1977

A Policy-oriented Analysis Of Hospital Costs

Alexander William Jenkins

A POLICY-ORIENTED ANALYSIS OF HOSPITAL COSTS

by

Alexander William Jenkins

Department of Economics

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Faculty of Graduate Studies

The University of Western Ontario

London, Ontario

March 1977

ABSTRACT

This thesis fulfills a threefold purpose. First, it develops an improved methodology for the statistical cost analysis of a sample of hospitals. Second, it investigates the cost structure of a sample of 101 Ontario hospitals during 1971. Third, it examines policy programmes which would increase hospital efficiency and reduce hospital output to socially optimal levels.

Improvements in the methodology of statistical cost analysis can be achieved in the areas of output definition and the specification and estimation of cost equations. Hospital output is services (case-types) whenever the hospital cost variable excludes (includes) the remuneration of attending physicians. Initial specification should formulate a total cost equation which reflects due attention paid to the form as well as the comprehensiveness of explanatory variables. Precise estimation should incorporate preliminary *a priori* aggregations, the principal components technique and stepwise regression.

Departmental costs are explained by services while all hospital costs are explained first by services then by case-types. The cost parameters estimated in departmental cost equations for individual services are employed as weights in constructing the output indices used in the service cost equation. All cost equations are efficiently estimated in an average cost form which avoids

heteroscedasticity. The degree of explanation of costs is greater and estimates of cost coefficients more precise than those of other cost studies. Services and case-types explain hospital costs equally well and indicate a similar structure of hospital costs. Capacity-related costs (40 percent) are somewhat larger while combined stay-related (35 percent) and admission-related (25 percent) costs are somewhat smaller than those indicated in other studies. Economies of scale are indicated for nuclear medicine and electrocardiography (ECG), but they are globally offset by diseconomies of scale in intra-hospital patient transportation and coordination so that standardized average costs in typical large (1,000-bed) hospitals exceed those of small (100-bed) hospitals by 8 percent. Scale-efficient hospitals providing both nuclear medicine and ECG are no larger than 100 beds.

Hospital regulators are encouraged to achieve cost savings through increased hospital efficiency and optimal output reductions. Increased scale-efficiency and the elimination of excess capacity would result in savings equivalent to 10 percent of all hospital costs. They are achieved through a centralization of nuclear medicine and ECG which exploits economies of scale; a hospital construction policy which favours small hospitals in order to avoid diseconomies of scale; and the elimination of as much as one-sixth of all hospital capacity (perhaps through selected hospital closure). Reductions of admissions and average length of stay (perhaps by 7 percent) would result in a net social benefit (perhaps 5 percent of all hospital costs) if they were achieved through capacity closure (rather than administrative fiat, reimbursement disincentives for hospitals or user charges for patients). Finally statistical cost estimates of the unit costs of individual services or case-types might be used in the incentive reimbursement of hospitals in order to promote increased allocative and X-efficiency. However, the possibility of inaccuracies in statistical cost estimates requires that the hospital regulator revise downward (upward) the reimbursement rate of any output for which unit costs are sufficiently overestimated (underestimated) as to result in an undesirable increase (decrease) in production.

ACKNOWLEDGEMENTS

This study would not have been possible without the assistance of a number of organizations and individuals. I would particularly like to thank the members of my dissertation committee at the University of Western Ontario - Drs. David Scheffman, John Palmer and David Burgess. I would also like to thank Mr. Donald Stuart of the Ontario Ministry of Health and those Ontario hospitals who permitted data access.

Financial assistance was provided by a Canada Council Doctoral Fellowship. Computing advice and facilities were provided by the Universities of Western Ontario and Alberta. Excellent secretarial assistance was provided by Ms. Mary Sorensen of the University of Alberta. Personal encouragement was provided by my wife, Kathleen, and by my infant daughters, Emily and Alanna.

TABLE OF CONTENTS

LIST OF TABLES

I. INTRODUCTION

The study of hospital economics is relatively new, there having elapsed little more than a decade since its birth. The structure of hospital costs is one issue, however, which is already controversial in that researchers such as Ingbar and Taylor (1968), Evans (1971) and Lave, Lave and Silverman (1972) differ in their views of the appropriate definition of hospital output, the importance of capacity costs and the existence of economies of scale. It is an important issue in view of the increasing involvement of government in the regulation of the hospital industry. Knowledge of the structure of hospital costs is vital to a cost-benefit analysis of policies which seek to curtail patient length of stay and admissions; to close designated hospitals; to reallocate special facilities among hospitals; to exercise control over the construction of additions to existing hospitals or of new hospitals; and even to adopt cost-based incentive reimbursement schemes for hospitals.

The purpose of this study is threefold: to develop an improved methodology for the determination of the structure of hospital costs in terms of the relative importance of stay-, admission- and capacity-related costs, the minimum fixed costs and any economies of scale associated with providing specific services and any overall economies of scale; to determine the cost structure of a sample of Ontario hospitals during 1971; and to examine the policy implications of the indicated structure of hospital costs. The vehicle used in achieving these objectives will be a statistical cost analysis[1] which incorporates considerable improvements over earlier cost studies in the areas of output definition and the specification and estimation of the hospital cost equation. Statistical cost analysis is thought to be superior to alternative methods of determining the structure of hospital costs such as engineering cost studies,[2] profitability studies[3] and the survivor technique[4] because it is better suited to the multi-product, not-for-profit nature of the hospital industry.[5]

The three major areas in which improvements can be made in hospital cost analysis are examined in the next three chapters. First, the choice of output definition is seen as dependent upon institutional arrangements between hospitals and attending physicians. If attending physicians are salaried or fee-for-service hospital staff members, as is approximately the case in the United Kingdom, then hospital output is "case-types". If attending physicians are private practitioners, as is approximately the case in Canada, then hospital output is "services" (which together with the inputs of attending physicians result in the production of illness remission for "case-types").

Second, the specification of hospital cost equations is examined. It should proceed systematically and without overdue regard for anticipated problems in estimation such as the inefficiency and imprecision arising from heteroscedasticity and multicollinearity respectively. In particular the hospital cost equation should be specified in a total cost form (although later modified to improve the efficiency of estimation) in order to ensure the most plausible forms for stay- and occupancy-related variables, and it should incorporate all plausible corrections for product heterogeneity (although later modified to increase the precision of estimation).

Third, the methodology used in estimating the hospital cost equation is examined. Ordinary least squares is typically applied to a weighted (usually average) form of the cost equation in order to avoid the inefficient estimation arising from heteroscedastic error terms, but only if statistical tests confirm the need for and validity of the weighting scheme. A synthesis of several aggregation techniques is applied to explanatory variables in order to reduce the imprecision of estimation arising from multicollinearity among the many plausible explanatory variables. Theoretically and empirically derived <u>a priori</u> information facilitate the aggregation of certain explanatory variables; for example, service output indices might be produced for a service cost analysis by weighting selected groups of services by the cost coefficients derived from departmental cost analyses. The principal components technique can be used to reduce the mean square error of parameter estimates. The stepwise regression technique can also be used, particularly if it avoids selectivity bias by evaluating explanatory variables in terms of their principal component parameter and standard error estimates.

A statistical cost analysis of a sample of 101 Ontario general hospitals is pursued in Chapter V using 1971 data. Improvements in the key areas of output definition and the specification and estimation of hospital cost equations are incorporated into the analysis of departmental, service and case-type costs. A superior degree of explanation and significance of cost coefficients are obtained for all cost categories. The degree of explanation of hospital costs and the insights into hospital cost structure are found to be relatively insensitive to the choice of the service or case-type output definition for the Ontario sample (but not necessarily for other samples). In relation to earlier cost studies, the indicated structure of hospital costs suggests larger capacity-related costs (40 percent of all hospital costs) and smaller stay- and admission-related costs (35 and 25 percent respectively of all hospital costs). Minimum fixed costs are apparent for one small department (nuclear medicine) and economies of scale for two small departments (nuclear medicine and ECG). They are, however, globally unimportant in relation to the diseconomies of scale thought to be associated with the large scale activities of intra-hospital patient transportation and general administration (standardized per diem costs are almost 8 percent higher in large 1,000-bed than in small 100-bed hospitals).

The policy implications of the indicated structure of hospital costs are discussed in a final chapter. They differ from those of other studies in that they are more specific and in one case contrasting (they recommend a small hospital construction policy). Cost savings which are of unambiguous net social benefit can be achieved through policy programmes which would increase hospital efficiency through a restructuring of special service and bed capacity. Economies of scale in ECG and nuclear medicine could be exploited through pooling arrangements within groups of geographically proximate hospitals. Diseconomies of scale in intra-hospital patient transportation and coordination could be avoided through a hospital construction policy which discouraged additions to existing large hospitals and ensured that new hospitals were small (100-300 beds). Excess special service and bed capacity (perhaps one-sixth of all capacity) could be reduced through closure of specified hospitals and increased utilization of other hospitals. Overall savings from increased scale-efficiency and the elimination of excess capacity would be about 10 percent of all hospital costs.

Cost savings which would be of less than proportionate net social benefit could be achieved through optimal output reductions (perhaps 7 percent). In particular, reductions in (elective) admissions and average length of stay could be achieved through administrative fiat, capacity closure, reimbursement disincentives for hospitals or user charges for patients. Since capacity costs are a substantial fraction (40 percent) of all hospital costs, output reductions would result in greatly enhanced savings if achieved through a proportionate capacity closure.

A final policy application of hospital cost analysis occurs in the incentive reimbursement of hospitals. Greater accuracy in estimating the unit costs of individual services or case-types ensures that a cost-based reimbursement scheme would provide incentives for increased allocative and X-efficiency. However, it is recognized that the hospital regulator must be prepared to monitor the output levels and decrease (increase) the reimbursement rates of services or case-types for which unit costs are sufficiently overestimated (underestimated) as to result in undesirably large increases (decreases) in output levels.

Footnotes

1. See Johnston (1960). All recent hospital cost studies utilize statistical cost analysis.
2. See Bain (1954) and Haldi and Whitcomb (1967).
3. See Scherer (1971, 79-82).
4. See Stigler (1958) and Shepherd (1967).
5. Hellinger (1975) examines specifications for hospital production functions. However, the not-for-profit nature of the hospital introduces the possibility of non-duality between production and cost functions.

II. OUTPUT AND COSTS IN HOSPITAL COST ANALYSIS

1. Introduction

In his discussion of hospital output, Sylvester Berki (1972, 31) asserts that "Many if not most hospital cost studies have floundered on the appropriate definition of those outputs whose costs are to be measured." Despite the perspicacity of his observation, however, he subsequently fails to generalize his discussion from a criticism of selected definitions of hospital output to the formulation of general criteria for choosing among alternative definitions of output for the analysis of the costs of a sample of hospitals.

Berki suggests that a "patient-day" definition of hospital output is inappropriate because patient-days of hospital care display considerable variation in per diem service levels or in the diagnoses of patients and are therefore heterogeneous in terms of resources consumed and related costs. He objects to a "service" definition of final hospital output on the grounds that services are an input (along with the resources of attending physicians) into the illness remission of patients suffering from illnesses falling into various diagnostic categories (or "case-types"). As an alternative, he advocates a "case-type" definition of final hospital output by virtue of its homogeneity and finality. He fails, however, to formulate general criteria for evaluating these three (and other) definitions of hospital output.

Such criteria might be stated in terms of two obvious requirements. First individual outputs in the vector of outputs defined to be that of the hospital must be homogeneous across all hospitals. Second, there must exist a correspondence between the vector of input levels associated with the observed level of hospital costs and that associated with the output - patient-days, services or case-types - defined to be no more and no less that of the hospital.[1] The first criterion would clearly disqualify a patient-day definition of hospital output since the resources consumed in the per diem treatment of seriously ill patients exceed those consumed in the per diem treatment of other patients. Although the first criterion is typically met by a sufficiently disaggregated taxonomy of services or case-types, the second criterion does provide a basis for choosing between the service and case-type definitions of hospital output. It would not, however, always favour the case-type definition of output advocated by Berki. It would in fact favour the service definition of output if observed hospital costs excluded the remuneration of attending physicians.

Having stated the general criteria which should pertain to the selection of a definition of output appropriate to hospital cost analysis, it is the purpose of this chapter to apply these criteria in providing rules for the selection of an appropriate definition of hospital output (especially services or case-types). The primary result is straightforward: if observed hospital costs totally exclude the remuneration of attending physicians, hospital output is services; if observed hospital costs totally include the remuneration of attending physicians, hospital output is case-types. Since, however, observed hospital costs often include some but not all of the remuneration of attending physicians, the cost variable may have to be computed as observed hospital costs less any reimbursement of attending physicians by hospitals (in a service cost analysis) or plus any reimbursement of attending physicians from non-hospital sources (in a case-type cost analysis).

In order to demonstrate the significance of these rules an examination of the primary organizational features of the hospital is required. The hospital is a composite of a number of factors of production, including the labour of nursing staff, technicians and maintenance personnel; material such as medication, dressings and food; capital such as building, beds and X-ray equipment; miscellaneous inputs such as land and energy; and the time inputs of staff physicians.

Physicians in general, whether salaried or fee-for-service hospital staff members or private practitioners who bill patients or their insurance agencies separately, fulfill three broadly-defined functions within the hospital. First, pathologists, radiologists and cardiologists undertake highly technical duties, primarily interpretative, in the laboratory, radiology and electrocardiography (ECG) departments. They are generally salaried or fee-for-service hospital staff members. Second, surgeons, obstetricians and anaesthesiologists undertake therapeutic procedures which, in combination with the operating and delivery room facilities and nursing staff provided by the hospital, result in surgical operations and deliveries. They are usually salaried staff members in British hospitals but private practitioners in North American hospitals.[2] Third, attending physicians, who may be general practitioners or specialists, are responsible for determining the length of stay of patients, the diagnostic tests, for example X-rays, laboratory tests and ECG's, which patients are to receive while in hospital, the final diagnosis of patients, and the therapy which patients are to receive, for example therapeutic X-ray, physiotherapy, surgery or delivery. In brief, attending physicians are responsible for combining their resources, primarily in the form of their time, with hospital services, for which they generate the demand, in order to produce illness remission in patients. As is the case for surgeons, obstetricians and anaesthesiologists, attending physicians are generally salaried staff members in British hospitals and private practitioners in North American hospitals.

Whether the factor composite which is designated as a hospital includes or excludes the resources of attending physicians is undoubtedly the organizational feature which most differentiates the hospital industries of the various national economies. It is also the consideration which most effects the appropriateness of possible definitions of hospital output for cost analysis.

If the hospital factor composite excludes the resources of attending physicians, and this exclusion is in turn reflected in hospital costs, then <u>services</u> are the end result of hospital activity and the definition of final output appropriate to the analysis of observed hospital costs. Services include the days of ward care, laboratory tests, surgery, delivery,[3] physiotherapy <u>et cetera</u> provided by the hospital.

Since both the resources of attending physicians and hospital services are necessary in producing illness remission in patients, it is only when the hospital factor composite includes the resources of attending physicians, and this inclusion is in turn reflected in hospital costs, that <u>case-types</u>, or, more properly, the illness remission of patients in such diagnostic or case-type categories as pneumonia, cancer and appendicitis, are the end result of hospital activity and the definition of hospital output which is relevant for hospital cost analysis.

Keeping in mind these distinctions, the discussion of the next section will examine the definition of output in hospital cost analysis. It is suggested that many cost studies have selected questionable output definitions, which is not surprising in view of their failure to apply the criterion of correspondence between the input levels associated with hospital output and those associated with the hospital cost variable. It will also be seen that in most organizational contexts, which in general lie between the stylized North American and British hospital industries, a conceptual correspondence between output and the dependent cost variables must be contrived through adjustments (in the form of the addition or subtraction of the remuneration of attending physicians derived from non-hospital or hospital sources respectively) to the reported values of hospital costs prior to cost analysis. The dependent cost variable in a case-type cost analysis should, in theory, include the remuneration of all attending physicians. It should exclude the remuneration of all attending physicians in a service cost analysis.

A concluding section will develop two models of hospital behaviour which further clarify the contexts within which the service and case-type definitions of output are appropriate. The two models also demonstrate the manner in which hospital output-cost relations are behavioural rather than purely technical.

2. Output and Cost Variables in Hospital Cost Analysis

Although several, ostensibly different definitions of hospital output have been employed in hospital cost studies, only the two already mentioned, services and case-types, are conceptually distinguishable. For example, Martin Feldstein (1968) suggests the case (admission) and patient-week as possible definitions of hospital output; since, however, he corrects for case-mix in both his per case and per patient-week cost equations, he implicitly uses a case-type output definition. Both Berry (1970) and Francisco (1970) use a patient-day definition of hospital output; since both authors group hospitals for regression purposes by similarity of facilities or service capabilities, they implicitly use a service output definition. Cohen (1970) uses a weighted patient-day definition of hospital output; since the components of the weighted patient-day output index are service output levels, he also uses a service output definition.[4]

That the output definitions used in hospital cost studies may not always correspond on a conceptual level to the hospital costs being analyzed is hardly surprising. Bearing in mind the organizational differences between British and North American hospitals, especially in terms of the preponderance of salaried staff members among the attending physicians in British hospitals, it seems implausible that the analysts of British and American hospital costs are both correct in using case-type output definitions. Since British hospital costs include the remuneration of attending physicians, Feldstein is correct in using the case-type output definition. Since North American hospital costs largely exclude the remuneration of attending physicians, Evans (1971) and Lave, Lave and Silverman (1972) are incorrect in their use of a case-type output definition to the extent that observed hospital costs understate actual case-type costs.

