) or high (eight patients or more during the 3-year period) volume ones. Comparison of hospital cost per patient (adjusted for case-mix and severity-of-illness) in low and high volume surgeons for emergency and non-emergency neurosurgical patients.	Adjusted cost per patient was significantly higher for low-volume surgeons than for high-volume surgeons in both cases, i.e. emergency and non-emergency
Fordham et al. (1992)	Neonatal Care	The study covers 17 hospitals in one RHA. Data on the total costs of running Neonatal Units were obtained along with the number of days in intensive and special care. Three simple cost models are employed. The first assumes that average cost is a linear function of cot days. The second includes a quadratic term for days. The third, a variable measuring teaching activity. In all models a case-mix variable (proportion of intensive care cases) is incorporated. Note that it would have been more appropriate to disaggregate cases in two intensive care variables but this was not attempted due to data limitations.	The superior quadratic cost model has an R^2 of 0.756. Economies of scale exist in the provision of both special and intensive care but savings may not be enough to offset the extra costs of travel and access if services were centralised beyond 4000 cot days.
Okunade (1993)	Hospital Pharmacies	The study is considered to an extent more reliable than other studies reviewed at this point and is presented in detail in TABLE 2	See TABLE 2.
Culler et al. (1995)	Knee Replacement Surgery	Estimation of a cost function for knee replacement (KR) surgery, in which the patient is the unit of analysis (267917 observations). The equation includes variables for demographic characteristics (e.g. black) and the health status (e.g. severity, complications) of the patients, hospital characteristics (e.g. wage rate, teaching status and volume and size variables) and environmental factors that affect cost (e.g. dummies for location). The equation is logarithmic with respect to costs, volume and size variables. The latter include second order terms and an interaction variable. There might be large measurement error in the volume measure employed.	Adjusted R^2 is 0.18-0.24 in the various equations. Economies exist up for all but the largest 5% of urban and rural hospitals.
Arndt, Bradbury & Golec (1995)	Cholecystectomy Prostatectomy Hysterectomy Intervertebral Disk Excision	Econometric analysis of the effect of surgeon's volume on hospital average charges, after controlling for factors such as patient age, sex, severity of illness, insurance plan and hospital. One dummy is included for each hospital in order to account for the different charges to cost ratios employed by the 43 Pennsylvania hospitals. Surgeons who perform less than 10 procedures of the specific type are excluded in order to mitigate biases from surgeons being simultaneously low and high volume ones in different hospitals.	Significant negative association between volume and charges for all procedures but hysterectomy. For the latter no significant relationship is evidenced. R^2 exceeds 0.80 in all equations.

APPENDIX I - REVIEW PROTOCOL

Background

There is a voluminous international literature which has explored empirically the impact of the scale and scope of hospital production on unit and total hospital costs respectively. Most of the studies have employed the hospital as the unit of observation in order to measure both scale and scope economies of various provided services. Still other researchers have focused on a particular hospital ward (e.g. maternity ward) or service (e.g. heart surgery) in order to assess in isolation the existence of scale phenomena.

A preliminary literature review was conducted as part of a dissertation for the MSc programme in Health Economics at the University of York, jointly supervised by CRD. The time constraints of that attempt did not permit an exhaustive identification of relevant studies and methodologies. As will be explained, the task of recovering relevant studies is extremely difficult in our context. Moreover, that work had the wider objective of evaluating economies in the health care sector as a whole. The early search for relevant reading material will be narrowed to consider only hospital studies. The work will be updated and extended to retrieve as many existing empirical studies as possible and to develop additional criteria for the assessment of the quality of individual studies.

Questions

The review aims to answer the following questions:

1) Are there overall long-run economies of scale or diseconomies in the production of a typical (average-sized) hospital? At which level of production are economies exhausted and diseconomies start to operate? In other words, can we also identify the optimal size of a hospital in general?

2) Is there evidence to suggest that economies of scale exist for some services whilst diseconomies for other (i.e. are there service-specific economies of scale)?

3) Are there cost savings from bundling some hospital services together? That is, are there economies of scope implying that providing them within the same hospital is cheaper than within more than one hospital units? Is it that for some other services diseconomies of scope may exist?

4) Could the observed relations between scale or scope of production and costs be in fact attributable to bias or other limitations of the techniques employed?

Methodology Of The Review

1 Identification of studies

The following searches were carried out:

1) UK database BIDS ISI (1981-1996),
2) Economic Literature Index (1969-1996),
3) Dissertation Abstracts Database (1968-1996),
4) Medline Express (1980-1996),
5) ABI Inform (1980-1996),
6) Health Planning and Administration (1980-1996),
7) NTIS (1980-1996),
8) Embase (1980-1996),
9) Economics Working Paper Archive at the Internet (econwpa.wustl.edu),
10) DEA WWW bibliography at the Internet (http:\ www.emp.pdx.edu),
11) Two printed bibliographies were also used to trace the early studies (1968 - early 1980s),
12) The Health Econometrics mail-base as well as some CRD contacts,
13) The identified articles themselves were scrutinised to reveal new relevant studies,
14) Recent issues (1995-1996) of three key journals (Applied Economics, Journal of Health Economics, Health Economics) were handsearched.

An important step for the identification of the relevant studies relates to the «key words» employed in the electronic database searches. The actual search strategy is shown in Appendix II.

2 Inclusion criteria

Studies to be evaluated must be both empirical and relevant. A study will be considered to be relevant if it reports empirical findings on economies of scale and/or scope in general hospitals. Or if it reports economies of scale estimates in specialised hospitals, particular wards within a general hospital (e.g. maternity ward) or in a particular hospital service (e.g. heart surgery). Of course there may be some work that focuses on other aspects of hospital performance (e.g. input substitution) which also provides estimates on economies as a by-product. That is, although inference can be drawn from their models the authors do not explicitly discuss, interpret or even measure the phenomena of interest from the estimated parameters. These studies are also relevant but nevertheless more difficult to trace. The types of studies we are looking for are econometric, statistical and mathematical programming analyses of any kind and in any language conducted during the period 1967-1996.

3 Exclusion criteria

Articles which do not report empirical findings on economies of scale or scope in hospitals or in the provision of specific hospital services are excluded from the evaluation. Note that psychiatric hospitals, health maintenance organisations, nursing homes and home health agencies will not be considered to be relevant. In addition, studies which do not attempt to remedy the case-mix effect will be excluded since the consequent bias on the measurement of scale effects has been shown to be severe. Although some form of case-mix adjustment is the minimum quality requirement for a study design to be included in the evaluation, a study might still be rejected in the final assessment if necessary. In the case of data envelopment analyses it is self-evident that the so-called Constant Returns to Scale (CRS) models which, by definition, cannot measure returns to scale and optimal hospital size are excluded.

It should also be emphasised that theoretical papers discussing the merits and limitations of the methodologies employed can be of use while developing the quality criteria. This is also the case for advances made in applications at the non-hospital sector. And some empirical studies that analyse other factors (e.g. case-mix, technical efficiency) which are associated with the reliability of the measurement of scale and scope effects are also valuable. So, despite our narrow inclusion criteria with respect to the empirical studies to be assessed, we aimed at recovering as much research as possible that may contribute towards the configuration of the evaluation process it self.