Other North American hospital cost analysts have used a service definition of hospital output. However, to the extent that observed hospital costs include even a small fraction of the remuneration of attending physicians, they overstate actual service costs. For example, since Kushner (1969) fails to exclude from his cost variable all remuneration of attending physicians (that falling under the category of the office of the medical staff), he also fails to establish a conceptual correspondence in terms of inputs between his (service) output and cost variables.

Although case-types are a more appropriate definition of final output within the context of the entire health sector, they are only appropriate for hospital cost analysis in two instances. They are appropriate first in the institutional settings for which observed hospital costs include the remuneration of all attending physicians; or second, if the cost variable to be explained is computed as observed hospital costs plus an estimate of the remuneration of any non-staff attending physicians (i.e. private practitioners).

Services can be defined as hospital output in two instances. They are appropriate first in institutional settings for which observed hospital costs exclude the remuneration of attending physicians; or second, if the cost variable to be explained is computed as observed hospital costs less the remuneration of any staff attending physicians. If physicians performing therapeutic procedures, especially surgeons, obstetricians and anaesthesiologists, are salaried staff members, surgery and delivery are defined as hospital services; otherwise the provision of operating and delivery room facilities and nursing staff are defined as hospital services.

The consequences of either an inappropriate choice of output definition, given the reported value of hospital costs, or alternatively an incorrect computation of the cost variable for a given choice of output definition, are twofold. If the case-type cost variable (either observed hospital costs alone or augmented by an estimate of the remuneration of non-staff attending physicians) understates actual case-type costs by failing to include the remuneration of all non-staff attending physicians, then the case-type cost analysis will generate downwardly biased case-type cost coefficients. In addition, inter-hospital comparisons of costs, standardized for case-types, will be biased since hospitals for which omitted remuneration is relatively small will appear more costly.

If the service cost variable (either observed hospital costs alone or reduced by an estimate of the remuneration of any staff attending physicians) overstates actual service costs by failing to exclude the remuneration of all staff attending physicians, then a service cost analysis will generate upwardly biased service cost coefficients. In addition, inter-hospital comparisons of costs, standardized for services, will be biased since hospitals for which the included remuneration is large will appear more costly.

An additional problem in output definition, or rather in the formulation of the dependent cost variable for the given output definition, is specific to case-type cost analysis. Most case-type cost analysts have limited their cost analysis to an examination of in-patient costs since case-type classifications typically extend to in-patients only. However, the in-patient cost variable has been computed in a manner which to some extent invalidates the cost analysis. In-patient costs have been computed as observed hospital costs less an estimate of out-patient costs. They therefore understate the actual case-related costs of those patients admitted as in-patients during their episodes of illness by the costs of any out-patient care which is used as a substitute, usually pre-admission or post-discharge, for in-patient care.[5]

The consequences of the failure of in-patient cost estimates to include all case-related costs, or alternatively, to exclude from observed hospital costs only the cost of treating patients on a purely out-patient basis during their episodes of illness, are twofold. First, case-type cost parameters are downwardly biased.

Second, inter-hospital comparisons of in-patient costs, standardized for case-types, are biased since hospitals failing to make relatively frequent partial substitutions of out-patient for in-patient care use more resources in in-patient care and therefore appear more costly.

Case-type cost studies should take one of two possible approaches to avoid this pitfall. First, they should endeavour to classify by case-type both in-patients and those patients treated on a purely out-patient basis, and then analyze all hospital costs. If this is not possible and the analysis is restricted to in-patient costs, they should compute case-type in-patient costs as observed hospital costs less only the costs of treating any patients on a purely out-patient basis during their episodes of illness.

The next section will formalize much of the previous discussion of service and case-type output definition by developing two alternative models of hospital behaviour. It will also make explicit the manner in which hospital cost-output relations, or cost functions, are behavioural rather than purely technical. To the extent that cost functions are behavioural rather than technical, the meaning of cost functions for policy purposes becomes ambiguous. However, the scope for technically inefficient behaviour may be circumscribed, particularly through the reimbursement mechanism.

3. The Behavioural Nature of the Hospital Cost-Output Relation

The behavioural nature of hospital cost functions was originally postulated by Evans (1971, 200-201). He recognizes "the possibility of systematic divergences between observed and "minimal" costs due to the behaviour patterns of hospitals." He asserts that the not-for-profit nature of the hospital will allow the hospital decision-maker to indulge himself in forms of behaviour which increase hospital costs. Hence the empirical results of cost analysis can only be viewed as "behavioural" cost-output relations.

Evans is not alone in his appreciation of the cost implications of the not-for-profit nature of hospitals. Other authors have, in fact, been concerned with the cost, output, quality, growth and even pricing implications of the not-for-profit nature of the hospital. For example, Newhouse (1970) was concerned with the implications of the utility maximization of the decision-maker in the not-for-profit hospital. He constructs a model in which the hospital decision-maker maximizes his utility which depends upon output quantity (patient-days) and quality (as measured by the utilization of certain factors of production, notably professional personnel and complex capital assets). He views the optimization as being constrained by the demand for hospital care and both minimum quality and profit restrictions. He predicts that costs are typically higher in not-for-profit than in proprietary hospitals.

Joseph (1974) constructs a similar but more elaborate model in which the expected number of patients turned away as a result of insufficient excess capacity becomes an additional variable affecting (dis)utility. Hospitals are more costly within the not-for-profit context because not only do decision-makers gain satisfaction from the quality of hospital care, but they also maintain excess capacity in order to avoid the dissatisfaction associated with turning patients away.

Most other models of hospital behaviour are similar to those of Newhouse and Joseph, usually differing only in terms of their designation of prestigious, quality-producing inputs. They are summarized by both Joseph and Jacobs (1974).

One model differs from these models because it questions the inevitability of the enhanced quality and inefficiency outcomes of the not-for-profit nature of the hospital. It does this through its explanation of how the not-for-profit characteristic of hospitals became commonplace even before the era of public ownership of hospitals. Pauly and Redisch (1973) suggest an alternative model of not-for-profit behaviour. They view the hospital as a physicians' cooperative which provides services which, together with the time inputs of physicians, result in the illness remission of patients, maximum income for physicians, and a break-even for hospitals. Since physicians have de facto control of the hospital, they will limit the extent to which the hospital administrator indulges himself in enhanced quality and resultant cost increases, since physicians' incomes would otherwise be reduced. The model is invaluable in that it not only provides an economic rationale for the not-for-profit characteristic and suggests limitations on the behavioural scope for hospital inefficiency, but it also examines the relationship, long-neglected both in the modelling and cost analysis of hospitals, between hospitals and physicians.[6]

In order both to examine further the question of output definition in hospital cost analysis and also to facilitate the discussion of the behavioural nature of hospital cost functions, two models of hospital behaviour will be developed. They admit the possibility of not-for-profit behaviour and attempt to capture the polar extremes of the organizational arrangements between hospitals and attending physicians. In the first model, the "service centre" model, attending physicians are private practitioners. In the second model, the "complete organism"[7] model, they are salaried hospital house staff.

A number of economic agents have some dealings in the hospital industry. Attending physicians undertake personal diagnostic and prescriptive activities and in so doing generate the demand for hospital services.

Referring physicians have exclusive dominion in determining whether patients are to be admitted and their actions are exogenous to both the hospital and attending physicians. Patients play a purely passive role and consent to any tests or treatment recommended by referring and attending physicians. Government or some other third party establishes reimbursement guidelines for hospital services or case-types. Finally, charitable agencies, through the financial limits of their philanthropy, together with solvency requirements, establish a minimum profit condition for the hospital.

The first model of hospital behaviour, the service centre model, views each economic agent as acting independently. In particular, attending physicians are private practitioners who direct bill the patient or his insurance agency for their personal services. The second model, the complete organism model, combines within the hospital its resources as a service centre with those of attending physicians. Attending physicians are salaried or fee-for-service hospital staff members. The first model approximates the institutional setting of North American hospitals, the second that of British hospitals.

The service centre model restricts the hospital decision-maker to maximizing his utility, U, through selection of I input levels, $\{F_i\}$, both productive, $\{F_i^p\}$, and unproductive, $\{F_i^u\}$. Productive and unproductive input levels affect utility both directly and indirectly through their effect on profits, P. Unproductive input levels might take the form of excess capacity, deluxe inputs, or overpaid staff members, and their indirect contribution to utility through profits is always negative. Hospital output is defined as the vector of J service output levels, $\{S_j^o\}$, which are exogenously determined by attending physicians. Technology is captured by a transformation function, T, of productive inputs into services. Profits are determined as the excess of hospital reimbursement, $\sum_{j=1}^{J} p_j S_j^o$, over costs, $\sum_{i=1}^{I} r_i F_i$. Minimum profits are defined as P_m. In symbols, the hospital decision-maker seeks to

$$\text{maximize} \quad U\,(P,\ \{F_i^p\},\ \ \{F_i^u\})$$

$$\text{w.r.t.} \quad \{F_i^{p,u}\}$$

$$\text{s.t.}\ (1) \quad T\,(\{F_i^p\},\ \{S_j^o\}) = 0$$

$$(2) \quad P = \sum_{j=1}^{J} p_j S_j^o - \sum_{i=1}^{I} r_i F_i$$

$$(3) \quad P \geqq P_m$$

For well-behaved preferences and technology, there exist optimal input levels, $\{F_i^* (\{S_j^o\}; \{r_i\}, \{p_j\}, P_m)\}$ i.e. the sum of $\{F_i^{p*}\}$ and $\{F_i^{u*}\}$, and a behavioural rather than a technical relation between costs and service output, $\sum_{i=1}^{I} r_i F_i^*$. This cost-output relation implies that service output gives rise to only those costs which are reflected in payments to those factors which constitute the hospital as a service centre; specifically, service costs exclude the remuneration of attending physicians. The cost function will be behavioural to the extent that the choice of input levels is allocatively and X-inefficient. Inefficiency occurs whenever outlays on productive inputs exceed minimum costs and non-productive input levels are positive. The extent of this inefficiency is, however, constrained both by the importance of profits to utility levels and by the levels of reimbursement rates and minimum profits.

The complete organism model ascribes greater latitude to the hospital decision-maker in his utility-maximization. The model allows him to select productive and unproductive levels of the I input levels $\{F_i^p\}$ and $\{F_i^u\}$, normally associated with the hospital in its role of a service centre; productive and unproductive levels of K attending physician inputs, $\{A_k^p\}$ and $\{A_k^u\}$; and productive and unproductive (totally unnecessary) levels of J services, $\{S_j^p\}$ and $\{S_j^u\}$. Technology, T, dictates the transformation of productive services and physician inputs into L exogenously-determined case-type output levels, $\{C_\ell^o\}$. Profits, P, are the excess of reimbursement, $\sum_{\ell=1}^{L} p_\ell C_\ell^o$, over costs $\sum_{i=1}^{I} r_i F_i + \sum_{k=1}^{K} q_k A_k$, and they must exceed some minimum level, P_m. In symbols, the decision-maker seeks to

$$\text{maximize} \quad U\left(P, \{F_i^p\}, \{F_i^u\}, \{A_k^p\}, \{A_k^u\}, \{S_j^p\}, \{S_j^u\}\right)$$

w.r.t. $\{F_i^{p,u}\}$, $\{A_k^{p,u}\}$, $\{S_j^{p,u}\}$

s.t. (1)' $T \ (\{F_i^p\}, \{S_j\}) = 0$

(2)' $T_A \ (\{S_j^p\}, \{A_k^p\}, \{C_\ell^o\}) = 0$

(3)' $P = \sum_{\ell=1}^{L} p_\ell C_\ell^o - \sum_{i=1}^{I} r_i F_i - \sum_{k=1}^{K} q_k A_k$

(4)' $P \geqq P_m$

The process involved in equation (2)' includes the activity of service demand generation, which, in contrast to the assumption of the service centre model, is entirely within hospital control. As a result, the hospital decision-maker has greater scope for his indulgence in utility-maximizing allocative and X-inefficiency.

For well-behaved preferences and technology, there exist optimal service centre and attending physician input levels,

$$\{F_i^* \ (\{C_\ell^o\}; \{r_i\}, \{q_k\}, \{p_\ell\}, P_m)\} \text{ and}$$

$$\{A_k^* \ (\{C_\ell^o\}; \{r_i\}, \{q_k\}, \{p_\ell\}, P_m)\} \text{ respectively,}$$

as well as optimal service output levels,

$$\{S_j^* \ (\{C_\ell^o\}; \{r_i\}, \{q_k\}, \{p_\ell\}, P_m)\}.$$

The relationship between case-type output levels and hospital costs, $\sum_{i=1}^{I} r_i F_i^* + \sum_{k=1}^{K} q_k A_k^*$ (which exceed service costs, $\sum_{i=1}^{I} r_i F_i^*$), implies that case-type output levels give rise to those costs reflected in payments to both service centre inputs and attending physicians. The cost function is behavioural rather than technical to the extent that the choice of input levels is allocatively and X-inefficient. Inefficiency occurs whenever outlays on productive inputs exceed minimum costs and unproductive input and service levels are positive. Although the scope for inefficiency is greater in the complete organism model, it is nonetheless constrained by the importance of profits to utility levels and by the levels of reimbursement rates and minimum profits.

Rarely will actual institutional arrangements in the hospital industry under investigation fulfill the assumptions of either the service centre or complete organism model, particularly in terms of the complete exclusion from or inclusion in reported hospital costs of the remuneration of attending physicians. Therefore a cost analysis must adjust the reported value of hospital costs downward by the reported value of any included remuneration, or upward by an estimate of excluded remuner- if the respective service or case-type definitions of output are to be employed.

Inter-hospital cost differences (even standardized for output) cannot necessarily be interpreted as arising from <u>inherent</u> efficiency differences,[8] since differences in reimbursement schemes could account for such differences. Cost comparisons of hospitals in different political jurisdictions, for example those undertaken by Berry (1970) in his study of a large sample of U.S. hospitals, are illegitimate to the extent that different reimbursement schemes prevail in different states. Clearly cost analysis should be restricted to a sample subjected to the same reimbursement scheme, or a sample in which corrections can be made for type of reimbursement scheme.

4. Summary

A definition of hospital output which is suitable for cost analysis must meet two requirements. First, individual outputs must be homogeneous across hospitals. Second, the vector of input levels which are reflected in the hospital cost variable must correspond to the vector of input levels which result in the vector of output levels. Services and case-types are alternative definitions of hospital output which generally meet the first criterion. Services meet the second criterion only if the hospital cost variable completely excludes the remuneration of attending physicians. Case-types meet the second criterion only if the hospital cost variable includes the remuneration of all staff and non-staff attending physicians. In an analysis of in-patient costs only, case-types meet the second criterion only if the in-patient cost variable also includes the costs of any treatment received by in-patients during their episodes of illness, especially pre-admission and post-discharge, in out-patient departments.

Models of hospital behaviour can be developed in order to depict the relationship between services and case-types and to suggest the "behavioural" rather than purely technical nature of the hospital cost equation. Service centre models exclude attending physicians and require that hospital output be services. Complete organism models include attending physicians who combine their resources with the services produced by the hospital in order to produce case-types. Although not-for-profit behaviour introduces the scope for allocative and X-inefficiency into both models, any residual profit motivation, along with restrictive reimbursement guidelines and minimum profitability, limit the extent to which the hospital cost function is behavioural and therefore vindicate a technical interpretation of the results of hospital cost analysis.

Footnotes

1. Pauly and Redisch (1973, 98) raise the same issue when they assert that "empirically observed cost curves [relating hospital costs to case-type output] may be misleading, if we add [or rather recognize that hospital costs fail to fully reflect] physician input."
2. See Klarman (1965, 132-133) for a comparison of the British and U.S. hospital industries.
3. If the hospital fails to maintain a salaried staff of surgeons, obstetricians and anaesthesiologists, it can only provide the facilities and nursing staff of operating and delivery room rather than the full-fledged services of surgery and delivery.
4. The weighted patient-day output, WPD_j, of the j th hospital is computed using information on hospital output of I services, S_{ij}, and the average cost of each service, $\overline{AC}_i$, for the entire sample. Specifically,

$$WPD_j = \sum_{i=1}^{I} (\overline{AC}_i / AC_1)\, S_{ij} \quad \text{where}$$

   i = 1 for ward care as measured by patient-days.
5. Davis and Russell (1972) have detected such substitution possibilities.
6. Melvin Reder (1965) is exceptional in this respect.
7. Philip Jacobs (1974) distinguishes between "organism" and "exchange" models of hospital behaviour in which

decision-making is respectively centralized and decentralized among hospital trustees, administrators, staff, patients and physicians. Since many organism models exclude physicians, the term "complete organism" was devised in order to differentiate organism models which include physicians.

8. For example those cost differences arising from returns to scale or factor market imperfections.

III. THE SPECIFICATION OF HOSPITAL COST EQUATIONS

1. Introduction

The specification of hospital cost equations creates problems in hospital cost analysis in addition to those associated with output definition. Specified forms of hospital cost equations in the literature have been unsatisfactory in several important respects. They have employed questionable forms for certain key variables, especially occupancy rate and length of stay. Some have de-emphasized the overwhelming importance of attaining product homogeneity through a sufficient number of output corrections. Finally, departmental, service and case-type cost equations are open to particular criticisms of their specifications. For example, case-type cost studies have not always examined with sufficient care the need for and selection of corrections for illness severity beyond those implied by the case-type classification of patients.

Each of the two general deficiencies will be discussed and remedies proposed in the sections which follow. A concluding section will discuss problems in specification which are specific to departmental, service and case-type cost analysis.

2. The Form of Variables Included in Hospital Cost Equations

In order to better establish the importance of specifying plausible forms for variables in cost equations, our discussion begins by specifying a total cost equation which assumes linearity in special services or case-types similar to that implicit in most previous hospital cost studies. Unlike these studies, however, it extends the linearity of total costs in capacity and output to the total rated patient-days (or rated bed-days), patient-days and admissions variables as well as the special service and case-type output corrections.

The justification for the linear form of total cost equations is twofold. First, to the extent that there are constant returns to scale and no interactions in production, or increasing or decreasing returns to scale and production interactions which are controlled for regression purposes, a total cost equation which is linear in output and capacity is quite realistic. Second, to the extent that inter-hospital variations in output mix are small, the linear form of the total cost equation is a good approximation of underlying cost conditions.

Whether hospital output is defined as services or case-types, it is well recognized that there exist the possibilities of factor indivisibilities, short-term or quasi-fixed capacity, returns to scale, and interactions in production. These give rise to respective minimum fixed costs (mfc), capacity costs, economies of scale, and interdependent marginal costs among outputs. If the total cost equation is linear in output, the variables required to control for each of these sources of inter-hospital cost variation include dummy constants, capacity measures such as rated patient-days (R), polynomials in output or capacity, and cross-product output terms, respectively.

In addition, if hospital output is defined to be services, marginal costs arise from the provision of ward service as measured by total patient-days (P), the service "admission and discharge" as measured by total admissions (A), and other (special) services, (S_j, j = 1, ...J), such as laboratory tests, X-rays and deliveries. If hospital output is defined to be case-types, marginal costs arise from the treatment of all patients (on the assumption of a standard diagnosis such as "normal delivery") as measured by total admissions (A), any differential treatment required by patients of non-standard diagnostic category, as measured by non-standard case-types, (C_ℓ, ℓ = 1, ...L), and, assuming complementarity between ward and special services, the total days of stay of all patients (P), since length of stay is only imperfectly related to case-type classification due to differences in illness severity.

A service cost equation might therefore be written as follows:

$$(1) \quad TC_s = a_1R + bP + cA + mfc + \sum_{n=2}^{N} a_nR^n + \sum_{j=1}^{J} d_jS_j + \sum_{i=1}^{I} \sum_{j=1}^{J} e_{ij}S_iS_j \;\; (i \neq j) + \text{error}$$

Similarly a case-type cost equation might be written as follows:

$$(2)\quad TC_c = a_1'R + b'P + c'A + mfc'$$

$$+ \sum_{n=2}^{N'} a_n'R^n + \sum_{\ell=1}^{L} d_\ell'C_\ell + \sum_{k=1}^{L}\sum_{\substack{\ell=1 \\ k\neq\ell}}^{L} e'_{k\ell}C_kC_\ell + \text{error}$$

Heteroscedasticity[1] of the error terms is to be expected, since large hospitals will surely have unexplained costs of a greater magnitude than those of small hospitals. Therefore efficient estimation[2] typically requires the reformulation of total cost equations prior to regression. The usual result is an average cost equation-per rated patient-day (TC/R), per patient-day or per diem (TC/P), or per admission (TC/A).

For simplicity of exposition, let us consider versions of the service and case-type cost equations which are truncated to include only the first three terms. Then equations (1) and (2) can be expressed in average cost form in order to determine the implied forms of key variables such as occupancy rate and length of stay. By dividing truncated total cost equations by R, P and A respectively, and remembering that occupancy or utilization rate is defined as U=P/R, and length of stay is defined as L=P/A, the following average cost equations result:

$$(3)\quad TC/R = a+bU+cU/L$$

$$(4)\quad TC/P = a/U+b+c/L$$

$$(5)\quad TC/A = aL/U+bL+c$$

It is obvious that there is no average cost equation which admits linear forms in <u>both</u> occupancy rate <u>and</u> length of stay. The per diem cost equation admits neither. Average cost equations which are linear in <u>both</u> occupancy rate <u>and</u> length of stay are therefore inconsistent with a (highly plausible) total cost equation which is linear in rated patient-days, actual patient-days and admissions.

The alternative specifications of average cost equations appearing in studies done by authors such as Ingbar and Taylor (1968), Ro (1968), Evans (1971) and Lave <u>et. al.</u> (1972) are questionable in several respects. First, since they employ linear forms for both occupancy rate and length of stay in their average cost equations, they are generally inconsistent with a total cost equation which is linear in rated patient-days, actual patient-days and admissions. Second, at least two authors recognize but are unable to correct the shortcomings of using a linear specification for occupancy rate in per diem and per admission cost equations. Both Ro (1968, 227) and Evans (1971, 210) recognize that the linear corrections for occupancy rate in their per diem and per admission cost equations fail to reflect the liklihood that the reduction in average costs associated with a given increase (5 percent) in occupancy rates will be larger when occupancy rates are low (50 percent) rather than high (90 percent). They might have avoided this shortcoming by specifying reciprocal rather than linear forms of occupancy rate. Third, most authors fail to adequately justify the form of variables appearing in their specifications. Evidence for this assertion is provided by comparing the two alternative cost equations used by most authors (for example, the per rated patient-day and per diem cost equations specified by Ingbar and Taylor). Both equations typically employ the same (linear) form of occupancy rate and length of stay, and thereby suggest without justification and perhaps without the intention of the authors that two different total cost equations apply to the same sample of hospitals.

In order to rectify this deficiency in specification, cost studies should undertake the initial specification of their cost equation in total form. The linear form suggested above is an obvious possibility. Only afterwards should they weight the total cost equations for efficient estimation.

3. Product Homogeneity in Hospital Cost Analysis

Product homogeneity across hospitals in cost analysis can be achieved by including sufficiently numerous and detailed output variables in cost equations to control for all output-related sources of systematic inter-hospital cost variation. This procedure is desirable not only for the estimation of the structure of hospital costs but also for inter-hospital comparisons of standardized costs, although the latter may still be possible if information is available on the nature of any remaining product heterogeneity. For example, if large hospitals appear to have lower standardized costs on the basis of an insufficiently detailed output classification and if it is known that large hospitals produce outputs of higher quality and costliness than do small hospitals, then large hospitals are unambiguously less costly than small hospitals.

Product homogeneity may not always be attainable in hospital cost analysis for two reasons. First, the most detailed output classification available may still be inadequate for product homogeneity. Second, the number of output corrections actually employed may be less than that of the most detailed output classification available because the latter is thought to be so large in relation to sample size as to result in multicollinearity and the imprecise estimation it entails, or even insufficient degrees of freedom for estimation.

The hospital cost analyst typically has no control over the availability of a detailed output classification, at least in the short run. If the best available output classification is still too aggregated, he may proceed with his examination of hospital costs, perhaps employing quality indices to control for unremoved product heterogeneity, and always issuing the necessary caveats. If sufficient information is available on any residual product heterogeneity, he may still be able to make some inter-hospital comparisons of standardized costs.

The cost analyst should never sacrifice output corrections in order to mitigate the impact of insufficient degrees of freedom or multicollinearity. He should employ one or more legitimate aggregation techniques (discussed in the next chapter) to reduce the dimensionality of the data matrix and achieve more precise estimation, especially those techniques that generate and use a priori information or aggregate independent variables using sound statistical techniques. He should never ignore "minor" output corrections or aggregate independent variables on the basis of highly subjective or arbitrary criteria.

Many examples of such questionable aggregation techniques can be found throughout the hospital cost literature. Martin Feldstein (1968, 35-39) begins his case-type cost analysis by specifying 33 detailed case-types which he then reduces to 9 aggregated case-types in order to both permit computation by a small-capacity computer and increase the precision of estimation which was otherwise limited by multicollinearity. Although he suggests the possibility of aggregating the detailed case-types on the basis of similarity of the cost coefficients from the full regression, he chooses instead to aggregate case-types on the basis of his (layman's?) perception of similarity of medical character. By using such an arbitrary criterion for aggregating case-types he may forego any reasonable degree of product homogeneity in his cost analysis.

A further example is found in the study done by Ro (1968). Ro begins his service cost analysis by specifying 25 output variables which he believes will affect hospital costs. He subdivides these variables into six groups: size-volume, capacity utilization, scope of services, technology, exogenous and (teaching) dummy. He requires that a minimum of one variable from each group be included in the cost equation, and then undertakes stepwise regression on the remaining variables in each group. He does this to reduce the number of variables and attendant multicollinearity without as much "selectivity bias"[4] as usually accompanies stepwise regression. However, the crudeness of his aggregations is recognized, especially in the preliminary categorization of variables, since he resorts to using a quality index in both his per diem and per admission cost equation. He measures quality inversely as the number of patient-days per hospital personnel, and in so doing he unforunately corrects for an input (labour) and therefore biases his cost analysis to the extent that labour and other factors of production are substitutes.

A further example of questionable aggregation technique which results in a loss of product homogeneity is found in the study done by Kushner (1969). He regresses "hotel-type" hospital costs (ward nursing costs, dietary expenses, and laundry and linen and housekeeping costs) on alternative polynomials in actual and rated patient-days. By so doing he fails to recognize that some components of hotel-type costs are incurred in the provision of services other than ward service, especially laboratory, radiology, operating and delivery room services. In addition, he regresses all hospital costs on broadly-defined services, for example, days of ward care, X-rays and laboratory tests undifferentiated by type. In neither of his cost analyses did he investigate all possible output corrections. He reduces multicollinearity but fails to achieve product homogeneity.

Other authors have shown much more concern for the need to achieve product homogeneity and have therefore selected alternative aggregation techniques for dealing with multicollinearity. However, they have not always investigated all possible corrections necessary for product homogeneity. For example, Evans fails to examine corrections in his case-type cost analysis for indirect teaching costs or the costs of preferred accommodation. Perhaps the enthusiasm of many cost analysts for the speculative introduction of such "characteristic"[5] variables is dampened by their awareness of multicollinearity.

Clearly the initial specification of hospital cost equations should include all available output corrections while temporarily ignoring the possibility of imprecise estimation arising from multicollinearity. If it does not, product homogeneity may be unnecessarily sacrificed, since valid empirical techniques do exist for reducing the dimensionality of the data matrix to a level commensurate with precise (and relatively unbiased) estimation.

4. <u>Departmental, Service and Case-Type Cost Equations</u>

All cost equations should adhere to the general guidelines for specification outlined in the previous sections. In addition, departmental, service and case-type cost equations present specification problems particular to organizational levels and institutional contexts.

a. <u>Departmental Cost Equations</u>

Individual hospital departments include ward services, newborn nursery, laboratory, radiology, physical medicine and rehabilitation, emergency ward, operating and recovery room, and delivery room. Minor departments are sometimes separately organized for electrocardiography (ECG), electroencephalography (EEG), and nuclear medicine. Whenever minor departments are not separately organized, major departments, especially the laboratory and radiology department, provide any related services.

In most North American hospitals, the operating and delivery room provide the facilities and nursing staff used by surgeons and obstetricians who are privileged private practitioners. In those hospitals, however, where surgeons and obstetricians are salaried hospital staff members, the departments should be designated as surgery and delivery respectively. Costs should include the remuneration of medical as well as nursing staff, and output is defined as completed surgical and obstetrical procedures (differentiated by type) rather than the provision of facilities and nursing staff.

Individual hospital departments provide a number of services and their cost equations should therefore incorporate detailed output corrections. For example, the laboratory provides various types of laboratory tests (bacteriology, biochemistry, blood bank, cytopathology, haematology, histopathology and urinalysis); autopsies; and sometimes ECG, EEG and nuclear medicine. Physical medicine and rehabilitation provides physiotherapy and occupational and speech therapy. By way of contrast, Feldstein (1968, 81-86) regresses departmental costs on polynomials in bed capacity and cases per bed per year rather than detailed service outputs. Ingbar and Taylor (1968, 67-73) regress departmental costs on the same set of highly aggregated service variables used in their service cost regression. Cohen (1973) regresses departmental costs on a polynomial in a uni-dimensional output measure (laboratory tests for laboratory, X-rays for radiology) and teaching affiliation.

The service outputs of departments are not always readily identifiable. ECG, EEG and nuclear medicine are provided by the laboratory in one hospital, by the radiology department in another, and by a separately organized department in a third. Needless to say, outputs must be assigned to individual departments prior to cost analysis. Certain obstetrical and minor surgical procedures may be performed in the operating room, delivery room, or emergency ward, depending upon the hospital. If it is impossible to identify the department providing these facilities or services, departments may have to be aggregated in order to produce an unambiguous correspondence between output and costs.

The largest and most complex department within the hospital is that of ward service. Not only do wards provide stay-related services proportional to patient-days, whether short- or long-term, intensive care, medical, surgical, obstetric, psychiatric or paediatric, but they provide a number of admission-related services which include the services "admission and discharge" and trips off-ward to special service departments such as radiology or the operating room.

It is especially important to include in departmental cost equations variables which reflect departmental cost structures: for example, capacity measures such as rated patient-days for ward service, bassinets set up for newborn nursery, and number of operating rooms for surgery; a dummy variable valued at unity if the department exists, zero otherwise; a polynomial in departmental output or capacity; departmental teaching measures, for example the number of technicians trained by the laboratory; output variables for closely related departments to capture any interactions; and proxies for unremoved product heterogeneity, for example medical teaching affiliation. The signs, magnitudes and significances of the estimated coefficients of these variables provide insights into the overall structure of hospital costs which might be impossible from the service cost equation as a result of insufficient degrees of freedom or irremediable multicollinearity. These insights, moreover, can be used to restrict the service cost equation to an empirically manageable form.

Ideally the costs of individual departments should include both direct staff and material and indirect overhead costs. However, the latter, if reported, may be unreliable, and if unreported, may not be imputable with acceptable accuracy. Overhead costs are evaluated by Anderson (1974) as a markup, determined by prior measurement, on direct costs; and by Cohen (1973) as the same fraction of all hospital overhead costs as that of departmental to hospital income. In the absence of an accurate estimate of overhead (indirect) costs of each department, Feldstein is constrained to comparing reported (direct) departmental costs with output. This last procedure provides some insights into the structure of departmental costs (i.e. the structure of the direct cost component) and these insights will be unbiased to the extent that overheads are complements with direct inputs, or differences in substitution across hospitals and services are small and randomly distributed. In addition, to the extent that the distribution of overheads among the outputs of a department is approximated by a constant mark-up on direct costs, direct cost equations imply restrictions on the form of the service cost equation.

b. The Service Cost Equation

The initial specification of the service cost equation should include all variables thought to contribute to departmental costs, as well as any variables affecting overall hospital costs, for example teaching activity or preferred accommodation. It should adhere to the general guidelines specified in the previous sections. The specification may be subsequently modified by aggregations which incorporate the results of the departmental cost regressions.

Service costs should exclude from total hospital costs the remuneration of attending physicians and any other costs unrelated to service production, for example the costs of education, research, and (self-financing) coffee shops and parking. They should also be adjusted for any factor price differences other than those arising from differences in factor quality. For example, the interest component of capital costs may be highly variable among hospitals because some hospitals have ready access to low or even zero-price charitable funds. Hence service costs might best exclude interest costs, particularly if they are small and infrequently observed and if capital is thought to be complementary to other factors of production.

c.. The Case-Type Cost Equation

Case-type output variables include the numbers of patients treated in each homogenous case-type category, as well as corrections for the structure of hospital costs similar to those outlined in the previous sections, especially capacity measures, dummy variables and polynomials in output or size for specific case-types. If the case-type classification is of such a sufficiently aggregated nature that all hospitals report approximately proportionate positive output in all case-type categories, a single capacity measure (rated patient-days), dummy variable (valued at unity for all hospitals) and polynomial (in total admissions, patient-days or rated patient-days) are plausible corrections for the structure of hospital costs. If inter-hospital comparisons of standardized costs are to be interpreted as inter-hospital differences in efficiency, the value of the capacity variable must be viewed as exogenous to hospitals, perhaps as a result of either exogeneity of case-mix and patient length of stay or the administrative fiat of the hospital regulator. Otherwise the maintenance by the hospital of capacity considered by the regulator to be excessive for its case load should be viewed as an indication of inefficiency in addition to that suggested by positive cost residuals.