4 Methodological quality assessment

Each econometric cost or production study satisfying the inclusion criteria will be assessed according to the following criteria:

a) unit of measurement of hospital output,
b) choice of functional form,
c) adjustment for heterogeneity of output (case-mix),
d) derivation of long-run scale estimates,
e) inclusion of input prices (inputs in production studies),
f) treatment of uncertainty,
g) adjustment for quality of care,
h) choice of model variables,
i) regular behaviour of estimated cost functions (in flexible models),
j) regulatory environment and cost minimisation (in flexible cost models).

Although these form our basic assessment criteria other parameters (e.g. sample means, endogeneity of variables, multicollinearity, diagnosis of models using goodness-of-fit measures or specification tests) will also be considered.

Each survival-type study will be assessed by the following criteria:

Control for various factors - other than size or volume - that may determine survival or growth (e.g. case-mix, quality).

Goodness-of-fit will also be discussed.

Each DEA model according to the criteria:

a) unit of measurement of hospital output,
b) adjustment for heterogeneity of output (case-mix),
c) adjustment for input differences,
d) adjustment for quality of care (including reservation quality),
e) treatment of errors in the model specification and the data (e.g. sensitivity analysis to different specifications of outputs, identification and elimination of outliers).

Each statistical study according to the criteria:

a) adjustment for various factors (e.g. case-mix) that might confound the relation of interest,

b) theoretical validity of findings (i.e. is the study measuring what it claims to be measuring, namely economies as defined by economic theory?).

5 Data extraction and synthesis

Data extraction:

For each relevant study data will be extracted in a systematic way in order to highlight the important aspects of the methodologies used, the evidence on scale and scope effects as well as any other findings that may assist the evaluation procedure. The data extraction sheets are shown in Appendix III.

Data synthesis:

The quality of individual studies will be assessed using the criteria mentioned. Sensitivity analysis will be pursued in order to see whether the better-quality studies satisfying the developed criteria yield similar results. If results differ across the more general methodologies, the more reliable ones will be identified - if possible - in light of their relative strengths and weaknesses. Finally, we will check whether any observed discrepancies in different studies with respect to the operation of economies or diseconomies of scale, for the average hospital, can in fact be reconciled once the different sample mean levels of output are considered.

6 Conclusions

On the basis of the preceding analysis conclusions will be drawn about the relationship between the scale and scope of hospital production and the incurred costs. The strength of the evidence will also be critically assessed taking into account the current limitations of the literature.

APPENDIX II - SEARCH STRATEGY

Search Strategy

The following databases were searched:

- Social Science Citation Index (BIDS ISI) for 1983-1996
- Science Citation Index (BIDS ISI) for 1983-1996.
- ISTC (BIDS-ISI) for 1983-1996
- ABI Inform (searched on the OHIO-Link) for 1980-1996 was searched by a third-party in the USA
- Economic Literature Index (Knight Ridder Dialog) for 1969-1994
- Medline (Ovid CD-ROM) for 1980-1996
- Economics Working Paper Archive
- Dissertation Abstracts (Knight Ridder Dialog) for 1980-1996
- Health Planning and Administration (Knight Ridder Dialog) for 1980-1996
- NTIS (Knight Ridder Dialog) for 1980-1996
- Embase (Knight Ridder Datatstar) for 1980-1996

The searches were carried out by a variety of people, including the researcher, CRD information staff and third parties. Detailed descriptions of the strategies are given below:

1. Social Science Citation Index (BIDS ISI) for 1981-1996 and Science Citation Index (BIDS ISI) for 1981-1996

The search syntax below reflects that used by the BIDS system.

Set	Search terms:
1	Economies + Scale
2	Economies + Scope
3	Efficient Scale
4	(Optimum,optimal)+(Scale,Size)
5	Scale + Efficiency
6	Returns + Scale

7	Cost + Complementarities
8	Economies + Specializ(s)ation
9	Economies + Expansion
10	Expansion + Scale
11	Scale + Operation*
12	Mergers+ Trust*
13	Consolidation,consolidations
14	Cost + (Relations,relationships)
15	Cost + (Function,functions)
16	Cost + Analysis
17	Cost + Model
18	Cost + Regression
19	Cost + Frontier
20	Econometric + Estimation + Cost
21	Estimated + Cost
22	Production + Efficiency
23	Production + Returns
24	Production + (Function,functions)
25	Production + Analysis
26	Production + Model
27	Production + Regression
28	Production + Correspondance*
29	Survival,survivor
30	Data + Envelopment
31	Most + Productive + Scale + Size
32	Non + Parametric
33	Nonparametric
34	Cost + Frontier
35	Production + Frontier
36	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20
37	21,22,23,24,25,26,27,28,29,30,31,32,33,34,35
38	36,37
39	Hospital,HOSPITALS
40	Secondary + Care
41	Health + Care + (Institutions,Organizations)
42	Inpatient,OUTPATIENT
43	39,40,41,42
44	38+43

A second level search for Program - Specific Economies of Scale was carried out as follows:

Set search terms

1 Economies + Scale
2 Economies + Scope
3 Returns + Scale
4 1,2,3
5 Patients,Treatment(s),Cardiology,Pediatrics,Obstetrics,Gynaecology
6 Emergency,Trauma,Cardiology,Renal,(General+Surgery),Intensive+Care, Cancer,Radiotherapy,Chemotherapy
7 Geriatrics,Psychiatric
8 5,6,7
9 4+8

2. Index of Scientific and Technical Proceedings (BIDS-ISI) for 1983-1996

The search syntax below reflects that used by the BIDS system.

Set Search terms

1 Economies + (Scale,scope)
2 (Efficient,Optimum,optimal)+(Scale,Size)
3 Scale + (Efficiency,returns)
4 Cost +(Complementarities,Relations,relationships)
5 Economies + (Specialization,specialisation,expansion)
6 Expansion + Scale
7 Scale + Operation*
8 Mergers+ Trust*
9 Consolidation*
10 Cost + (Function,functions,analysis,model,regression,frontier)
11 Econometric + Estimation + Cost
12 Estimated + Cost
13 Production+(Efficiency,returns,Function,functions,analysis,Model,regression, correspondance*)
14 Survival,survivor
15 Data + Envelopment
16 Productive + Scale + Size
17 (Non + Parametric),nonparametric
18 (Cost,production) + Frontier
19 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18
20 Hospital,HOSPITALS,(Secondary + Care)
21 Health + Care + (Institutions,Organizations,organisations)
22` Inpatient*,OUTPATIENT*
23 20,21,22
24 19+23

Second level search for Program - Specific Economies of Scale:

1 Economies + Scale
2 Economies + Scope
3 Returns + Scale
4 1,2,3

3. Economic Literature Index

The Economic Literature Index were initially searched by the researcher in 1995 for search periods covering the years 1969-1994. Search terms:

1 Economies + Scale
2 Economies + Scope
3 Efficient + Scale
4. Optimum (-al) + Scale (Size)
5 Scale + Efficiency, Returns + Scale, Mergers
6 Multihospital + Systems
7 Multihospital + Arrangements
8 Chains + Performance
9 Consolidation(s)
10 Cost + Function(s)
11 Cost + Analysis
12 Cost + Regression
13 Econometric + Estimation + Cost
14 Estimated + Cost
15 Production + Efficiency
16 Production + Returns
17 Production + Function(s)
18 Production + Analysis
19 Production + Regression
20 Survival (vor)
21 Data + Envelopment
22 Most + Productive + Scale + Size,
23 mpss
24 Non + Parametric
25 Nonparametric
26 Cost + Frontier
27 Production + Frontier.
28 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20
29 21,22,23,24,25,26,27,28
30 28,29
31 Hospital(s)
32 Secondary + Care
33 Tertiary + Care
34 30+(31,32,33)

4. Medline

This was searched using the OVID CD-ROM version and the search terms are as follows (OVID syntax is reproduced)

Set	Search terms
001	(economies adj2 (scale or scope)).ti,ab,sh.
002	(efficient adj scale).ti,ab,sh.
003	(production adj (function$ or correspondance$)).ti,ab,sh.
004	(scale and efficiency).ti,ab,sh.
005	((return or returns) adj2 scale).ti,ab,sh.
006	(cost adj2 complementarities).ti,ab,sh.
007	(cost adj2 function?).ti,ab,sh.
008	((cost or costs or production) adj frontier$).ti,ab,sh.
009	or/1-8
010	((optimum or optimal) adj (size or scale)).ti,ab,sh.
011	(economies and speciali#ation).ti,ab,sh.
012	(economies and expansion).ti,ab,sh.
013	(expansion and scale).ti,ab,sh.
014	(scale adj2 (operation or operations)).ti,ab,sh.
015	(econometric adj2 estimat$).ti,ab,sh.
016	(estimated adj (cost or costs)).ti,ab,sh.
017	(size adj2 expenditures).ti,ab,sh.
018	(production and efficiency).ti,ab,sh.
019	(production and returns).ti,ab,sh.
020	(production and regression).ti,ab,sh.
021	(data adj envelopment).ti,ab,sh.
022	mergers.tw.
023	HEALTH FACILITY MERGER/ec
024	models,econometric
025	(survival adj analysis).ti,ab,sh.
026	(productive adj scale adj size).ti,ab,sh.
027	mes.tw.
028	mpss.tw. not exp "methylprednisolone"/
029	(non adj parametric).ti,ab,sh.
030	nonparametric.tw.
031	translog.tw.
032	((cost or costs) adj2 analy$).ti,ab,sh.
033	or/10,18-20,29-30
034	32 and 33
035	or/11-17,21-28,31
036	exp primary health care/
037	physicians, family/
038	exp leper colonies/
039	exp medical office buildings/
040	exp nurseries/
041	exp physicians' offices/
042	exp residential facilities/

043 FAMILY PRACTICE/
044 home care services/
045 exp ambulatory care facilities/
046 or/36-45
047 exp "hospitalization"/
048 exp health facilities/
049 exp economics, hospital/
050 (hospital or hospitals or hospitali#ation).tw.
051 (patient or patients).tw.
052 or/47-51
053 or/34-35
054 46 not (52 and 46)
055 (53 and 52) not 54

5. ABI Inform

This search was carried out by a third party.

Search terms:

1 Economies + Scale
2 Economies + Scope
3 Returns + Scale
4 Data + Envelopment
5 Survival (vor)
6 Mergers
7 Multi-hospital + Arrangements,
8 1,2,3,4,5,6,7
9 Hospital(s).
10 8+9

6. Economics Working Paper Archive

This did not reveal any references. The search terms used were:

hospital
secondary + care
tertiary + care
patients
treatments

7. Dissertation Abstracts, Health Planning and Administration, Embase, Dissertation Abstracts and NTIS were searched simultaneously on Knght Ridder Dialog Onesearch with Medline, to eliminate duplicate records.

Set	Search terms
1	ECONOMIES(2W)(SCALE OR SCOPE)
2	EFFICIENT()SCALE
3	PRODUCTION()(FUNCTION? OR CORRESPONDANCE?)
4	SCALE AND EFFICIENCY AND (COST()ANALYSIS)
5	RETURNS(2W)SCALE
6	COST(2W)COMPLEMENTARITIES
7	COST(2W)FUNCTION?
8	(COST OR PRODUCTION)()FRONTIER?
9	1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8
10	(OPTIMUM OR OPTIMAL)()(SIZE OR SCALE)
11	ECONOMIES AND SPECIALI?ATION
12	ECONOMIES AND EXPANSION
13	ECONOMETRIC(2W)ESTIMAT?
14	ESTIMATED()(COST OR COSTS)
15	SIZE(W2)EXPENDITURES
16	PRODUCTION AND (EFFICIENCY OR RETURNS OR REGRESSION)
17	PRODUCTION(L)(EFFICIENCY OR RETURNS OR REGRESSION)
18	PRODUCTION(5W)(EFFICIENCY OR RETURNS OR REGRESSION)
19	DATA()ENVELOPMENT
20	HEALTH FACILITY MERGER(L)EC/DE
21	PRODUCTIVE()SCALE()SIZE
22	MES
23	MPSS
24	(S22 OR S23) NOT METHYLPREDNISOLONE
25	NON()PARAMETRIC OR NONPARAMETRIC
26	TRANSLOG
27	10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 17 OR 18 OR 19 OR 20 OR 21 OR 24 OR 25 OR 26
28	PRIMARY HEALTH CARE!
29	DC=N1.20.665.380.710?
30	PHYSICIANS()FAMILY/DE
31	DC=Q1.550.45? OR GENERAL()PRACTITIONER/DE
32	MEDICAL OFFICE BUILDINGS!
33	NURSERIES!
34	PHYSICIANS OFFICES!
35	RESIDENTIAL FACILITIES!
36	FAMILY()PRACTICE/DE
37	HOME()CARE()SERVICES/DE
38	AMBULATORY()CARE()SERVICES!
39	AMBULATORY CARE SERVICES!
40	PRIMARY()CARE OR GENERAL()PRACTI? OR FAMILY()PRACTICE
41	s28-s40
42	HOSPITALIZATION! OR HEALTH FACILITIES!

43 ECONOMICS HOSPITAL! OR DC=N4.380.400?
44 ECONOMICS()HOSPITAL!
45 HOSPITAL OR HOSPITALS OR HOSPITALI?ATION
46 PATIENT OR PATIENTS OR INPATIENTS OR INPATIENT
47 DC=N1.10.400? OR DC=I1.655.530.400? OR DC=N1.20.655.380.405?
48 S42-S47
49 S41 NOT (S41 AND S48)
50 (S48 AND S27) NOT S49
51 S9 OR S50
52 S51/1980-1996
53 REMOVE DUPLICATES ON SET 52

A subsequent search was carried out on all the databases for the following additional terms: Set 7

Set search terms
1 MULTI()HOSPITAL?
2 MULTI()INSTITUTIONAL()SYSTEMS?/DE
3 MULTIHOSPITAL()SYSTEM/DE
4 MULTIHOSPITAL?
5 S1 OR S2 OR S3 OR S4
6 ECONOMIES OR ECONOMIC?
7 TC=0139
8 DC=I1.250?
9 S2(L)EC
10 S6-S8
11 S5 AND S10
12 S11 OR S9
13 remove duplicates of set 12

8. DEA WWW bibliography

This consisted of a list of references and hence no key words had to be chosen.