Specifications of case-type cost equations should carefully investigate corrections for illness severity beyond those implied by the case-type classification of patients in order to avoid biasing the cost analysis. They should avoid using corrections for illness severity, for example the variables "numbers of patients having major or minor surgery" employed by Lave, Lave and Silverman (1972), which also correct for a service (surgery). To the extent that surgery is a substitute for other services, especially chemo-, radio- and physiotherapy, in treating certain case-types, case-type cost coefficients will be downwardly biased and inter-hospital comparisons of standardized costs will be biased in favour of hospitals making such substitutions. Should a need be indicated for such ambiguous corrections for illness severity, a critical reappraisal of the case-type output classification might be in order. For example, the case-type classification used by Lave <u>et. al.</u> (which employed only 17 categories) seems overly aggregative when compared with the 33 categories employed by Feldstein (1968) and the 41 employed by Evans (1971).

The most frequently utilized correction for illness severity in case-type cost analysis has been the total patient-days variable, which takes the form of a length of stay correction in average cost equations. While it is true that extended lengths of stay, corrected for case-mix, may signify that ward care is substituted for special services within a hospital, the value[6] and statistical significance of the coefficient of the length of stay variable used by Evans and by Lave <u>et. al.</u> and the excess length of stay variable used in the per admission cost analysis of Chapter V suggest that an extended length of stay is primarily the result of illness severity. Length of stay has proved to be a necessary correction for illness severity in all case-type cost analyses to date, regardless of the number of categories in the case-type classification.

An additional correction for illness severity is implied by corrections for the age-sex composition of hospital patients. However, detailed age-sex corrections by case-type are so numerous as to be empirically unmanageable, since the number of corrections equals the product of the number of age-sex and case-type classifications. Aggregated age-sex corrections unrelated to case-type were used by Evans in his per admission cost equation, but their impact on hospital costs is necessarily ambiguous. Since the primary impact of age-sex composition on costs appears to occur through length of stay differences, as was evident in Evans' study when he found that age-sex corrections were insignificant in his per diem cost equation, then the need for age-sex corrections is not obvious in any event.

In order that the case-type cost coefficients reflect stay- and capacity-related costs as well as admission-related costs, the corrections for total patient-days and rated patient-days might be alternatively expressed as "excess patient-days" and "unused rated patient-days" respectively. The former is computed for each hospital by algebraically summing the differences between the length of stay of each patient and the average length of stay for patients of that diagnosis across all hospitals. The latter is computed for each hospital as the difference between rated and total patient-days. The cost coefficient of the excess patient-day variable will reflect the per diem costs of both the normal ward and extra ward and special services provided to patients during extra days of stay.

Ideally the case-type cost variable should include the remuneration of all attending physicians as well as the service-related costs of the hospital. The case-type output classification should extend to both those patients admitted as in-patients during their episodes of illness and those patients treated on a purely out-patient basis. In practice this has not been possible due to lack of data; in particular, case-type costs have excluded some of the remuneration of attending physicians, and the cost analysis has been confined to in-patient costs alone.

If case-type costs are computed as observed hospital costs, and if observed hospital costs include the major portion of the remuneration of all attending physicians (i.e. salaried staff outnumber private practitioners), then the case-type cost analysis will minimally understate true case-type cost coefficients and will permit valid inter-hospital comparisons of standardized costs if the omitted remuneration is small and randomly distributed across hospitals. If however, observed hospital costs exclude the major portion of the remuneration of attending physicians (i.e. private practitioners outnumber salaried staff), and particularly when the omitted remuneration is not randomly distributed (e.g. only teaching hospitals maintain salaried staff), case-type costs should be computed as observed hospital costs less the remuneration of salaried medical staff. The case-type cost analysis will more substantially understate true case-type cost coefficients but will permit valid inter-hospital comparisons of standardized costs if any substitution between physician and hospital services in treating case-types is small and randomly distributed across hospitals.

If the case-type classification extends to in-patients alone, then the case-type cost analysis might be restricted to in-patients alone. However, the in-patient case-type cost variable should include not only reported in-patient costs but also the costs of any out-patient treatment which is received, usually pre-admission or post-discharge, by in-patients during their episodes of illness. In practice it may not be possible to identify the recipient of out-patient care, so case-type in-patient costs are frequently computed as observed hospital costs less an estimate of all out-patient costs. Estimated out-patient costs might be computed as out-patient income or out-patient services weighted by the service cost parameters estimated in a prior service cost analysis. They should not be computed as the reported values of organized out-patient and emergency costs alone, since out-patient costs also include the much larger costs of out-patient services provided in special service departments such as the laboratory and radiology department.[7] The value of in-patient costs computed as total less out-patient costs understates actual in-patient case-type costs and cost parameters, but if the understatement is small and randomly distributed across hospitals, valid inter-hospital comparisons of standardized costs may still be possible.

5. Summary

The specification of hospital cost equations should proceed without overdue regard for anticipated problems in estimation (although care must subsequently be taken to weight the total cost equation in order to reduce the inefficiency of estimation associated with heteroscedasticity and to aggregate explanatory variables in order to reduce the imprecision of estimation associated with multicollinearity). The specification of a total rather than average cost equation seems more likely to ensure plausible forms for such key variables as occupancy rate and length of stay in the (equivalent) average cost equation. Given the plausibility of a total cost equation which is linear in rated patient-days, actual patient-days and admissions, the linear form of both occupancy rate and length of stay in per rated patient-day, per diem and per admission cost equations is questionable; reciprocal forms of occupancy rate and length of stay might be especially desirable in the per diem cost equation. The specification of the hospital cost equation should initially include all plausible corrections for output heterogeneity and only afterward should relatively unbiased aggregations be pursued in order to increase the precision of estimation while maintaining product homogeneity.

Departmental, service and case-type cost equations must control for the structure of hospital costs by admitting corrections for minimum fixed costs, capacity costs and economies of scale as well as for output mix. Cost variables should be adjusted to offset any factor price differences other than those arising from differences in factor quality. Departments provide services and departmental costs should include a reliable imputation of overheads if possible. Service costs should entirely exclude the remuneration of attending physicians. Case-type costs should include the remuneration of all staff <u>and</u> non-staff attending physicians. In addition, a case-type analysis of in-patient costs should ensure that the in-patient cost variable includes any episode-related out-patient costs. Finally, in order to avoid biasing the cost analysis, case-type cost equations should carefully investigate (and avoid if possible) any corrections for illness severity (such as the incidence of surgery) which potentially correct for an input (especially services which are substitutes for other services, for example some types of surgery with radio-therapy) into the illness remission of case-types.

Footnotes

1. This is reflected in variance-covariance matrices for error terms which are diagonal but not scalar-diagonal matrices. Typically a priori information indicates a relationship between the variances of the error term and one or more of the independent variables. There are several statistical tests for this phenomenon. See Johnston (1972, 214-221).
2. Efficient estimation requires the application of generalized least squares. See Johnston (1972, 208-211).
3. Multicollinearity exists when there is significant correlation between (linear combinations of) independent variables. See Johnston (1972, 159-168).
4. Selectivity bias arises when regressors are non-orthogonal and the results of stepwise regression, in terms of which variables are retained as being statistically significant on the basis of an F-test of incremental R^2, are sensitive to the order in which additional variables are introduced into the regression.
5. This term was introduced by Lave, Lave and Silverman (1972) to designate those variables which were not, strictly speaking, output corrections (either case-types or services).
6. The values of coefficients are highly comparable to the marginal costs of a patient-day of ward care which are estimated in the service cost analysis of

Chapter V.

7. Evans (1971) appears to have missed this point and so his estimate of out-patient costs grossly understates actual out-patient costs.

IV. MULTICOLLINEARITY IN HOSPITAL COST ANALYSIS

1. Introduction

Multicollinearity[1] is a third problem in hospital cost analysis. It arises whenever, for a given sample, explanatory variables are highly collinear or, alternatively, (linear combinations of) explanatory variables are highly correlated. It can best be detected through an examination of the size distribution of the eigenvalues associated with the correlation matrix of the explanatory variables.[2] If the number of eigenvalues (there are as many eigenvalues as explanatory variables) found to be insignificantly different from zero[3] is large in relation to the number of explanatory variables, then a high degree of multicollinearity is present. If one or more eigenvalues is zero, then multicollinearity is perfect.

The consequences of multicollinearity are several. As Johnston (1972, 160) points out:

> "1. The precision of estimation falls so that it becomes very difficult, if not impossible to disentangle the relative influences of the various X [explanatory] variables. The loss of precision has three aspects: specific estimates may have very large errors; these errors may be highly correlated, one with another; and the sampling variances of the coefficients will be very large.
>
> 2. Investigators are sometimes led to drop variables incorrectly from an analysis because their coefficients are not significantly different from zero, but the true situation may be not that a variable has no effect but simply that the set of sample

data has not enabled us to pick it up.

3. Estimates of coefficients become very sensitive to particular sets of sample data, and the addition of a few more observations can sometimes produce dramatic shifts in some of the coefficients."

Multicollinearity is an especially serious problem in hospital cost analysis as a result of the large number of initially specified explanatory variables in relation to sample size. For example, Ingbar and Taylor (1968) initially specify over 100 service output variables as potentially determining 144 observed values of hospital costs; Cohen (1970) defines 15 service output variables (which were aggregates of 26 more homogeneous services) in his explanation of 46 observed values of hospital costs; Evans (1971) specifies 41 case-type, 20 age-sex, and several characteristic variables (length of stay, occupancy rate, rated bed capacity) as determinants of 185 observed values of hospital costs; and Lave, Lave and Silverman (1972) specify 17 case-type and 15 characteristic variables as determinants of 65 observed values of hospital costs.

Since multicollinearity is the result of both the data and the specified form of the cost equation, the hospital cost analyst may respond to it by expanding his data base or changing the cost equation so that regression estimates of key parameters are sharpened. However, he may not be able to pursue the former course of action, since "in many situations the econometrician has no control over his data";[4] moreover, he may find it difficult to determine the manner in which any additional data can be used to reduce the standard errors of estimate of parameters; and he may in any event find that additional data reduce multicollinearity little or not at all.

The hospital cost analyst might in fact react to the suggestion that he expand his data in the same manner in which the French mob must have reacted to a recommendation that they eat cake instead of bread. He is already confronted with a data set which barely provides sufficient degrees of freedom to estimate the initially specified form of the model, let alone minimize multicollinearity. He therefore typically acquires and uses all available data even before he begins to worry about multicollinearity. Hence he must resort to respecifying the cost equation if he is to reduce multicollinearity.

The cost analyst seeks a revised specification of his cost equation which permits more precise estimates of key parameters (i.e. parameters have smaller mean square errors). He typically modifies his specification by aggregating explanatory variables through the use of one or more of the techniques discussed in the next section. His aggregation technique can be defended in terms of the largeness of the reduction in the variance of parameter estimates in relation to any bias (i.e. any <u>net</u> reduction in the mean square errors of parameter estimates).

Too little attention has been given to a constructive criticism of aggregation technique in hospital cost analysis. Criticism of the many techniques will be provided in the next section. An improved response to multicollinearity, which synthesizes several aggregation techniques, will be outlined in a concluding section.

2. Aggregation Techniques: Responses to Multicollinearity

There are almost as many different approaches to dealing with the multicollinearity problem as there are hospital cost studies. Eight approaches can be differentiated, including the intuitive a priori aggregation employed by Feldstein and Kushner (mentioned in the previous chapter). Each seems to rely almost exclusively on a single aggregation technique. Each will be discussed within the context of the study in which it is (best) applied. Particular attention is paid to the extent of any bias introduced into estimates of cost parameters by aggregation techniques.

a. Intuitive A Priori Aggregation

Many cost analysts undertake "plausible" implicit a priori aggregations of explanatory variables by opting for an output classification less detailed than its available alternatives. For example, Evans (1971) employs an output classification which includes 41 case-types which are an aggregation provided by the Ontario Hospital Services Commission of 260 Ontario Broad Code case-types.

Lave, Lave and Silverman (1972) use a case-type classification which aggregates the 1,000 I.C.D.A. (International Classification of Diseases, Adapted) case-types into a mere 17 case-types. These authors are justified in opting for the aggregated case-type classifications only if the criteria used by medical experts in ge[illegible] case-type aggregations coincide with similar mar[illegible]st criteria.

In some cases hospital cost an[illegible] deliberately aggregate explanatory variables usi[illegible] other criteria than their own perceptions of similar costliness. For example, Martin Feldstein aggregates 33 case-types into 9 on the basis of what he calls "medical similarity". Kushner aggregates the detailed service outputs reported by Ontario hospitals into broadly defined service categories by excluding available output corrections for the different types of laboratory tests, X-rays, _et cetera_. Hospital cost analysts can be criticized whenever they make "intuitive" _a priori_ aggregations which _unnecessarily_ detract from product homogeneity and thereby introduce the possibility of substantial bias into their analysis.

b. The Facilities Approach

Berry (1970) and Francisco (1970) use the facilities approach in their attempts to finesse the problem of multicollinearity. This approach involves subdividing the hospital sample into sub-samples within which hospitals provide either the same combination or total number of "facilities" (the capabilities of providing such broadly-defined services as surgery, laboratory, X-ray, *et cetera*). They then regress, for hospitals in each sub-sample, per diem costs on a constant, occupancy rate, and a reciprocal term and polynomial in total patient-days. For example, Francisco subdivides a sample of 1,328 U.S. hospitals into 25 groups (each having at least 30 hospitals) on the basis of their providing the same combination of 17 different facilities. He points out the incidence of a negative coefficient for total patient-days in 22 of the 25 estimated per diem cost equations in order to demonstrate the existence of economies of scale in the general hospital industry.

The facilities approach makes one very critical assumption which seems unrealistic. It assumes that hospitals providing the same combination of facilities all provide services in the same proportions, for example .10 physiotherapy treatments, .03 X-ray treatments, .09 ECG's, *et cetera* per patient-day. Unless this assumption is borne out by sample properties, the facilities approach provides a biased and inadequate explanation of hospital costs, since hospitals providing a greater per diem volume of each service will appear more costly, and the $\bar{R}^2$ will be low. A modified version of the facilities approach undertaken by Francisco illustrates this last point: his regression of per diem costs (for the entire sample) on 17 dummy variables (one for each total number of facilities) yields an $\bar{R}^2$ of only .5, which is inferior (assuming comparable samples) to the $\bar{R}^2$ of .7 achieved by Ingbar and Taylor (1968) through their application of factor analysis in their study of the per diem costs of hospital services.

c. Stepwise Regression

Stepwise regression begins by regressing hospital costs on a subset of all possible explanatory variables, in particular those that are "theoretically obvious". It then includes additional explanatory variables if they increase the R^2 of the equation in a statistically significant manner. The technique is valid if explanatory variables are orthogonal or, alternatively, if multicollinearity is absent. However, it admits "selectivity bias" into the regression in that it will incorrectly reject certain explanatory variables in the presence of multicollinearity; moreover, the identity of rejected variables is sensitive to the initial choice of "theoretically obvious" variables.

Almost all hospital cost studies use stepwise regression in that they reject any statistically insignificant variables for which inclusion cannot be strongly justified on a theoretical basis. For example, most studies reject a dummy constant in equivalent total cost form (or reciprocal output term in average cost form); some reject polynomials in output; and others reject such characteristic variables as rural location or religious affiliation. They are justified in doing this only to the extent that multicollinearity is weak and little theoretical support can be found for including statistically insignificant variables in the regression.

One study, however, relies almost entirely on stepwise regression to reduce multicollinearity. Ro (1968) begins by subdividing 27 service and characteristic variables into 6 groups: size-volume, capacity utilization, scope of services, technology, exogeneous and dummy (education). He selects one variable from each group for regression purposes, that which is most highly correlated with per diem costs. He then introduces (and rejects) all remaining variables using the stepwise technique.

Several criticisms are possible. The output corrections implied by the 27 explanatory variables are overly crude; for example, X-rays are undifferentiated by type (tuberculosis, fluoroscopic, cineradiographic *et cetera*). Some output corrections are questionable as they correct for inputs rather than outputs. For example, personnel per 100 beds corrects for factor mix as well as output complexity and, as a result, the cost analysis is biased in that hospitals substituting labour for capital appear less costly. The choice of 6 groups is arbitrary, as is the rationale for initially including in the regression only one variable from each group. Finally, the stepwise technique is generally suspect in view of the presence of multicollinearity.

d. Factor Analysis

Factor analysis[6] is a specific type of cluster analysis.[7] It seeks to replace the original set of independent variables with a subset which are both orthogonal and capture the preponderance of the variation in the original set of independent variables.