APPENDIX III - DATA EXTRACTION SHEETS

DATA EXTRACTION SHEET 1: COST AND PRODUCTION FUNCTION STUDIES

1. Bibliographic information

a) Author:

b) Title:

c) Year of publication:

d) Journal:

2. Type of analysis(-es):

Cost analysis:

Production analysis:

3. Unit of analysis:

General hospital (e.g. short-term acute hospital):

Speciality:

Other (e.g. patient, ward):

4 Characteristics of the model(s):

a) Structural (e.g. only outputs and input prices in the cost function):

Behavioural: (i.e. augmented by other independent variables):

b) Flexible functional form (e.g. translog):

Restrictive functional form (e.g. Cobb-Douglas):

c) Short-run function:

Long-run function:

5. Objective(s) and details on the methodology of the study:

6. Description of the sample:

Discussion of the quality of the data:

Origin of data (country, location):

Number of observations (and number of parameters to be estimated):

Means of selected variables (mainly volumes, size):

Type of included hospitals (e.g. teaching, for-profit):

Cross-sectional or time-series (and years):

7. Estimation method:

8. Model diagnosis:

Theoretical consistency of estimation results (e.g. regularity conditions):

Goodness-of-fit measures and mispecification tests:

Presence and treatment of multicollinearity:

Presence and treatment of heteroscedasticity:

Presence and treatment of serial-correlation in time-series:

Presence and treatment of endogeneity of variables (e.g. inputs, outputs):

9. Unit of measurement of hospital output (e.g. days, cases):

10. Adjustment for:

Output heterogeneity (case-mix, teaching output etc.):

Quality of care:

Uncertainty (standby capacity):

11. Inclusion of input prices or justification for exclusion (cost functions):

Inclusion of inputs (production functions):

12. Other included control variables (e.g. market competition variables):

13. Theoretical justification or empirical treatment of allocative efficiency hypothesis (e.g. cost-based or prospective reimbursement regime):

14. Empirical findings:

a) On economies of (returns to) scale

Computed measure of economies, its value and interpretation (e.g. short-run, long-run and global, local):

Theoretical validity of the measure:

Pattern of economies or returns to scale:

Optimal size or volume level of operation:

Statistical significance of the measure (i.e. is the null of constant returns rejected?):

b) On economies of scope

Method of calculation (e.g. weak cost complementarities or difference):

Nature of the measure (e.g. global, local):

Pairs for which weak cost complementarities exist:

Pairs for which weak cost complementarities are absent:

Pairs of services exhibiting economies (e.g. paediatrics, all other outputs):

Pairs of services exhibiting diseconomies:

Pairs of services exhibiting neither economies nor diseconomies:

Statistical significance of the measure (i.e. is the null of no scope effects rejected?):

15. Findings useful for the validity assessment process

Assumptions and restricted models tested (e.g. separability assumption):

Identification of potential (additional) validity assessment criteria:

Other:

DATA EXTRACTION SHEET 2: DATA ENVELOPMENT ANALYSES

1. Bibliographic information

a) Author:

b) Title:

c) Year of publication:

d) Journal:

2. Unit of analysis:

General hospital (e.g. short-term acute hospital):

Speciality:

Other (e.g. aggregates of hospitals in a region, ward):

3. Objective(s) and details of the methodology of the study:

4. Description of the sample:

Discussion of the quality of the data:

Origin of data (country, location):

Number of observations:

+

Means of selected variables (mainly bed size):

Type of included hospitals (e.g. teaching, for-profit):

5. Model diagnosis (e.g. misspecification, measurement error):

6. Unit of measurement of hospital output (e.g. cases or patient days):

7. Adjustment for:

Output heterogeneity (e.g. case-mix, teaching):
Quality of care differences:

Uncertainty (reservation quality):

Other differences (e.g. for-profit, not-for-profit status):

Input heterogeneity:

8. Findings on returns to scale:

Pattern of returns to scale:

Computed measure and its interpretation (e.g. most productive scale size, largest-size efficient scale):

Optimal facility size:

9. Findings useful for the validity assessment process:

DATA EXTRACTION SHEET 3: SURVIVAL ANALYSES

1. Bibliographic information

a) Author:

b) Title:

c) Year of publication:

d) Journal:

2. Model type:

Univariate (i.e. descriptive of changes in market shares of different bed size categories over time):

Multivariate (i.e. many regressors thought to influence survival or growth):

3. Objectives and methodology (e.g. continuous growth model estimated by OLS, survival binary model i.e. logistic regression, functional form):

4. Description of sample (observations, country, year):

5. Included variables:

Dependent variable:

Independent variables (e.g. size, controls for case-mix, quality of care):

6. Model diagnosis (e.g. goodness of fit):

7. Findings on economies of scale (range of optimal hospital sizes):

8. Findings useful for the validity assessment process:

DATA EXTRACTION SHEET 4: STATISTICAL STUDIES

1. Bibliographic information

a) Author:

b) Year of publication:

c) Title:

d) Journal:

2. Unit of analysis:

3. Objectives and methodology:

4. Description of sample:

5. Adjustment for various confounding factors (e.g. case-mix):

6. Theoretical validity of findings on economies (is the technique measuring economies appropriately i.e. as defined by economic theory?)

7. Findings useful for the validity assessment process:

REFERENCES

Ablett, J.R. (1993) Economies d' Echelle et de Gamme en Milieu Hospitalier Belge: Une Etude Transversale, *Cahiers Economiques de Bruxelles*, no.139, 3eme trimestre, pp.369-399.

Ahern, M. M. (1988) *The Effect of Medicare DRG Reimbursement on the Structure of Hospital Costs*, Unpublished PhD Dissertation, The Florida State University.

Alchian, A. (1959) Costs and Outputs, in M. Abramovitz (ed.), *The Allocation of Economic Resources: Essays in Honor of B.F. Haley*, pp.23-40. Standford: Standford University Press.

Aletras, V.H. (1995) *A Critical Review of the Empirical Evidence on Economies of Scale and Scope in Health Care with Emphasis on Hospital-based Studies*, Unpublished Msc Dissertation, University of York.

Anderson, D.L. (1974) Economic Aspects of Clinical Laboratory Service, pp. 118-139 in R.D. Fraser (ed.) *Health Economics Symposium: Proceedings of the First Canadian Conference,* Sept. 4-6.

Arndt, M., Bradbury, R.C., Golec, J.H. (1995) Surgeon Volume and Hospital Resource Utilization, *Inquiry*, vol.32, Winter,pp. 407-416.

Atkinson, S.E., Halvorsen, R. (1992) Estimating a Non-minimum Cost Function for Hospitals: Comment, *Southern Economic Journal*, vol.58, no.4, pp.1114-1117.

Ault, D., Johnson, T.Jr. (1979) A Study of the Determinants of Variations in the Cost Structures of USA Hospitals, *Journal of Behavioural Economics*, vol.8, pp.79-113.

Banker, R.D. (1984) Measuring Most Productive Scale Size Using Data Envelopment Analysis, *European Journal of Operational Research*, vol.17, pp.35-44.

Banker, R.D., Das, S., Datar, S.M. (1989) Analysis of Cost Variances for Management Control in Hospitals, *Research in Governmental and Nonprofit Accounting*, vol.5, pp.269-291.

Banker, R. D., Conrad, R.F., Strauss, R.P. (1986) A Comparative Application of Data Envelopment Analysis and Translog Methods: An Illustrative Study of Hospital Production, *Management Science*, vol.32, no.1, pp.30-44.

Banker, R.D. (1989) Econometric Estimation and Data Envelopment Analysis, *Research in Governmental and Nonprofit Accounting*, vol.5, pp.231-243.

Banker, R.D., Gadh, V., Goor, W. (1989) *A Monte Carlo Comparison of Efficiency Estimation Methods*, Working Paper, Carnegie Mellon University.

Banks, D.A. (1993) Voluntary and Proprietary Hospital Behavioural Response to Socio-Economic Stimuli, *Applied Economics*, vol.25, pp.853-868.

Barer, M.L. (1982) Case Mix Adjustment in Hospital Cost Analysis: Information Theory Revisited, *Journal of Health Economics*, vol.1, pp.53-80.

Baumol, W.J. (1977) On the Proper Tests for Natural Monopoly in a Multiproduct Industry, *American Economic Review*, vol.67, pp.809-822.

Baumol, W.J., Panzar, J.C., Willig, R.D. (1982) *Contestable Markets and the Theory of Industrial Structure*, Harcourt Brace Jovanovich, New York.

Bays, C. W. (1980) Specification Error in the Estimation of Hospital Cost Functions, *Review of Economics and Statistics*, vol.62, no.2, pp.302-305.

Bays, C. W. (1986) The Determinants of Hospital Size: A Survivor Analysis, *Applied Economics*, vol.18, pp.359-377

Becker, E.R., Sloan, F.A. (1985) Hospital Ownership and Performance, *Inquiry (USA)*, vol.23, no.1, pp.21-36.

Berki, S.E. (1972) *Hospital Economics*, Lexington Books.

Berry, R.E. Jr. (1967) Returns to Scale in the Production of Hospital Services, *Health Services Research*, vol.2, no.2, pp.123-139.

Berry, R.E. Jr. (1970) Product Heterogeneity and Hospital Cost Analysis, *Inquiry (USA)*, vol.7, no.1 (March), pp.67-75.

Berry, R.E. Jr. (1974) Cost and Efficiency in the Production of Hospital Services, *Midland Memorial Fund Quarterly (MMFQ)*, Health and Society, Summer.

Braeutigam, R.R., Daughety, A.F. (1983) On the Estimation of Returns to Scale Using Variable Cost Functions, *Economics Letters*, vol.11, pp.25-31.

Breyer, F. (1987) "The Specification of a Hospital Cost Function: A Comment on the Recent Literature", *Journal of Health Economics*, vol.6, pp.147-157.

Brown, R.S., Caves, D.W., Christensen, L.R. (1979) Modelling the Structure and Production for Multiproduct Firms, *Southern Economic Journal*, vol.49, pp.256-273.

Brown, M.C. (1980) Production and Cost Relations of Newfoundland's Cottage Hospitals, *Inquiry (USA)*, vol.17, pp.268-277.

Bruning, E.R., Register, C.A. (1989) Technical Efficiency within Hospitals: Do Profit Incentives Matter?, *Applied Economics*, vol.21, pp.1217-1233.

Burns, A.M. (1982) *The Economics of the Connecticut Hospital Industry: Competition, Structure, and Performance*, Unpublished PhD Dissertation, The University of Connecticut.

Butler, J.R.G. (1995) *Hospital Cost Analysis*, Kluwer Academic Publishers, Dordrecht / Boston / London.

Byrnes, P., Valdmanis, V. (1994) Analyzing Technical and Allocative Efficiency of Hospitals, in A. Charnes, W.W. Cooper, A.Y. Lewin and L.M. Seiford (ed.) *Data Envelopment Analysis: Theory, Methodology and Applications*, Kluwer.

Carr, J.W., Feldstein, P.J. (1967) The relationship of cost to hospital size, *Inquiry (USA)*, vol.4, pp.45-65.

Caves, D.W., Christensen, L.R. (1976) *Modelling the Structure of Production in the US Railroad Industry*, paper presented at the winter meetings of the Econometric Society, Sept.

Chambers, R.G. (1988) *Applied Production Analysis: A Dual Approach*, Cambridge University Press, Cambridge / New York / New Rochelle / Melbourne / Sydney.

Charnes A., Cooper W.W., Lewin A.Y., Seiford, L.M. (1994) *Data Envelopment Analysis: Theory, Methodology and Applications*, Kluwer Academic Publishers.

Chernichovsky, D., Zmora, I. (1986) A Hedonic Prices Approach to Hospitalization Costs, *Journal of Health Economics*, vol.5, pp.179-191.

Christensen, L.R., Jorgenson, D.W., Lau, L.J. (1973) Transcendental Logarithmic Production Frontiers, *Review of Economics and Statistics*, vol.55, pp.28-45.

Cohen, H.A. (1967) Variations in Cost Among Hospitals of Different Sizes, *Southern Economic Journal*, vol.33, pp.355-366.

Cohen, H.A. (1970) Hospital Cost Curves with Emphasis on Measuring Patient Care Output, in H.E. Klarman (ed.) *Empirical Studies in Health Economics: Proceedings of the Second Conference on the Economics of Health*, John Hopkins Press, Baltimore and London.

Collins, S.R. (1994) *Estimation of a Multiproduct Translog Cost Function for Maryland Hospitals*, Unpublished PhD Dissertation, The George Washington University.

Conrad, R.F., Strauss, R.P. (1983) A Multiple-Output Multiple-Input Model of the Hospital Industry in North Carolina, *Applied Economics*, vol.15, pp.341-352.

Cowing, T.G., Holtmann, A.G., Powers, S. (1983) Hospital Cost Analysis: a Survey and Evaluation of Recent Studies, *Advances in Health Economics and Health Economics Research*, vol.4, pp.257-303.

Cowing, T.G., Holtmann, A.G. (1983) Multiproduct Short-Run Hospital Cost Functions: Empirical Evidence and Policy Implications From Cross-section Data, *Southern Economic Journal*, vol.49 (Jan.), pp.637-653.

Coyne, J.S. (1982) Hospital Performance in Multihospital Systems: A Comparative Study of System and Independent Hospitals, *Health Services Research*, vol.17, no.4 (Winter), pp.303-329.

Culler, S.D., Holmes, A.M., Gutierrez, B. (1995) Expected Hospital Costs of Knee Replacement for Rural Residents by Location of Service, *Medical Care*, vol.33, no.12, pp.1188-1209.

Culyer, A.J., Wiseman, J., Drummond, M.F., West, P.A. (1978) What Accounts for the Higher Costs of Teaching Hospitals?, *Social and Economic Administration*, vol.12, no.1 (Spring), pp.20-29.

Custer, W.S., Willke R.J. (1991) Teaching Hospital Cost: The Effects of Medical Staff Characteristics, *Health Services Research*, vol.25, no.6, pp.831-857.