Factor analysis begins by generating a normalized data matrix of n observations on k explanatory variables, X. It next computes the correlation matrix of the data matrix, X'X. It then computes the k x k matrix of normalized eigenvectors, A, and k eigenvalues, Λ (in diagonal form), where $A'X'XA = \Lambda$. It then computes the mutually orthogonal factors or principal components of the data, $Z = XA$, which fully reflect the variation in the data. It then "identifies" ℓ factors ($\ell \leq k$) for which factor loadings of (or correlations with) explanatory variables are large (for example in excess of .3). Finally it retains for cost regressions the ℓ explanatory variables, one per identified factor, which have the highest factor loadings.[8]

Both Ingbar and Taylor (1968) and Lave, Lave and Silverman (1972) use factor analysis in their cost studies. The latter apply the technique to case-type variables only. They all admit that the criteria for identifying factors are subjective. However, they fail to fully recognize the bias created by the under-representation in the cost regression of those variables which are poorly correlated with identified factors. They also fail to obtain information on the structure of hospital costs which is as interpretable or as detailed as possible for policy applications (see subsection g. below).

e. Average Cost Aggregation

Cohen (1970) aggregates the 15 service outputs of each sample hospital, S_{ij}, where i refers to service and j to hospital, to produce a weighted patient-day output index, where weights are the average cost of each service in relation to that of a day of ward care (the service for which i = 1). He implicitly proceeds as follows: first, the total cost of the jth hospital, TC_j, are related to service output in the following manner:

$$(1) \quad TC_j = mfc + \sum_{i=1}^{15} a_i S_{ij} + u_j$$

where a_i is the average unit cost for all 46 sample hospitals of providing the ith service and u_j is a random error. In symbols,

$$(2) \quad a_i = \sum_{j=1}^{46} TC_{ij} / \sum_{j=1}^{46} S_{ij}.$$

Equation (1) can be rewritten as follows:

$$(3) \quad TC_j = mfc + a_1 \sum_{i=1}^{15} (a_i/a_1) S_{ij} + u_j$$

or

$$(4) \quad TC_j = mfc + a_1 WPD_j + u_j$$

where WPD_j is the weighted patient-day output of the jth hospital. Equation (4) is now estimated.

A number of criticisms are possible. First, by implicitly opting for the cost model given by equation (1), Cohen assumes an overly simplistic structure of hospital costs. Specifically he assumes negligible capacity costs or equal utilization rates for services; factor indivisibilities for only those services which are common to all hospitals; and when he specifies an additional squared term in weighted patient-day output, returns to scale which are of a similar magnitude for all services. Second, he accepts as accurate the reported full costs of each service by hospitals, TC_{ij}; to the extent that overheads are on average imputed by hospitals to specific services in an arbitrary manner, he employs inaccurate average cost weights, a_i, and therefore admits bias into the cost regression.

f. Similar Marginal Cost Aggregation

Although Martin Feldstein (1968, 39) first suggests the use of similar marginal cost aggregation, Lave, Lave and Silverman (1972) are the first to employ it in hospital cost analysis. They begin by regressing the 1967 per admission costs for 65 hospitals on the full set of 17 case-type and 15 characteristic variables. They then sort the case-type variables into 5 groups on the basis of similar magnitudes of estimated cost coefficients. Specifically they view 11 case-types as having costs which are insignificantly different from the numeraire case-type; 3 case-types as having costs which are \$5-\$10 above those of the numeraire; one case-type as having costs $25 above those of the numeraire; and two case-types as having costs $45-$184 below those of the numeraire. For a different data set (1968, in order to eliminate pretest bias), they then add total case-types within each group for each hospital in order to generate the five aggregated case-type variables used in the regression of 1968 per admission costs.

Similar marginal cost aggregation as used in this study is of questionable value since it seems especially likely to introduce bias into the regression. Three reasons are apparent. First, the use of a specific number (5) of similar marginal cost categories receives no thorough justification in terms of a trade-off between reduction of the variances and unbiasedness of the estimated cost parameters of aggregate variables. Second, the use of the imprecise case-type cost estimates for 1967 as weights is questionable. Third, aggregation by similar rather than actual point estimates of marginal costs is also questionable since it "throws away" information. In addition, the application of the aggregation technique to case-type variables does nothing to deal with multicollinearity among the 15 characteristic variables.

g. Principal Components

Evans (1971) and Lave, Lave and Silverman use the principal components technique. Prior to regression the former replaces 40 case-type output corrections with the 10 principal components (or factors) corresponding to the 10 largest eigenvalues, and the 20 age-sex corrections with 6 principal components corresponding to the 6 largest eigenvalues. The latter replace 16 case-type output corrections with 5 principal components corresponding to the largest 5 eigenvalues. Neither replaces the characteristic variables with principal components and neither retrieves estimates for the cost coefficients of the original set of case-type and age-sex variables.

The rationale for using the principal components technique is straightforward: the number of mutually orthogonal principal components used in the regression, ℓ, is generally less than the number of explanatory variables, k. If the exclusion of one or more principal components results in a reduction in the variances of the parameter estimates which is large in relation to any bias, multicollinearity is less and estimation of the parameters of characteristic variables is more precise (parameter estimates have smaller mean square errors) than would be achieved using straightforward ordinary least squares (OLS).

The technique as used by Evans and Lave <u>et. al.</u> proceeds as follows. First it is conjectured that the vector of n hospital costs (in deviation form), y, is linearly related to the normalized data matrix, X, of n observations on k case-type variables, and to the data matrix (also in deviation form), W, of n observations on k' characteristic variables.

(5) $y = XB + W\delta + \text{error}$

Since by the definition of the principal components,

(6) $Z = XA$, then for A orthogonal (which must be the case since A is the matrix of eigenvectors of X'X),

(7) $X = ZA'$.

Substitution of (7) into (5) gives the result:

(8) $y = ZA'B + W\delta + \text{error}$

or (9) $y = Z\gamma + W\delta + \text{error}$

Now select an $n \times \ell$ subset of the principal components, Z*, which captures a large portion of the variation in Z (and X)—which will be the case if selected principal components have relatively large eigenvalues—and, more important, explains y in a manner which is no worse in a statistically significant sense than does Z (and X)—which will be the case if excluded principal components are statistically insignificant. Finally regress an equation of the form:

(10) $y = Z^*\gamma^* + W\delta^* + \text{error}^*$

If the principal component aggregations have been undertaken in a manner which results in reductions in the variances of the estimates of δ (the parameters of the characteristic variables) which are large in relation to any bias (for example, by including in the regression only those principal components which have large eigenvalues or are statistically significant), estimates of δ will be more precise (have smaller mean square errors) than those found using OLS alone.

It should be noted at this point that it is possible to retrieve "principal component" estimates of the parameters (B) and standard errors for the case-type variables. Such a procedure would greatly increase the interpretability of regression results, and would encourage the application of the principal components technique to all collinear variables (both X and W). Although neither Evans or Lave et. al. use this procedure, Kendall (1957) recommends it while McCallum (1970) and Cheng and Iglarsh (1976) suggest that its judicious use will result in estimates of the parameters of the original model (in this case, the vector B) which are more precise (have smaller mean square errors) than the straightforward OLS estimates.[9] However, Johnston (1972, 329-331) points out that the retrieval procedure should be reserved for applications where multicollinearity is less than perfect.

The retrieval or "unscrambling" step of the principal components technique proceeds as follows. First, since it is definitionally the case that (for $\hat{\gamma}^*$ the OLS estimates of γ),

$$(11) \quad \hat{\gamma}^* = A^{*\prime} \hat{B}^*,$$

then for A^* orthonormal (any n x ℓ partition of A is necessarily orthonormal)

$$(12) \quad \hat{B}^* = A^*\hat{\gamma}^*$$

and $\hat{B}^*$ are the principal component estimates of B.

Next it is noted that the variance-covariance matrix of $\hat{B}^*$, Σ_{BB}, is related to that of $\hat{\gamma}^*$, $\Sigma_{\gamma\gamma}$, as follows:

$$(13) \quad \Sigma_{BB} = A^*\Sigma_{\gamma\gamma}A^{*\prime}$$

and the vector of standard errors of estimate of $\hat{B}^*$ (S_B) is formed from the square roots of the diagonal elements of Σ_{BB}. Finally, the T-statistics of the principal component estimates of B are found by pairwise division of the elements of $\hat{B}^*$ by those of S_B.

A number of criticisms are possible of the manner in which Evans and Lave et. al. employ the principal components technique. The criteria which they use in determining the identity of the principal components which are to be included in and those that are to be excluded from the regression seem vague or arbitrary. For example, Evans fails to specify the level of statistical significance used in excluding principal component regressors (surprisingly large totals of 30 case-type and 13 age-sex principal components are excluded). Lave et. al. fail to even discuss the criteria used in excluding 11 of 16 principal components. In both cases, the authors include a number of principal components which captures a mere 70 percent of the variation in the explanatory variables.

Neither study takes any steps to reduce multicollinearity among the characteristic variables. They should logically have extended the application of the principal components technique to encompass the characteristic as well as case-type and age-sex variables. They might then have used the "unscrambling" process implied by equations (12) and (13) to retrieve parameter estimates and standard errors for all explanatory variables. By so doing they would have sharpened the estimates of the coefficients of the characteristic variables while maintaining interpretability and even extending it to the case-type cost coefficients.

h. Information Theory

Evans and Walker (1972) apply this technique in their case-type cost analysis of Ontario and B.C. hospitals. They construct complexity and specialization indices for each sample hospital. They compute the former by weighting each case-type by the extent to which its incidence throughout the sample is non-uniform, with non-uniformly distributed (infrequent) case-types receiving a large weight, and uniformly distributed (common) case-types receiving a low weight. They compute the latter by weighting each case-type by the logarithm of its frequency in individual hospitals in relation to its frequency in the entire sample, so that hospitals treating disproportionate numbers of infrequent case-types have a high specialization index. They then regress per admission costs for each hospital on the complexity and/or specialization index along with a number of characteristic variables such as teaching affiliation and occupancy rate.

Several criticisms are possible. First, the authors attach weights to case-types which are the outcomes of a very strong and potentially unrealistic assumption: that the information function is logarithmic. Second, by employing the complexity index in the cost equation they implicitly assume that the complexity weight of each case-type has an impact on marginal case-type costs which is invariant with respect to case-type. For example, if only one hospital in the sample treats case-type x (which is k percent of all hospitalized case-types), while another is the only hospital treating case-type y (which is also k percent of all hospitalized case-types), then the weights for case-types x and y would be identical. Unless the two case-types had similar marginal costs, case-type cost analysis would be biased. Third, by employing the specialization index in their cost equation, the authors implicitly anticipate that hospital costs are always higher whenever disproportionate numbers of unusual case-types are treated. Since infrequency and costliness of case-types are conceivably related through a complex rather than simple monotonic functional form, the relevance of the specialization index to an explanation of hospital costs is questionable. Finally, the technique is perhaps overly aggregative in dealing with multicollinearity in that it reduces information on hospital output of 98 case-types to one or two output indices.

3. A Synthesized Approach to Multicollinearity

No single aggregation technique seems to be entirely satisfactory in modifying the hospital cost equation so as to reduce the standard errors of estimate of cost parameters in a relatively unbiased manner. A new approach is needed. It must avoid both the tendency to rely on a single technique and the subjectivity or arbitrariness characteristic of the application of most techniques.

The approach found to be most satisfactory (in the next chapter) in dealing with multicollinearity begins by specifying a cost equation which is as detailed as the data permit. In particular, it specifies output and related structural variables for homogeneous case-types or services. It should employ a case-type or service taxonomy which adheres to the principle of similarity of production technique.

The approach next undertakes any exclusions or aggregations which can be strongly justified on a theoretical level. For example, if medical science suggests that two case-type taxonomies, one less detailed than the other (compare the 41 Relative Stay Index and 260 Ontario Board Code case-types), explain hospital costs in a similar manner, then it should employ the less detailed taxonomy for regression purposes. Similarly, it might use a standardized or weighted service output taxonomy (previously developed by medical science) for certain broadly-defined types of laboratory tests or physiotherapy procedures, rather than more detailed output taxonomies. Finally, if hospitals are entitled to reimbursement at competitive wage rates for both lay and non-lay staff members, even though portions of the salaries of the latter are appropriated for "good works", it can initially exclude corrections for religious affiliation on an *a priori* basis.[10]

The approach then undertakes any exclusions or aggregations which can be soundly justified on an empirical basis. For example, it can restrict the form of the service cost equation by using any insights into the costs of different types of laboratory services (haematology, blood bank, urinalysis, autopsy, *et cetera*) provided by analysis of the costs of the laboratory department. It can aggregate into an output index the different types of laboratory services by similar or actual values of marginal costs in relation to those of a departmental numeraire (such as a standard unit of urinalysis). It might exclude some output corrections altogether if their cost coefficients in the departmental cost analysis were small and statistically insignificant. However, if reported departmental costs exclude overheads, it must subsequently ensure that aggregations and exclusions remain valid in the regression of all hospital costs. It can do this by including corrections for non-numeraire or excluded outputs in a stepwise manner designed to detect statistically significant differences in the incidences of overheads among the outputs of a department.

The approach must next construct a set of principal components which correspond to all explanatory variables emerging from the preliminary aggregation stage. It regresses hospital costs on as many principal components as are important in explaining the variation among the explanatory variables (i.e. have relatively large eigenvalues) and, in a statistically significant manner, the variation in hospital costs. It might determine statistical significance through an F-test of the incremental R^2 associated with the introduction of additional principal component regressors (in descending order of eigenvalue, since eigenvalues reflect the contribution of related principal components to the variation in the explanatory variables). It then computes principal component estimates of the cost coefficients of the explanatory variables, $\hat{B}^*$, by linearly transforming the estimated cost coefficients of the principal components, $\hat{\gamma}^*$ (subsection 2.g. confirms that $\hat{B}^* = A^*\hat{\gamma}^*$, where A^* are the relevant eigenvectors of the correlation matrix of the explanatory variables). It also retrieves standard errors of estimate and T-statistics[11] for the $\hat{B}^*$'s (again see subsection 2.g.).

The objective of the application of the principal components technique as an intermediate step in estimation is to generate estimates of B which are more precise or which have smaller mean square errors than the straightforward OLS estimates. The minimum mean square error (of parameter estimates) criterion for determining the identity of the principal components to be included in the regression is difficult to apply in the usual case of unknown true parameter values (B). Guidelines suggested by McCallum (1970, 111) and Gheng and Iglarsh (1976, 230-233) for applying the criterion include recommendations that included principal components demonstrate statistical significance in explaining the dependent variable, excluded principal components have relatively small eigenvalues, and the principal component estimates ($\hat{B}^*$) fall within the pre-specified confidence intervals for parameter values which are suggested by <u>a priori</u> information. The first two guidelines are synthesized by the requirement that included principal components be statistically significant on the basis of an F-test of the incremental R^2 associated with groups of principal components introduced into the regression in descending order of eigenvalue. The third guideline appears in this study to be redundant since the further addition (or deletion) of principal components seems unwarranted in terms of any need to alter the signs of coefficients to a priori signs or the magnitudes of coefficients to values more consistent with a priori information. For example, the principal component estimates found for the service cost equation by including the first 9 principal components as regressors (justified by the first two guidelines) are as or more consistent with expectations of a priori sign and with parameter values estimated in departmental cost analyses than are the principal component estimates found by including more (or less) principal components as regressors.

The approach next excludes from further consideration any explanatory variables for which a strong theoretical case and statistical significance are absent. It also repeats the regression, including and testing for the significance of any explanatory variables which were excluded from the initial regression as a result of a weak theoretical or empirical justification (and retains them if statistically significant). It is justified in using both backward and forward stepwise regression to the extent that preliminary aggregations and the principal components technique reduce multicollinearity in a relatively unbiased manner. The approach then estimates the cost equation in its final form, again using principal components as an intermediate stage in estimation.

The overall approach seems to avoid the subjectivity and arbitrariness associated with any single technique. It is able to do this because it applies each technique in an objective manner. For example, principal components are retained in the regression if they are significant at the 5 percent level. This objectivity is in turn made possible by the fact that the overall approach is, through its diversity, able to avoid placing the entire onus in dealing with multicollinearity upon a single aggregation technique. Diversity therefore facilitates the production of more precise estimates of the hospital cost structure.

4. Summary

Multicollinearity results in imprecise estimation. It is especially severe in hospital cost analysis because the number of output corrections is usually large in relation to sample size. Approaches to dealing with multicollinearity (aside from modifying the composition of the sample) in hospital cost analysis include the use of aggregation techniques such as intuitive *a priori* aggregation, the facilities approach, stepwise regression, factor analysis, average and marginal cost aggregation, principal components and information theory. Each has deficiencies but the most serious is the likelihood that its exclusive use introduces substantial bias into estimated cost coefficients whenever a large reduction is to be undertaken in the variances of parameter estimates (which are large as a result of multicollinearity). In addition, most techniques result in an insufficiently interpretable cost analysis.

An alternative approach to dealing with multicollinearity synthesizes several aggregation techniques and thereby reduces the need for using any single technique to such an extent as to introduce substantial bias into the estimation. Variables for which the theoretical case for inclusion is weak, for example religious affiliation, are initially excluded from the regression. Groups of service output variables are aggregated into output indices for the service cost analysis by weighting them by the cost coefficients estimated for them in prior departmental cost analyses. Principal components is used as an intermediate step (in order to maintain interpretability) in the estimation of all departmental, service and case-type cost equations. Finally, stepwise regression is used, both backward and forward, to eliminate any statistically insignificant and theoretically unjustifiable explanatory variables. The synthesized approach is thought to result in parameter estimates which are more precise (have smaller mean square errors) than those achieved using ordinary least squares either straightforwardly or in conjunction with any single aggregation technique.