Derveaux, B., Leleu, H., Lebrun, T., Boussemart, J.P. (1994) Construction d' un Indice de Productivite pour le Secteur Hospitalier Public, Version Provisoire, *Xvemes Journees des Economistes de la Sante*, 20 et 21 Janvier.

Donaldson, C., Gerard, K. (1992) *Economics of Health Care Financing: The Visible Hand,* Macmillan Press Ltd., London.

Dranove, D., Shanley, M. (1995) Cost Reductions or Reputation Enhancement as Motives for Mergers: The Logic of Multihospital Systems, *Strategic Management Journal*, vol.16, pp.55-74.

Eakin, K.B., Kniesner, T.J. (1988) Estimating a Non-minimum Cost Function for Hospitals, *Southern Economic Journal*, vol.54, pp.583-597.

Eakin, K. B. (1991) Allocative Inefficiency in the Production of Hospital Services, *Southern Economic Journal*, vol.58, no.1, pp.240-248.

Eakin, K.B., Kniesner, T.J. (1992) Estimating a Non-minimum Cost Function for Hospitals: Reply, *Southern Economic Journal*, vol.58, no.4, pp.1118-121

Ellis, R.P. (1993) *Hospital Cost Function when Firms may Not Try to Minimise Total Costs*, Working Paper, Sept., Boston University

Ermann, D., Gabel, J. (1986) Multihospital Systems: Issues and Empirical Findings, *Health Affairs*, vol.3, no.1 (Spring), pp.50-64.

Evans, R.G. (1971) "Behavioural" Cost Functions for Hospitals, *Canadian Journal of Economics*, vol.4, no.2 (May), pp.198-215.

Evans, R.G., Walker, H.D. (1972) Information Theory and the Analysis of Hospital Cost Structure, *Canadian Journal of Economics*, vol.5, no.3, pp.398-418.

Fare, R., Grosskopf, S., Lovell, C.A.K. (1985) The Measurement of Efficiency of Production, Kluwer - Nijhoff Publishing, Boston / Dordrecht / Lancaster.

Feldstein, M.S. (1967) *Economic Analysis for Health Services Efficiency: Econometric Studies of the British National Health Service*, North-Holland Publ. Co., Amsterdam.

Feldstein, M.S., Schuttinga, J. (1977) Hospital Costs in Massachusetts: A Methodological Study, *Inquiry (USA),* vol.14 (March), pp.22-31.

Feldstein, P.J. (1983) *Health Care Economics* (2nd ed.), John Wiley & Sons.

Finch, L.E., Christianson, J.B. (1981) Rural Hospital Costs: An Analysis with Policy Implications, *Public Health Reports*, vol.96, no.5 (Sept.-Oct.), pp.423-433.

Finkler, S.A. (1979) Cost - Effectiveness of Regionalization: The Heart Surgery Example, *Inquiry*, vol.16 (Fall), pp.264-270.

Finkler, S.A. (1981) Cost - Effectiveness of Regionalization: Further Results for Heart Surgery, *Health Services Research*, vol.16, no.3 (Fall), pp.325-333.

Fleming, S.T. (1991) The Relationship Between Quality and Cost: Pure and Simple?, *Inquiry (USA)*, vol.28 (Spring), pp.29-38.

Folland, S., Goodman, A.C., Stano, M. (1993) *The Economics of Health and Health Care*, Macmillan Publishing Company, New York / Toronto / Oxford / Singapore / Sydney.

Fordham, R., Field, D.J., Hodges, S., Normand, C., Mason, E., Burton, P., Yates, J., Male, S. (1992) Cost of Neonatal Care Across a Regional Health Authority, *Journal of Public Health Medicine*, vol.14, no.2, pp.127-130.

Fottler, M.D., Rock, W.K. (1974) Some Correlates of Hospital Costs in Public and Private Hospital Systems: New York City, *Quarterly Review of Economics and Business*, vol.14, no.1, pp.39-53.

Fournier, G.M., Mitchell, J.M. (1992) Hospital Costs and Competition for Services: A Multiproduct Analysis, *Review of Economics and Statistics*, vol.74, no.4, pp.627-634.

Francisco, E.W. (1970) Analysis of Cost Variations Among Short-term General Hospitals, in H.E. Klarman (ed.) *Empirical Studies in Health Economics: Proceedings of the Second Conference on the Economics of Health*, John Hopkins Press, Baltimore and London.

Frank, R.G. (1988) On Making the Illustration Illustrative: A Comment on Banker, Conrad and Strauss, *Management Science*, vol.34, no.8 (Aug.), pp.1026-1029.

Frech, H. E. III, Mobley, L.R. (1995) Resolving the Impasse on Hospital Scale Economies: A New Approach, *Applied Economics*, vol.27, pp.286-296.

Friedman, B., Pauly, M.V. (1981) Cost Functions for a Service Firm with Variable Quality and Stochastic Demand: The Case of Hospitals, *The Review of Economics and Statistics*, vol.63, pp.620-624.

Friedman, B., Pauly, M.V. (1983) A New Approach to Hospital Cost Functions and Some Issues in Revenue Regulation, *Health Care Financing Review*, vol.4, no.3, March, pp.105-114.

Ganley, J.A., Cubbin, J.S. (1992) *Public Sector Efficiency Measurement: Applications of Data Envelopment Analysis*, North-Holland.

Gaynor, M., Anderson, G.F. (1995) Uncertain Demand, the Structure of Hospital Costs, and the Cost of Empty Hospital Beds, *Journal of Health Economics*, vol.14, pp.291-317.

Goodman,C. (1993), *Literature Searching and Evidence Interpretation for Assessing Health Care Practices*, The Swedish Council on Technology Assessment in Health Care.

Grannemann, W., Brown, R.S., Pauly, M.V. (1986) Estimating Hospital Costs: A Multiple-output Analysis, *Journal of Health Economics*, vol.5, pp.107-127.

Gravelle, H., Rees, R. (1992) *Microeconomics* (2nd ed.), Longman, London / New York.

Grosskopf, S., Valdmanis, V. (1987) Measuring Hospital Performance: A Non-parametric Approach, *Journal of Health Economics*, vol.6, pp.89-107.

Grosskopf, S., Valdmanis, V. (1993) Evaluating Hospital Performance with Case-Mix-Adjusted Outputs, *Medical Care*, vol.31, no.6, pp.525-532.

Gruca, T.S., Nath, D. (1994) Regulatory Change, Constraints on Adaptation and Organisational Failure: An Empirical Analysis of Acute Care Hospitals, *Strategic Management Journal*, vol.15, pp.345-363.

Harris, J.E. (1977) The Internal Organization of Hospitals: Some Economic Implications, *Bell Journal of Economics*, vol.8, Autumn, pp.467-482.

Hefty, T.R. (1969) Returns to Scale in Hospitals: A Critical Review of Recent Research, *Health Services Research*, vol.4, pp.267-280.

Hellinger, F.J. (1975) Specification of a Hospital Production Function, *Applied Economics*, vol.7, pp.149-160.

Hornbrook, M.C., Monheit, A.C. (1985) The Contribution of Case-mix Severity to the Hospital Cost-output Relation, *Inquiry (USA)*, vol.22 (Fall), pp.259-271.

Hu, T. (1971) Hospital Costs and Pricing Behaviour: The Maternity Ward, *Inquiry (USA)*, vol.8, no.4, pp.19-26.