Footnotes

1. See Johnston (1972, 159-168) for a general discussion of multicollinearity and a review of the literature.
2. See Johnston (1972, 102-105) for a discussion of the manner in which the eigenvalues (and eigenvectors) of a matrix are computed. The terms "eigen", "latent", and "characteristic" are interchangeable when referring to roots (values) and vectors.
3. See Johnston (1972, 327) for a test which identifies eigenvalues which are (sufficiently alike to be) treated as having true values which are equal to zero. The existence of k_o such eigenvalues suggests that variation of the explanatory variables is negligible in the corresponding k_o dimensions.
4. See Johnston (1972, 165-168) for a discussion of this point and the use of additional data in "best breaking the multicollinearity deadlock".
5. See Wonnacott and Wonnacott (1970, 309-312).
6. See Scott (1966).
7. See Bijnen (1973).
8. This description of the procedure is developed in Johnston (1972, 322-331).
9. See the next section for a discussion of the criteria which should be met in the "judicious" use of the principal components technique.
10. An objection can arise if it is conceivable that religious-affiliated hospitals show a "labour bias".

in production in order to augment funds available for good works. If this were the case, religious-affiliated hospitals would be allocatively inefficient and more (rather than less) costly than other hospitals. However, the observation that religious-affiliated hospitals employ large numbers of fully-paid (lay) nursing personnel suggests that low-wage non-lay nursing personnel are scarce. Therefore religious-affiliated hospitals are confronted with the same marginal conditions for the employment of labour as are other hospitals. This possibility, along with minimum (reported) profitability constraints, any difficulty in substituting labour for other factors of production, and the statistical insignificance for religious affiliation in explaining costs (found in the next chapter), all suggest that "labour bias" is negligible and that religious-affiliated hospitals have the same cost structure as other hospitals.

11. The ratios of principal component estimates to standard errors are, strictly speaking, interpretable as T-statistics only to the extent that principal component estimates are relatively unbiased.

V. A HOSPITAL COST ANALYSIS, ONTARIO, 1971

1. Introduction

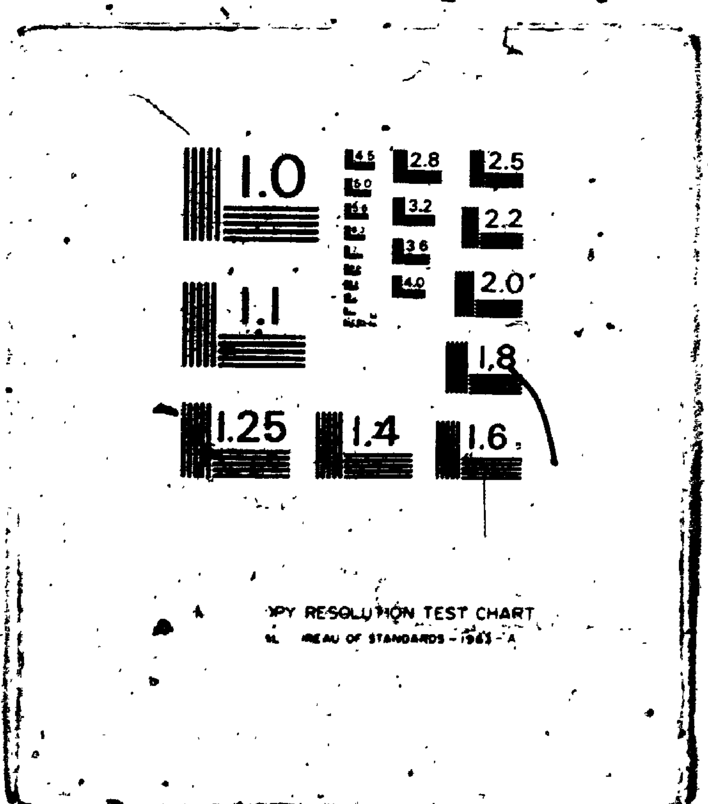
The statistical cost analysis of a sample of hospitals should not only accurately explain the structure of hospital costs but also should provide information which is sufficiently reliable as to be employed in the cost-benefit analysis of hospital policy issues. It is therefore of considerable importance that the cost analysis successfully overcome problems in the areas of output definition and the specification and estimation of the hospital cost equation.

Solutions to these problems in hospital cost analysis are outlined in the previous three chapters. In particular, if and only if the cost variable excludes the remuneration of attending physicians should services be defined as hospital output. If the cost variable includes the remuneration of attending physicians, then case-types are defined as hospital output. The appropriate specification of hospital cost equations initially formulates them in a total cost form which incorporates all plausible output corrections. Only afterward does it modify the cost equation in order to deal with heteroscedasticity (by weighting, usually resulting in an average cost form) and with multicollinearity (by aggregating). Finally, multicollinearity should be confronted with a diversified aggregation approach which incorporates theoretically and empirically justified a priori aggregations, the principal components technique employed as an intermediate step in estimation, and stepwise regression.

Alternative departmental, service and case-type explanations of hospital costs can be compared with those explanations of hospital costs achieved in other studies in order to determine the merits of the proposed solutions to problems in hospital cost analysis. They also provide insights into the structure of hospital costs, especially the relative importance of stay-, admission- and capacity-related costs, minimum fixed costs, economies of scale and any indirect teaching costs. These insights are of considerable importance for hospital policy.

Our cost analysis of a selected sample of Ontario hospitals during 1971 will investigate and compare three cost-output relationships. First, it will examine the relationship between the service output and reported costs of individual hospital departments such as ward service, laboratory and radiology. Second, it will examine the relationship between service output and costs which are computed as reported hospital costs excluding the remuneration of attending physicians. Third, it will examine the relationship between case-type output and costs. However, as a result of data deficiencies it will restrict the case-type cost analysis to in-patients and to the hospital component of case-type costs (which excludes the physician component of case-type costs - the

2 OF/DE 3



ARC - 374

remuneration of attending physicians).

The examination of departmental costs has a threefold purpose. First, it provides insights into the structure of reported departmental costs. Second, it facilitates empirically justified *a priori* aggregations in the service cost equation in the form of output indices and exclusions. Third, it provides a means of confirming the insights into the structure of hospital costs which are suggested by the service cost analysis.

The service cost analysis is the most important of the three cost analyses for two reasons. First, the data are most favourable to accurate measurement of the service output and cost variables. Second, our hospital sample approximates the "service centre" rather than the "complete organism" model in that the preponderance of attending physicians are private practitioners. The service cost analysis is therefore thought to provide the most accurate insights into the structure of hospital costs.

The case-type cost analysis is restricted to an examination of the *hospital* component of case-type costs since data on the remuneration of attending physicians are generally unavailable. The analysis of the hospital component of case-type costs (which, as a result of data deficiences, computationally coincides with service costs) has a threefold purpose. First, it provides insights into the structure of the largest (hospital) component of case-type costs (which are unbaised to the extent that hospital and physician inputs are complements in the production of case-type illness remission, or any inter-hospital differences in substitution small and randomly distributed, and any inter-hospital differences in the behaviour of attending physicians are small and randomly distributed). Second, it provides a means of further confirming the insights into the structure of hospital costs which are suggested by the service cost equation. Third, it facilitates an evaluation of the empirical validity of the alternative service and case-type output definitions in explaining related hospital costs.

The discussion will proceed by first outlining the preliminary stages of the cost analysis: the choice of sample, data sources and properties, cost and output variables and regression technique. It then presents and interprets the results of the departmental, service and case-type cost regressions. Finally, it examines the relative empirical appropriateness of the service and case-type definitions of output in hospital cost analysis. An examination of the policy implications of the estimated structure of hospital costs is postponed until the next chapter.

2. Preliminary Stages of Cost Analysis

a. Sample Selection

Hospital cost analysis seeks to examine the relationship between hospital costs and homogeneous outputs for a given level of technology and factor prices. Hence, it must ensure not only that individual service or case-type outputs are homogeneous (of similar quality), but also that sample hospitals face the same technology and factor prices. It should therefore draw the sample from the same geopolitical area. In addition, although it can be pooled cross-sectional time series (observations on a number of hospitals are made at different points in time), it typically is purely cross-sectional (observations on a number of hospitals are made at the same point in time), since in the former case there arises the need for additional corrections reflecting the change over time in technology, factor prices and output quality.

A number of readily verifiable properties should be associated with the chosen geopolitical area. Technology should be uniform as a result of both good informational flows and similar legal restrictions on feasible hospital production techniques. Factor markets should be homogeneous as a result of similar educational and licensing requirements and industry-wide unionization (or lack of it) for non-medical staff; and identical prices (as a result of similar taxes, tariffs and transportation costs) should prevail for non-labour inputs. In addition, similar quality of service and case-type output for all hospitals is further assured by uniform legal restrictions, reimbursement guidelines and patient characteristics.

The cost analysis should exclude or make necessary allowances for any candidate hospitals within the chosen geopolitical area which present anomalies in terms of technology, factor prices and output quality. It might restrict the hospital sample to short-term general hospitals by excluding chronic care, psychiatric and other specialized hospitals (for which costs could be separately analyzed). It might restrict the sample to a size range for which service or case-type output is relatively uniform.

The sample employed in this study consists of 101 Ontario Hospital Services Commission (OHSC) Group A and B general hospitals during 1971. All but 10 of the total of 111 OHSC Group A (teaching) and B (non-teaching) hospitals (ranging in size from 100 to 1,000 beds) granted the permission necessary for data access. Approximately one hospital in six is a teaching hospital involved in the training of medical interns and residents. Almost all hospitals are involved in the training of nurses, nursing aides and assistants, and laboratory and radiology technicians. Almost one-half of all hospitals have some minor involvement in long-term chronic and convalescent care.

The sample is deliberately drawn from a single geopolitical area-the province of Ontario. Since Ontario (which provides the largest provincial sample of hospitals within Canada) is relatively homogeneous with respect to technology, factor markets, population and legislative environment[1] its general hospitals have the properties desirable in cost analysis. However, small hospitals of less than 100 beds are excluded from the sample as there is reason to believe that they produce a lower quality output than larger hospitals. The major reasons for this belief are the differentiation in roles between small and large hospitals which is suggested by the failure of small hospitals to always provide such basic services as laboratory or to treat certain complex case-types; the frequency with which small hospitals transfer seriously ill case-types to larger, better-equipped hospitals; the observation that small hospitals are under-involved in teaching programmes and over-involved in long-term chronic care; and the separate categorization by the OHSC of small hospitals as Group C.

b. Data Sources and Properties

Data for the year 1971 are drawn from a number of sources. Data on service production, costs and characteristic variables are obtained from Statistics Canada Forms HS-1 and HS-2[2] which are submitted annually by all Canadian hospitals. Summaries were obtained from the OHSC (since subsumed by the Ontario Ministry of Health) of the number of surgical procedures (in each of 26 categories) performed in each sample hospital. Summaries of total admissions and days of stay for patient in each of 41 case-type categories are available for each hospital in the Relative Stay Index Study.[3]

A number of consultations were undertaken in order to clarify reporting and medical terminology and hospital practices. At various times throughout the study consultations were held with officials of the OHSC; hospital accountants, nursing staff and laboratory and radiology personnel of hospitals in London, Ontario; and physicians in the Department of Epidemiology of the University of Western Ontario.

The data have a number of desirable properties. Reporting conventions are well-established, as Ontario hospitals have been required to submit annual reports during the past several decades and to follow standardized accounting procedures.[4] Instructions and definitions for the annual reports are sufficiently detailed and comprehensive to be intelligible to even outside researchers. The data are quite detailed if for no other reason than their use in the reimbursement of hospitals since the introduction of the provincial hospitalization scheme during the late 1950's.

However, the data are still deficient in several respects. Some hospitals report output but not costs for some departments; for example, one hospital failed to report the incidence of nursing costs among departments such as ward service, newborn nursery, delivery room *et cetera*. Clearly the departmental cost analysis has to exclude such hospitals.

Other hospitals report some of the costs for a particular service under one departmental category and the remaining costs under another. For example, some hospitals reported medical staff salaries and fees for the inter-pretation of ECG's as the only expenditures under the heading of ECG Departmental Expense. Since the remaining ECG costs (technician salaries and supplies and other expenses) are certainly positive, a further examination of the data is in order. As ECG output in such cases is reported by the laboratory, it is apparent that the total expenditures reported by the laboratory include the non-medical portion of ECG costs. In such cases, the analysis of the costs of the ECG department must exclude any hospi-tals ascribing ECG costs to more than one department, and an analysis of laboratory costs must exclude an estimate of any incurred ECG costs.

As is discussed in Chapter II, the reported values of departmental costs exclude overheads such as house-keeping expense and building maintenance and depreciation. Since any attempt to impute overheads to departments seems arbitrary, departmental cost analysis is accordingly restricted to an examination of the relationship between the direct costs and service output of individual hospital departments.

The distinction made between in- and out-patient services in certain departments, especially the emergency ward, is sometimes arbitrary, since hospitals exercise some discretion in classifying patients for reimbursement purposes. If, for example, the services rendered to in-and out-patients by the emergency ward differ, and if patient classification is somewhat arbitrary, the analysis of the costs of the emergency ward is biased in that any net tendency to favour either the in-patient or out-patient classification will result in cost parameters for that classification which are biased in the direction of the cost parameters of the disfavoured classification.

The case-type output data fail to distinguish between patients having single and multiple diagnoses. They therefore do not permit a maximum degree of product homogeneity unless, of course, the unlikely possibility arises that the pattern of multiple diagnoses is similar among all hospitals.

The data report the remuneration of only those attending physicians who are salaried staff members. They therefore do not permit a good estimate of total case-type costs.

All deficiencies aside, the data are far superior to those available or used in previous hospital cost studies. They are, in fact, more than adequate in some cases. For example, the detailed surgical procedures data prove to be unnecessary in explaining the (hospital portion of the) costs of the operating room. Unfortunately, the sample size does not share a similar over-adequacy.

c. Cost Variables

Cost variables are needed for departmental, service and case-type cost analysis. The direct costs of individual departments are computed as the sum of reported nursing and technician salaries and supplies and other expense. In addition, laboratory, radiology, ECG, EEG and nuclear medicine costs include the reported remuneration of related medical staff (pathologists, radiologists and cardiologists). The costs of the operating and delivery room and emergency ward are combined since there is an overlap in their provision of certain obstetrical and minor surgical procedures. The costs of ECG, EEG and nuclear medicine are subjected to departmental cost analysis only if departments are separately organized. The costs of the laboratory and radiology department exclude estimates of the costs of any ECG, EEG or nuclear medicine they provide; these estimates are computed using the cost parameters obtained from the prior cost analysis of all separately organized ECG, EEG and nuclear medicine departments.

The costs of providing hospital services are computed as reported operating costs of the hospital less two major and several minor cost categories. The first major cost category, that of the office of the medical staff, is excluded since the diagnostic and prescriptive activities of service demand generation by attending physicians are thought to be largely complementary to service production in producing the illness remission of case-types.[5] The second major cost category, that of education, is excluded since the in-hospital training of medical interns and residents, nursing staff and technicians is only weakly related to service production. However, recognizing the possibility of substantial indirect costs of teaching programmes, a teaching activity variable is included as an explanatory variable in subsequent cost analysis. Several minor cost categories are excluded since they are unrelated (in their approximately self-financing form) to service production: coffee shop, parking and ambulance service. Other minor cost categories are excluded in order to avoid factor price variation: for example, interest costs are excluded since they are small and highly sensitive to considerations of external (charitable) financing; and anaesthetic gas costs are excluded as anaesthetic gases (largely complementary to other inputs) are often provided by private-practice anaesthesiologists rather than the hospital (as is suggested by the observation that a significant number of hospitals reported no expenditures for anaesthetic gases). One additional minor cost is excluded as a result of anticipated problems in estimation; since only 20 hospitals report output and costs in the provision of 23 different types of organized out-patient clinics, and since related costs are well less than one percent of all hospital costs, then the service cost analysis excludes the costs and output of the organized out-patient department.

Case-type costs are incurred as a result of the combined activities of service production (by hospitals) and service demand generation (by attending physicians). However, the reported values of hospital costs in the sample do not include (even a fraction of) the remuneration of attending physicians in 9 of 10 cases; in addition, estimates of excluded remuneration are unavailable. Therefore the hospital component of case-type costs, which is defined to exclude the remuneration of attending physicians, becomes the object of the case-type cost analysis.

The hospital component of case-type costs is computed as the reported operating costs of hospitals less the reported remuneration of attending physicians (found under the heading of the office of the medical staff), education costs and other minor cost categories similarly excluded in the service cost analysis. Although the first exclusion aggravates the downward bias in case-type cost coefficients, it minimizes or even eliminates the bias in inter-hospital comparisons of standardized costs, since hospitals reporting expenditures for the office of the medical staff will no longer appear more costly. Other exclusions have already been defended.

Since the case-type classification extends to in-patients only, the case-type cost analysis is further restricted to an examination of in-patient costs. In-patient costs are computed as case-type costs less an estimate of out-patient costs. Out-patient costs are alternatively computed as reported out-patient income or as an estimate of out-patient costs which is derived as the sum of the products of reported out-patient services and the out-patient service cost parameters estimated in a prior service cost analysis. The former estimate of out-patient costs averages one percent less than the latter; they both have per admission coefficients of variation of 33 percent; and they are significantly correlated in the per admission form (ρ = .79).

In-patient case-type costs are further adjusted to exclude an estimate of the costs of long-term chronic care. This is done for two reasons. First, the case-type classification includes short-term in-patients only. Second, no obvious correction for long-term chronic care seems appropriate, since the per diem services received by chronic care patients are neither identifiable nor homogeneous (in view of the variation in approved per diem charges of $20 to $50). Long-term chronic care costs are computed as the reported value of long-term chronic care income.

Exclusions from both service and case-type cost variables do limit the scope of the cost analysis. However, they may nonetheless permit a relatively <u>unbiased</u> examination of the <u>defined</u> level of costs which is ultimately more worthwhile than a biased examination of <u>observed</u> hospital costs. Moreover, they do not prevent a separate examination of the excluded cost categories in the event of subsequent data improvements.

d. Explanatory Variables

Explanatory variables are also specified for departmental, service and case-type cost equations. They are classified as either output or characteristic variables. Output variables capture the variable costs associated with services or case-types, the minimum fixed costs associated with any factor indivisibilities, the quasi-fixed costs associated with any short-term fixed capacity, and the economies of scale associated with any returns to scale. Characteristic variables capture cost differences arising from the spillover effects of teaching programmes, quality differences associated with medical affiliation or rural location, and any factor price differences associated with religious affiliation.

Reported services for individual departments generally permit detailed and usually obvious corrections for departmental output. For example, reported output for the following departments is differentiated by in-patient and out-patient classification: physical medicine and rehabilitation, which provides weighted units of physiotherapy, paid hours of occupational therapy and visits to speech therapy; the laboratory, which provides standard units of seven types of laboratory tests as well as autopsies; and the ward service department, which provides seven types of patient-day service, admission and discharge, and "trips off-ward" (computed as the number of round trips patients make to the radiology, EEG, physical medicine and rehabilitation, and other special service departments).

Capacity variables are available for a minority of departments. The capacity of the newborn nursery is reported as bassinets set up, that of the operating room as the number of operating rooms, and the capacity of ward services as rated patient-days (rated beds x 365), with additional corrections possible for rated obstetrical, intensive care and other patient-days. Proxies for capacity appear to take the form of rated patient-days in the case of physical medicine and rehabilitation and combined operating and delivery room and emergency ward (the rated patient-days variable is marginally significant and is responsible for about 10 percent of departmental costs). No capacity variables (or good proxies) are available for laboratory, radiology, ECG, EEG and nuclear medicine.

Characteristic variables prove to be generally unimportant at the departmental level with the single exception of medical teaching activity (which, as discussed below, probably proxies for unremoved product heterogeneity). The insignificance of rural location is not surprising in view of the exclusion of land from departmental costs. The insignificance of religious affiliation is also not surprising in view of the equal reimbursement of hospitals for lay and non-lay personnel. The insignificance of training activity in the laboratory and radiology is consistent with the view that students acquiring general training (for which skills are readily adaptable to other employers) pay for their training by accepting wages less than their productivity.[6]

Explanatory variables in the service cost equation initially include all those found to be statistically significant in the departmental cost equations. In order to reduce multicollinearity using the approach outlined in Section 3 of the previous chapter, variables for each of four (groups of related) departments (newborn nursery; laboratory and ECG; radiology, EEG and nuclear medicine; and physical medicine and rehabilitation) appear in aggregated form. They are combined into output indices, with weights equal to their marginal costs in relation to that of a departmental numeraire (a newborn patient-day, standard laboratory unit, outpatient X-ray and weighted physiotherapy unit respectively).

For example, the estimated form of the physical medicine and rehabilitation cost equation is given by equation (1),

$$(1) \quad tc_p = a_1PU + a_2OH + a_3SV + a_4R + \text{error}$$

where tc_p are the reported costs of physical medicine and rehabilitation, PU is weighted units of physiotherapy, OH is paid hours of occupational therapy, SV is visits to speech therapy, and R is rated patient-days (a proxy for capacity).

Equation (1) can be rewritten as follows:

$$(1)' \quad tc_p = a_1\left(PU + \frac{a_2}{a_1} OH + \frac{a_3}{a_1} SV + \frac{a_4}{a_1} R\right) + \text{error}$$

An output index for physical medicine and rahabilitation can therefore be computed as follows:

$$(2) \quad OI_p = PU + \frac{a_2}{a_1} OH + \frac{a_3}{a_1} SV + \frac{a_4}{a_1} R$$

Explanatory variables found to be statistically <u>insignificant</u> in departmental cost equations are introduced into the service cost equation in a stepwise manner. All except private and semi-private patient-days of preferred accommodation retain their insignificance and are therefore excluded from the final form of the service cost equation. In addition, variables representing numbers of autopsies, ECG's, EEG's, nuclear medicine visits, paid hours of occupational therapy and speech therapy visits are speculatively introduced into the service cost equation; their insignificance confirms the validity of the four output indices (a similar procedure warned against the use of output indices for ward service, operating and delivery room and emergency ward).

The explanatory variables included in the case-type cost equation are a total short-term admissions variable (defined residually as having a normal delivery diagnosis), the numbers of admissions in each of 40 non-standard case-type categories (see Table 5), long-term convalescent patient-days, private and semi-private patient-days, and several structural and characteristic variables. A typical non-standard case-type has unit (stay-, admission- and capacity-related) costs equal to the algebraic sum of the cost coefficient of the total short-term admissions variable and that of the relevant non-standard case-type output variable.

The long-term convalescent patient-days variable captures the treatment costs of patients other than short-term admissions (represented by the total admissions and case-type output variables) and long-term chronic care admissions (for which costs are excluded). Structural variables include a correction for capacity (unused rated patient-days, computed as rated less actual patient-days) and for returns to scale (squared rated patient-days). Characteristic variables are several: a correction for the indirect costs of medical teaching programmes (teaching expense); a correction for illness severity (rural rated patient-days, where hospitals are rural if located in a municipality of less than 20,000); and a further correction for illness severity (excess patient-days, corrected for case-mix). In symbols, the excess patient-days variable is computed as follows for the jth hospital.

$$(3)\quad XPD_j = \sum_{i=1}^{41} (P_{ij} - C_{ij} \sum_{k=1}^{101} P_{ik} / \sum_{k=1}^{101} C_{ik})$$

where XPD_j are the excess patient-days of hospital j, P_{ij} and C_{ij} are the total patient-days and patients respectively of case-type i in hospital j, and P_{ik} and C_{ik} are the total patient-days and patients respectively of case-type i in hospital k.

As is indicated by their insignificance in the per diem cost equation of Evans (1971), age-sex corrections are though to be unnecessary. Not only are they ambiguous in the aggregated form and overly numerous in the disaggregated form, but they are largely redundant as long as a correction is made in the cost equation for length of stay (excess patient-days).

Alternative corrections are available for both capacity and the illness severity reflected in differential length of stay-rated patient-days and actual patient-days respectively. However, the use of these alternative corrections does not permit case-type cost coefficients to reflect capacity- and stay-related costs as well as admission-related costs.

e. Regression Technique

The straightforward application of ordinary least squares (OLS) in estimating total cost equations generally results in inefficient and imprecise parameter estimates in hospital cost analysis. Inefficient parameter estimates occur as a result of heteroscedasticity - a diagonal but not scalar-diagonal matrix of residual variances. Imprecise parameter estimates occur as a result of multicollinearity among the explanatory variables.

Heteroscedasticity in hospital cost analysis is suspected for two reasons. First, intuition suggests that large hospitals have large total costs and therefore tend to have large unexplained cost components; for example, the cost impact of a strike by non-medical workers against a subset of hospitals will be greater for large than small hospitals. Second, statistical tests such as the Goldfield and Quandt or Glejser test[7] (using OLS or Theil BLUS residuals) suggest that the absolute magnitude of total cost residuals are directly related to hospital size.

The hospital cost analyst typically responds to heteroscedasticity by regressing the cost equation in an average form. For example, he specifies the radiology cost equation in a per X-ray form; laboratory costs in a per standard laboratory unit form; and service or case-type cost equations in a per diem, per rated patient-day or per admission form. By regressing average cost equations he is using weighted least squares (WLS). He is correct in doing this if the residuals of the average cost equation are homoscedastic. An iterative logarithmic Glejser test of the residuals (both OLS and BLUS) of average cost equations does confirm that the adoption of average cost forms is an appropriate response to heteroscedasticity. Hence all departmental, service and case-type equations are estimated in an average cost form.

Multicollinearity is evident in virtually all cost equations. It is especially severe if a small minority of the k principal components captures a large fraction of the variation among the k explanatory variables; alternatively, it is especially severe if the sum of a small minority of k eigenvalues approaches k. It is most severe in the case-type cost equation, for which as few as 16 of all 48 eigenvalues sum to 36. It is least severe in the departmental cost equations for which sample sizes are large in relation to the number of explanatory variables. It is of intermediate severity in the service cost equation for which the number of explanatory variables (18) is kept to a minimum through the judicious use of the output indices and exclusions implied by the departmental cost equations.

In addition to any aggregations and exclusions justified on a theoretical or empirical basis, the principal components technique is employed as an intermediate step in cost regression in order to deal with multicollinearity. All explanatory (independent) variables are represented by the corresponding set of mutually orthogonal principal components. The cost regression employs as regressors only as many principal components as are statistically significant[9] on the basis of an F-test of the incremental R^2 associated with groups of additional principal components introduced into the regression in descending order of eigenvalues. The "unscrambling" procedure (linear transformation using the relevant subset of eigenvectors of the correlation matrix of the explanatory variables) described in the previous chapter is then applied to the estimated coefficients and variances of the principal component regressors. The results are estimates of the coefficients of the explanatory variables which are more precise (have smaller mean square errors) than the straightforward OLS estimates. In addition, standard errors and T-statistics can also be computed. See the previous chapter, especially subsection 2.g., section 3, and footnote 11 for further explanation.

Finally, stepwise regression is used, first backward then forward, to evaluate explanatory variables for which the theoretical case for inclusion is weak. Backward stepwise regression is used extensively in departmental cost equations (for which estimation is reasonably precise). In particular, a number of initially specified output corrections, for example, obstetric, pediatric and psychiatric patient-days in the ward service cost equation, are eliminated as a result of their statistical insignificance. Forward stepwise regression is used extensively in the service cost equation in order to evaluate output indices, output variables initially excluded on the basis of insignificance in the departmental cost analysis (for example, private and semi-private patient-days of ward care) and characteristic variables (for example, medical teaching expense and rural rated patient-days). It is also used, but less extensively, in the case-type cost equation in order to evaluate characteristic variables such as excess patient-days and medical teaching expense and structural variables such as unused rated patient-days, squared rated patient-days and a constant term for minimum fixed costs.

3. Regression Results

a. General Interpretation of Tables

All cost equations are estimated in an average cost form (with the divisor of total costs and all independent variables (IV) being the regression weight, RW) and most employ the principal components technique as an intermediate stage in estimation. However, the independent variables and corresponding cost coefficients (CC) presented in Tables 2A-2I, 3 and 4 (see also Table 1 for a key to symbols and Table 5 for case-type definitions) can be interpreted as those of total cost equations. For example an estimate of the total costs (TC_j) of ward services in the jth hospital is obtained from Table 2A as the sum of the products of CC and IV.

$$(4) \quad TC_j \doteq \sum_{i=1}^{6} CC_i IV_{ij}$$

or \$10.21 times total patient-days, plus \$1.87 times rated patient-days, minus \$4.43 times long-term patient-days, and so on.

Important properties of the explanatory variables are also presented in the tables. The coefficients of variation (standard deviation divided by the mean) among sample hospitals of the weighted independent variables (IV divided by RW) are given by the values of WCV. The levels of statistical significance of the independent variables can be determined through an investigation of their T-statistics, TS. The percentage of the total costs

TABLE 1 - HOSPITAL COST ANALYSIS: KEY TO SYMBOLS

Symbol	Definition
RW	Regression weight: total cost and all independent variables are divided by RW prior to estimation
IV	Independent variable in the total cost equation
WCV	Weighted coefficient of variation (%) = $\frac{\text{Standard deviation}}{\text{mean}}$ of $\frac{\text{variable}}{\text{RW}}$
CC	Estimated cost coefficient ($)
TS	T-statistic of CC
PAC	Percent of all costs ascribed to variable
WCV_d	Coefficient of variation of average costs
N	Number of observations in the sample
NV	Number of independent variables in the total cost equation
NPC	Number of principal component variables used in the regression, plus one
PAV	Percent of variation in the weighted independent variables captured by NPC-1 principal component variables
$\bar{R}^2$	Coefficient of multiple correlation, corrected for degrees of freedom, of the regression (in weighted form)
PD	Patient-days
i/p	In-patient
o/p	Out-patient
OI	Output Index
NB	Newborn
OR	Operating Room

TABLE 2A - WARD SERVICE COST COEFFICIENTS

RW	IV	WCV	CC	TS	PAC
Total PD	Total PD		10.21	8.6	64
WCV_d = 11	Rated PD	10	1.87	2.2	14
N = 98	Long-term PD	171	-4.43	-5.6	-2
NV = 6	Intensive Care PD	83	19.04	2.1	2
NPC = 5	Admissions	20	7.05	2.5	4
PAV = 96	Trips off-ward	31	3.70	2.4	7
$\bar{R}^2$ = .55	(Rated PD)^2	68	.00000768	6.0	10

TABLE 2B - NEWBORN NURSERY COST COEFFICIENTS

RW	IV	WCV	CC	TS	PAC
Total NBPD	Total NBPD		10.18	12.0	73
WCV_d = 23	Bassinets Set up	37	1.48	4.5	22
N = 94	Medical Teaching Costs	273	.021	3.0	4
NV = 3					
NPC = 3					
PAV = 100					
$\bar{R}^2$ = .20					

TABLE 2C - OPERATING AND DELIVERY ROOM AND EMERGENCY WARD COST COEFFICIENTS

RW	IV	WCV	CC	TS	PAC
Rated PD	i/p surgical procedures	36	26.71	6.5	33
WCV_d = 26	o/p operations in the OR	130	19.40	3.7	5
N = 94	Number of OR (main suites)	29	7050.73	2.1	10
NV = 8	Deliveries	50	46.28	3.4	12
NPC = 7	Emergency Visits	45	3.75	6.5	22
PAV = 97	Rated PD		.250	1.0	7
$\bar{R}^2$ = .71	$(\text{Rated PD})^2$	68	.00000148	4.2	8
	Medical teaching costs	230	.055	2.5	3

TABLE 2D - PHYSICAL MEDICINE AND REHABILITATION COST COEFFICIENTS

RW	IV	WCV	CC	TS	PAC
Rated PD	Weighted Physiotherapy Units	35	.138	6.8	67
WCV_d = 48	Paid hours of occupational therapy	136	3.28	3.4	12
N = 92	Visits to speech therapy	261	7.04	2.1	5
NV = 4	Rated PD		.083	1.7	15
NPC = 4					
PAV = 100					
$\bar{R}^2$ = .56					

TABLE 2E - LABORATORY COST COEFFICIENTS

RW	IV	WCV	CC	TS	PAC
Total Standard Units	Total Standard Units		.071	4.8	63
WCV_d = 18	Haematology	26	.060	1.5	10
N = 93	Biochemistry	26	.047	1.8	12
NV = 7	Histopathology	44	.077	1.5	6
NPC = 7	Cytopathology	123	.257	2.3	4
PAV = 100	Automated Units	282	-.010	-1.2	-1
$\bar{R}^2$ = .16	Autopsies	66	145.50	1.9	5

TABLE 2F - RADIOLOGY COST COEFFICIENTS

RW	IV	WCV	CC	TS	PAC
Total X-rays	Total X-rays		6.66	18.7	73
WCV_d = 16	i/p X-rays	34	3.87	4.3	12
N = 93	Miniature T.B. Screening X-rays	117	-4.22	-3.2	-3
NV = 7	Fluoroscopic X-rays	38	7.33	4.4	8
NPC = 5	Cineradiographic X-rays	37	36.37	3.6	2
PAV = 85	$(\text{Total X-rays})^2$	67	.00000771	3.1	4
$\bar{R}^2$ = .47	Medical Teaching Costs	238	.064	7.9	4

TABLE 2G - ECG COST COEFFICIENTS

RW	IV	WCV	CC	TS	PAC
Total Standard Units	Total Standard Units		.247	15.6	110
WCV_d = 21	o/p Units	54	.167	2.5	14
N = 39	$(\text{Units})^2$	64	-.000000173	-3.7	-24
N = 3					
NPC = 3					
PAV = 100					
$\bar{R}^2$ = .32					