Ingbar, M.L., Taylor, L.D. (1968) *Hospital Costs in Massachusetts: an Econometric Study*, Harvard University Press, Cambridge / Massachusetts.

Jacobs, P. (1974) A Survey of Economic Models of Hospitals, *Inquiry (USA)*, vol.11 (June), pp.83-97.

Jenkins, A.W. (1980) Multiproduct Cost Analysis: Service and Case-type Cost Equations for Ontario Hospitals, *Applied Economics*, vol.12, pp.103-113.

Jensen, G.A., Morrisey, M.A. (1986) The Role of Physician in Hospital Production, *Review of Economics and Statistics*, vol.63, no.3, pp.432-442.

Johnston, J. (1960) *Statistical Cost Analysis*, McGraw-Hill Book Co., New York / Toronto / London.

Kemere, P. (1992) *The Structure of Hospital Costs: An Econometric Analysis of Short-term General Hospitals in Maryland*, Unpublished PhD Dissertation, Howard University, Washington, D.C.

Lave, J.R., Lave, L.B. (1970a) Estimated Cost Functions for Pennsylvania Hospitals, *Inquiry (USA)*, vol.7, no.2, pp.3-14.

Lave, J.R., Lave, L.B. (1970b) Hospital Cost Functions, *American Economic Review*, vol.60, pp.379-395.

Lave, J.R., Lave, L.B. (1970c) Economic Analysis for Health Service Efficiency: A Review Article, *Applied Economics*, vol.1, pp.293-305.

Lave, J.R., Lave, L.B., Silverman, L.P. (1972) Hospital Cost Estimation Controlling for Case-mix, *Applied Economics*, vol.4, pp.165-180.

Lavers, R.J., Whynes, D.K. (1978) A Production Function of English Maternity Hospitals, *Socio-economic Planning Sciences*, vol.12, no.2, pp.85-93.

Lee, M.L. (1971) A Conspicuous Production Theory of Hospital Behaviour", *Southern Economic Journal*, vol.38, pp.48-58.

Lee, M.L., Wallace, R.L. (1973) Problems in Estimating Multiproduct Cost Functions: An Application to Hospitals, *Western Economic Journal*, vol.11 (Sept.), pp.350-363.

Lee, D.W. (1987) *Efficient Bundling and Unbundling of Hospital Services: An Examination of Scale and Scope Economies in the Hospital Industry*, Unpublished PhD Dissertation, University of Illinois at Urbana - Champaign.

Levitz, G.S., Brooke, P.P. Jr. (1985) Independent versus System-Affiliated Hospitals: A Comparative Analysis of Financial Performance, Cost and Productivity, *Health Services Research*, vol.20, no.3, pp.315-339.

Lewin, L.S., Derzon, R.A., Margulies, R. (1981) Investor-owneds and Nonprofits Differ in Economic Performance, *Hospitals* (July), pp.52-58.

Lillie - Blanton, M., Felt, S., Redmon, P., Renn, S., Machlin, S., Wennar, E. (1992) Rural and Urban Hospital Closures, 1985-1988: Operating and Environmental Characteristics that Affect Risk, *Inquiry*, vol.29 (Fall), pp.332-344.

Long, M.J., Ament, R.P., Dreachslin, J.L., Kobrinski, E.J. (1985) A Reconsideration of Economies of Scale in the Health Care Field, *Health Policy*, vol.5, pp.25-44.

Lynk, W.J. (1995) The Creation of Economic Efficiencies in Hospital Mergers, *Journal of Health Economics*, vol.14, pp.507-530.

Lynch, J.R., Ozcan, Y.A. (1994) Hospital Closure: An Efficiency Analysis, *Hospital & Health Services Administration*, vol.39, no.2, pp.205-220.

Maddala, G.S. (1977) *Econometrics*, McGraw-Hill Book Co., New York.

Maindiratta, A. (1990) Largest Size-Efficient Scale and Size Efficiencies of Decision - Making Units in Data Envelopment Analysis, *Journal of Econometrics*, vol.46, pp.57-72.

Manheim, L.M., Shortell, S.M., McFall, S. (1989) The Effect of Investor-Owned Chain Acquisitions on Hospital Expenses and Staffing, *Health Services Research*, vol.24, no.4, pp.461-484.

Mann, J.K., Yett, D.E. (1968) The Analysis of Hospital Costs: A Review Article, *Journal of Business*, vol.41, pp.191-202.

Markham, B. (1995) Review of the Multi-hospital Arrangements Literature: Benefits, Disadvantages, and Lessons for Implementation, *Submitted to Forum.*

McGregor, M., Pelletier, G. (1978) Planning of Specialized Health Facilities: Size vs. Cost and Effectiveness in Heart Surgery, The New England Journal of Medicine, vol.299, p.179.

Mobley, L.R., (1990) *Multihospital Systems in California: Behavior and Efficiency*, Unpublished PhD Dissertation, University of California, Santa Barbara.

Mobley, L.R., Frech H.E. III (1994) Firm Growth and Failure in Increasingly Competitive Markets: Theory and Application to Hospital Markets, *Journal of the Economics of Business*, vol.1, no.1, pp.77-93.

Montfort, G. van (1981) Production Functions for General Hospitals, *Social Science and Medicine*, vol.15C, pp.87-98.

Mulligan, J.G. (1987) Hospital Scale Economies due to Stochastic Demand and Service, *Economic Letters*, vol.23, no.2, pp.193-197.

Mullner, R.M., Rydman, R.J., Whiteis, D.G. (1990) Rural Hospital Survival: An Analysis of Facilities and Services Correlated with Risk of Closure, *Hospital & Health Services Administration*, vol.35, no.1 (Spring), pp.121-137.

Munoz, E., Boiardo, R., Mulloy, K., Goldstein, J., Brewster, J.G., Tenenbaum, N., Wise, L. (1990a) Economies of Scale, Physician Volume for Orthopedic Surgical Patients, and the DRG Prospective Payment System, *Orthopedics*, vol.13, no.1 (Jan.), pp.39-44.

Munoz, E., Boiardo, R., Mulloy, K., Goldstein, J., Tenenbaum, N., Wise, L. (1990b) Economies of Scale, Physician Volume for Neurosurgery Patients, and the Diagnosis - Related Group Prospective Hospital Payment System, *Neurosurgery*, vol. 26, no.1 (Jan.), pp.156-161.

Munoz, E., Boiardo, R., Mulloy, K., Goldstein, J., Brewster, J.G., Wise, L. (1990c) Economies of Scale, Physician Volume for Urology Patients, and DRG Prospective Hospital Payment System, *Urology*, vol. XXXVI, no.5 (Nov.), pp.471-476.

Nelson R.A., (1985) Returns to Scale from Variable and Total Cost Functions, *Economics Letters*, no.18, pp.271-276.

Newhouse, J.P. (1970) Toward a Theory of Nonprofit Institutions: An Economic Model of a Hospital, *American Economic Review*, vol.60, March, pp.64-74.

NHS Centre for Reviews and Dissemination (1995) *Guidelines for Undertaking Systematic Reviews of Effectiveness*, version 2.1, 10th February.

NHS Centre for Reviews and Dissemination (1996) *Undertaking Systematic Reviews of Research on Effectiveness: CRD Guidelines for Those Carrying Out or Commissioning Reviews*, CRD Report Number 4, January.

Norman, M., Stoker, B. (1991) *Data Envelopment Analysis: The Assessment of Performance*, John Wiley & Sons.

Nunamaker, T.R. (1985) Using Data Envelopment Analysis to Measure the Efficiency of Non-profit Organizations: A Critical Evaluation, *Managerial and Decision Economics*, vol.6, no.1, pp.50-58.

Okunade, A.A. (1993) Production Cost Structure of USA Hospital Pharmacies: Time-Series, Cross-sectional Bed Size Evidence, *Journal of Applied Econometrics*, vol. 8, pp.277-294.

Pangilinan, M-B. (1991) *Production / Cost Inefficiency and Flexible Cost Functions: The Case of New York State Hospitals*, 1981-1987, Unpublished PhD Dissertation, State University of New York at Albany.

Panzar, J.C., Willig, R.D. (1977) Economies of Scale in Multi-output Production, *Quarterly Journal of Economics*, vol.91, pp.481-493.

Pattison, R.V., Katz, H.M. (1983) Investor-Owned and Not-For-Profit Hospitals: A Comparison Based on California Data, *The New England Journal of Medicine*, vol.309, no.6, pp.347-353.

Pauly, M. V., Redisch, M. (1973) The Not-for-profit Hospital as a Physicians' Co-operative, *The American Economic Review*, March, pp.87-99.

Pauly, M. V. (1978) Medical Staff Characteristics and Hospital Costs, *Journal of Human Resources*, supplement, pp.77-111.

Robinson, J.C., Luft, H.S. (1985) The Impact of Hospital Market Structure on Patient Volume, Average Length of Stay, and the Cost of Care, *Journal of Health Economics*, vol.4, pp.333-356.

Roddy, D.J. (1980) *Empirical Analysis of the Multiproduct Cost Structure of the Short Term General Hospital Industry*, Unpublished PhD Dissertation, University of Winsconsin - Madison.

Roller, L.-H. (1990a) Modelling Cost Structure: The Bell System Revisited, *Applied Economics*, vol.22, pp.1661-1674.

Roller, L.-H. (1990b) Proper Quadratic Cost Functions with an Application to the Bell System, *The Review of Economics and Statistics*, pp.202-210.

Rozek, R.P. (1988) A Nonparametric Test for Economies of Scope, *Applied Economics*, vol.20, pp.653-663.

Schaafsma, J. (1986) Average Hospital Size and the Total Operating Expenditures for B1 Beds Distributed over H1 Hospitals, *Applied Economics*, vol.18, pp.279-290.

Scott, A., Parkin, D. (1995) Investigating Hospital Efficiency in the New NHS: The Role of the Translog Cost Function, *Health Economics*, vol.4, pp.467-478.

Scuffham, P.A., Devlin, N.J., Jaforullah, M. (1996) The Structure of Costs and Production in New Zealand Public Hospitals: An Application of the Transcendental Logarithmic Variable Cost Function, *Applied Economics*, vol.28, pp.75-85.

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

Sherman, H.D. (1984) Hospital Efficiency Measurement and Evaluation: Empirical Test of a New Technique, *Medical Care*, vol.22, no.10, pp.922-938.

Simpson, J. (1995) A Note on Entry by Small Hospitals, *Journal of Health Economics*, vol.14, pp.107-113.

Sinay, U.A. (1994) *The Theory of Hospital Mergers: Multiproduct Scope and Scale Economies*, Unpublished PhD Dissertation, Saint Louis University.

Sloan, F.A., Becker, E.R. (1981) Internal Organization and Hospital Costs, *Inquiry (USA)*, Fall, pp.224-239.

Sloan, F.A., Feldman, R.D., Steinwald, B.A. (1983) Effects of Teaching on Hospital Costs, *Journal of Health Economics*, vol.2, pp.1-28.

Stigler, G. J., (1958) The Economies of Scale, *The Journal of Law and Econometrics*, vol.1, pp.54-71.

Tatchell, M. (1983) Measuring Hospital Output: A Review of the Service Mix and Case Mix Approaches, *Social Science and Medicine*, vol.17, pp.871-883.

Thanassoulis, E. (1993) A Comparison of Regression Analysis and Data Envelopment Analysis as Alternative Methods for Performance Assessments, *Journal of Operational Research Society*, vol.44, no.11, pp.1129-1144.

Treat, T.F. (1976) The Performance of Merging Hospitals, *Medical Care*, vol.XIV, no.3, pp.199-209.

Valdmanis, V. (1992) Sensitivity Analysis for DEA Models: An Empirical Example using Public vs. NFP Hospitals, *Journal of Public Economics*, vol.48, pp.185-205.

Vita, M.G. (1990) Exploring Hospital Production Relationships with Flexible Functional Forms, *Journal of Health Economics*, vol.9, pp.1-21.

Vita, M.G., Langenfeld, J., Pautler, P., Miller, L. (1991) Economic Analysis in Health Care Antitrust, *Journal of Contemporary Health, Law and Policy*, vol.7, pp.73-115.

Vitaliano, D.F. (1987) On the Estimation of Hospital Cost Functions, *Journal of Health Economics*, vol.6, pp.305-318.

Wagstaff, A. (1989a) Econometric Studies in Health Economics: A Survey of the British Literature, *Journal of Health Economics*, vol.8, pp.1-51.

Wagstaff, A. (1989b) Estimating Efficiency in the Hospital Sector: A Comparison of Three Statistical Cost Frontier Models, *Applied Economics*, vol.21, pp.659-672.

Wagstaff, A. and Lopez, G. (1996) Hospital Costs in Catalonia: a Stochastic Frontier Analysis. *Applied Economics Letters*, Vol.3, pp.471-474.

Wales, T.J. (1977) On the Flexibility of Flexible Functional Forms: An Empirical Analysis, *Journal of Econometrics*, vol.5, pp.183-193.

Wan, T., Vanostenberg, P.R., Salley, J.J., Singley, D.W., West, J.L. (1987) A Cost and Production Analysis of Hospital Dental Care Programs, *Public Health Reports*, vol.102, no.5 (Sept.-Oct.), pp.512-522.

Watts, C.A., Klastorin T.D. (1980) The Impact of Case Mix on Hospital Cost: A Comparative Analysis, *Inquiry (USA),* vol.17 (Winter), pp.357-367.

Weiss, L. W. (1964) The Survival Technique and the Extent of Suboptimal Capacity, *Journal of Political Economy*, vol. 72, pp.246-260.

Wilson, G.W., Jadlow, J.M. (1982) Competition, Profit Incentives, and Technical Efficiency in the Provision of Nuclear Medicine Services, *Bell Journal of Economics*, vol.13, no.2, pp.472-482.

Zaretsky, H.W. (1977) The Effects of Patient Mix and Service Mix on Hospital Costs and Productivity, *Topics in Health Care Financing*, vol.4, pp.63-83.

Zhu, J., Shen, Z.H. (1995) A Discussion of Testing DMU's Returns to Scale, *European Journal of Operational Research*, vol.81, pp.590-596.