TABLE 2H - EEG COST COEFFICIENTS

RW	IV	WCV	CC	TS	PAC
Total Standard Units	Total Standard Units		.153	8.0	81
WCV_d = 28	o/p Units	48	.083	2.1	19
N = 28					
NV = 2					
NPC = 2					
PAV = 100					
$\bar{R}^2$ = .14					

TABLE 21 - NUCLEAR MEDICINE COST COEFFICIENTS

RW	IV	WCV	CC	TS	PAC
Total Visits	Total Visits		21.15	15.7	128
$WCV_d = 64$	Unity	279	5702.23	11.3	7
N = 25	$(\text{Total Visits})^2$	87	-.000837	-3.9	-45
NV = 4	Medical Teaching Costs	176	.017	2.7	10
NPC = 4					
PAV = 100					
$\bar{R}^2 = .90$					

TABLE 3 - SERVICE COST COEFFICIENTS, ONTARIO, 1971

RW	IV	WCV	CC	TS	PAC
Total PD	Total PD		16.07	4.2	26
WCV_d = 16	Rated PD	10	9.20	3.8	18
N = 101	Long-term PD	167	-8.73	-5.0	-1
NV = 19	Intensive Care PD	85	48.89	2.2	2
NPC = 10	NB Nursery OI	45	11.87	2.5	2
PAV = 87	Semi-Private PD	28	6.67	4.0	5
$\bar{R}^2$ = .82	Private PD	61	10.87	3.0	2
	Admissions	20	21.25	2.6	3
	Trips off-ward	31	7.55	3.0	4
	i/p surgical procedures	33	34.90	2.4	3
	o/p operations in the OR	127	36.64	1.3	1
	Number of OR (main suites)	32	36663.33	3.6	4
	Deliveries	49	46.49	1.2	1
	Physical Medicine and Rehabilitation OI	39	.279	1.9	2
	Laboratory and ECG OI	32	.083	13.3	9
	Radiology, EEG and Nuclear Medicine OI	32	9.96	9.1	9
	Emergency Visits	46	4.20	1.9	2
	$(\text{Rated PD})^2$	67	.0000166	5.8	6
	Medical Teaching Costs	239	.780	8.0	2

TABLE 4 - CASE-TYPE COST COEFFICIENTS, IN-PATIENTS, ONTARIO, 1971*

RW	IV	WCV	CC[a]	TS[a]	PAC[a]	CC[b]	TS[b]	PAC[b]
Short-term Admissions	Short-term Admissions		290.23	7.5	57	287.91	7.4	56
	(Rated - Total) PD	80	17.29	9.1	6	17.33	9.1	6
WCV_d = 33	Excess PD	508	24.74	10.3	-	25.64	10.7	-
N = 101	Long-term Convalescent PD	973	14.09	8.7	1	14.50	8.9	-1
NV = 49	Semi-private PD	29	12.60	4.0	10	13.90	4.4	11
NPC = 25	Private PD	65	15.04	3.3	3	13.36	2.9	3
PAV = 91	Rated PD if Rural**	140	-.835	-2.0	-.3	-.969	-2.2	-.3
$\bar{R}^2$ = .92	$(\text{Rated PD})^2$	119	.0000147	16.6	6	.0000152	17.2	6
	Medical Teaching Expense	260	.584	10.6	2	.603	10.9	2
	1 - INFPAR	39	-46.70	-.2	-.1	-120.68	-.4	-.5
	2 - NEOMAL	42	1026.75	7.3	6	1044.06	7.4	6
	3 - NEOBEN	37	481.73	1.7	2	494.41	1.7	2
	4 - ASTHMA	46	-119.75	-.2	-.1	-401.13	-.3	-.5
	5 - DIAMET	25	441.68	.7	2	342.59	.5	1
	6 - BLOOD	32	648.01	.4	1	626.01	.4	1
	7 - PSYCHO	77	1143.47	2.5	2	1323.35	2.0	2
	8 - NEUROT	67	-23.75	-.2	-.1	48.23	.3	.2
	9 - VLNERV	42	223.82	.5	.5	138.60	.3	.3
	10 - OTHCNS	51	3112.47	5.9	5	3197.32	6.1	5
	11 - NERVES	120	353.59	.5	.2	161.06	.2	.1

(continued)

TABLE 4 (continued)

IV	WCV	CC^a	TS^a	PAC^a	CC^b	TS^b	PAC^b
12 - EYEAR	60	74.72	.3	.4	62.15	.2	.3
13 - HEART	28	-293.90	-1.7	-3	-266.71	-1.5	-3
14 - CIRCUL	38	797.47	3.3	4	771.66	3.2	4
15 - UPRESP	53	-183.08	-1.1	-.7	-224.57	-1.3	-.8
16 - PNEUMO	63	809.58	3.4	3	805.57	3.4	3
17 - BRONCH	49	529.02	1.2	1	473.17	.9	1
18 - TONSAD	42	-417.12	-2.3	-4	-457.78	-2.5	-4
19 - OTHRES	53	289.64	.6	1	445.69	.8	1
20 - TEETH	111	-904.33	-5.0	-2	-985.11	-5.4	-2
21 - UPPRGI	33	-693.80	-1.3	-2	-532.66	-1.0	-2
22 - APPEND	45	532.98	.8	1	819.85	1.2	2
23 - HERNIA	33	-233.25	-.6	-1	-313.56	-.7	-1
24 - OTHGIS	20	78.25	.3	1	174.28	.5	2
25 - URINAR	77	-19.81	-.2	-.2	.99	.1	.0
26 - MGENIT	39	1296.76	1.7	3	1356.93	1.8	3
27 - FGENIT	34	-206.72	-1.1	-2	-259.92	-1.4	-2
28 - PREGCY	32	-753.92	-1.4	-3	-727.71	-1.3	-3
29 - ABNDEL	57	339.50	2.7	3	346.75	2.7	3
30 - SKIN	36	2257.15	4.6	6	2230.18	4.5	6
31 - BONEJT	43	308.41	2.2	2	391.02	2.9	3
32 - MALFOR	54	-198.33	-.3	-.4	-64.45	-.1	-.1
33 - INFANT	113	109.56	.2	.2	367.02	.7	.5
34 - SYMPTM	47	207.77	.8	1	230.66	.8	1
35 - FRACT	32	303.60	1.4	2	316.38	1.4	2
36 - DISLOC	40	-1949.84	-1.6	-2	-1824.73	-1.5	.2
37 - INTINJ	40	815.06	1.1	2	535.16	.7	1
38 - EXTWND	42	-1356.28	-4.0	-4	-1463.66	-4.3	-4
39 - POISON	26	-1119.40	-1.7	-4	-1301.40	-2.0	-4
40 - NBMISC	33	-116.23	-2.1	-3	-141.95	-2.5	-3

* i/p costs = total costs - x. x^a = estimated o/p income. x^b = o/p income.
**Hospitals are rural if municipal population is less than 20,000.

TABLE 5 - CASE-TYPE OR DIAGNOSTIC CATEGORIES

Diagnostic Category		Includes	ICDA-8 Codes
1	INFPAR	Infective and parasitic diseases	000-136
2	NEOMAL	Malignant neoplasms	140-209
3	NEOBEN	Benign neoplasms and neoplasms of unspecified nature	210-239
4	ASTHMA	Asthma	493
5	DIAMET	Endocrine, nutritional and metabolic diseases	240-279
6	BLOOD	Diseases of the blood and blood forming organs	280-289
7	PSYCHO	Psychoses	290-299
8	NEUROT	Other mental disorders	300-315
9	VLNERV	Cerebrovascular diseases	430-438
10	OTHCNS	Inflammatory, hereditary, familial and other diseases of the central nervous system	320-349
11	NERVES	Diseases of nerves and peripheral ganglia	350-358
12	EYEAR	Diseases of the eye, ear, and mastoid process	360-389
13	HEART	Heart diseases	390-429
14	CIRCUL	Other diseases of circulatory system	440-458
15	UPRESP	Acute respiratory infections and influenza	460-474
16	PNEUMO	Pneumonia	480-486
17	BRONCH	Bronchitis and emphysema	490-492
18	TONSAD	Hypertrophy of tonsils and adenoids	500
19	OTHRES	Other diseases of respiratory system	501-519
20	TEETH	Diseases of the oral cavity, salivary glands and jaws	520-529

(Continued)

TABLE 5 (continued)

Diagnostic Category		Includes	ICDA-8 Codes
21	UPPRGI	Diseases of esophagus, stomach, and duodenum	530-537
22	APPEND	Appendicitis	540-543
23	HERNIA	Hernia of abdominal cavity	550-553
24	OTHGIS	Diseases of intestine, peritoneum, liver, gallbladder and pancreas	560-577
25	URINAR	Nephritis, nephrosis and other diseases of the urinary system	580-599
26	MGENIT	Diseases of male genital organs	600-607
27	FGENIT	Diseases of female genital organs	610-629
28	PREGCY	Complications of pregnancy	630-639
29	ABNDEL	Abortion, complications of delivery and of the puerperium	640-645, 651-678
30	SKIN	Diseases of the skin and subcutaneous tissue	680-709
31	BONEJT	Diseases of the musculoskeletal system and connective tissue	710-738
32	MALFOR	Congenital anomalies	740-759
33	INFANT	Certain causes of perinatal morbidity and mortality	760-779
34	SYMPTM	Symptoms, senility and ill-defined conditions	780-796
35	FRACT	Fractures	800-829
36	DISLOC	Dislocation without fracture and sprains	830-848
37	INTINJ	Internal Injuries	850-869
38	EXTWND	Lacerations, contusions, superficial injuries, effects of foreign bodies, burns, and injuries to nerves and spinal cord	870-959
39	POISON	Adverse effect of medicinal agents, toxic effect of non-medicinal substances and other adverse effects	960-999
40	NBMISC	Mature and immature newborn and special conditions and examinations without sickness	Y00-Y29
41	NORDEL	Normal delivery (Standard)	650

*Source: The Relative Stay Index Report, 1971. Toronto: QHSC, 1971, p. 11.

explained in each table by individual independent variables is given by PAC. For the ith independent variable,

$$(5) \quad PAC_i = \left(\sum_{j=1}^{N} CC_i IV_{ij} / \sum_{j=1}^{N} TC_j \right) \times 100\%$$

Large values of WCV and PAC for an independent variable suggest its importance in explaining the inter-hospital variation in total costs.

Other statistical properties are given at the left-hand side of each table. The coefficient of variation among sample hospitals of the weighted (average) cost variable (total costs divided by RW) is given by WCV_d. The sample size, N, is always less than the maximum of 101 for individual departments as a result of the non-existence of separately organized or reporting departments in some hospitals. The total number of independent variables employed in the total cost equation is NV. The number of principal components (having the smallest eigenvalues) excluded from the average cost regression is NV-NPC. The percentage of the variation in the weighted independent variables captured by the included principal components is PAV. Finally, the coefficient of multiple determination, corrected for degrees of freedom, $\bar{R}^2$, gives an indication of the goodness of fit of the average cost equation. A good fit for the equivalent total cost equation is indicated by a large value of the $\bar{R}^2$ in relation to WCV_d.

b. Departmental Cost Analysis

Most departmental cost equations are estimated using an obvious regression weight. For example, ward service uses total patient-days, laboratory uses standard laboratory units and radiology uses total X-rays. However, "combined" departments such as physical medicine and rehabilitation use rated patient-days, since no obvious basis can be established for choosing among the outputs of physiotherapy, occupational therapy and speech therapy, and rated patient-days does achieve statistical significance as a proxy for capacity. All cost equations seem to be efficiently estimated since a Glejser test of the average cost residuals indicates their homoscedasticity.

Departmental cost equations are estimated in a manner which excludes few (if any) principal components. This is not surprising in view of the fewness of explanatory variables in relation to sample size. The radiology cost equation is estimated with the most spectacular use of the principal components technique since two principal components are excluded on the basis of statistical insignificance at the 10 and 40 percent level.

All departmental cost coefficients are statistically significant (using a one-tailed test) at the 2.5 percent level with the general exception of the laboratory cost coefficients and the specific exception of those of the rated patient-days variables in Tables 2C and 2D. Laboratory cost coefficients are significant at the 5 to 10 percent level, while the coefficients of rated patient-days, a proxy for capacity, are significant at the 5 to 15 percent level.

Quasi-fixed or capacity costs are indicated for a department whenever an obvious or proxy capacity measure has a positive and statistically significant coefficient. Capacity costs are indicated by the rated patient-days variable for ward service, physical medicine and rehabilitation, and combined operating and delivery room and emergency ward; by bassinets set up for newborn nursery; and by the number of operating rooms for operating room services. Capacity costs form a significant part of all departmental costs: 14 percent for ward service, 15 percent for physical medicine and rehabilitation, a total of 37 percent for combined operating and delivery room and emergency ward, and 22 percent for newborn nursery.

Minimum fixed costs (reflecting factor indivisibilities) are apparent in only one department, that of nuclear medicine. Costs of approximately $5,700 are incurred by a hospital in providing the basic capability for nuclear medicine services. They are about 25 percent of all nuclear medicine costs in typical small output hospitals and 2.5 percent in typical large output hospitals.

A significant coefficient for a squared term in departmental output or rated patient-days can be viewed as indicating the absence of constant returns to scale. However, if the adequacy of departmental output corrections for product mix is suspect, it may also suggest unremoved product heterogeneity.

The negative cost coefficients for squared units of ECG and nuclear medicine visits are thought to indicate economies of scale. The reasons for this are threefold. First, output is thought to be highly standardized in these departments, especially ECG. Second, any residual product heterogeneity would tend to produce positive coefficients for squared output, since the consensus of opinion is that output complexity increases with departmental output. Third, increasing returns to scale due to specialization are quite plausible in highly technical production processes. The standardized per unit costs of ECG and nuclear medicine are approximately 40 percent less in typical large output than in typical small output hospitals.

The positive coefficient of squared rated patient-days for ward services is thought to reflect diseconomies of scale. The reasons for this assertion are twofold. First, departmental output is thought to be adequately standardized, since additional corrections (surgical, obstetric, psychiatric and pediatric or narrowly-defined diagnoses per Table 5) for patient-day output neither are statistically significant nor significantly affect the coefficients of Table 2A. Second, diseconomies of scale in ward service are plausible to the extent that inter-departmental coordination and transportation costs (between wards and special service departments) are disproportionately higher in large (1,000-bed) hospitals than in small (100-bed) hospitals.[11] Standardized per diem costs for ward services are 15 percent higher in large than in small hospitals.

The positive coefficient for the squared terms in X-rays (Table 2F) and rated patient-days (Table 2C) are thought to indicate unremoved product heterogeneity. The reasons for this are twofold. First, unremoved product heterogeneity is plausible for both radiology and combined operating and delivery room and emergency ward in view of the possible inadequacy of corrections for output mix. Second, the consensus is that output complexity and departmental output are positively correlated. Standardized unit costs for radiology and combined operating and delivery room and emergency ward are respectively 6 and 12 percent higher in typical large than in typical small output hospitals.

A positive and significant coefficient for medical teaching expense might be the result of the indirect costs of teaching programmes or unremoved product heterogeneity. It is thought that the former cause is improbable at the departmental level for two reasons. First, the indirect costs of teaching programmes primarily take the form of overhead costs which are excluded from reported departmental costs. Second, the teaching expense variable is not significant for two of the largest hospital departments (ward service and laboratory) for which indirect teaching costs should be at least as important as for other departments. Hence, as a result of output complexity, standardized unit costs in teaching hospitals are 20 percent higher for newborn nursery, 10 percent higher for radiology, 25 percent higher for nuclear medicine, and 10 percent higher for combined operating and delivery room and emergency ward.

The unit costs of services differ between in-patients and out-patients for several departments. The costs of providing the facilities and nursing staff of the operating room are about 35 percent higher for an in-patient than an out-patient operation. The cost of an X-ray are about 50 percent higher for an in-patient than an out-patient. The costs of standard units of ECG and EEG are respectively 70 and 50 percent higher for out-patients than in-patients. This last finding contrasts with the first two. It is especially surprising in view of the highly standardized output measures for ECG and EEG which are implied by the "standard unit" system. A possible explanation might be that the considerable involvement of the ECG and EEG departments in out-patient activities imposes extra scheduling and administrative costs at the <u>departmental</u> level.

The results of the departmental cost analysis are not readily comparable with those of Ingbar and Taylor (1968) and Cohen (1973), since Ingbar and Taylor explain the costs of individual departments by using the same set of service variables used in the explanation of all hospital costs, and Cohen fails to make very detailed (homogeneous) output corrections. However, the results of this cost analysis are superior in at least one respect - whenever dependent variables are comparable, cost equations have greater explanatory power. For example, the $\bar{R}^2$ of the radiology cost equation is .47 while that found by Cohen is only .04, and the $\bar{R}^2$ of the ward service cost equation is .55 while that found by Ingbar and Taylor is only .46. In addition, the significance levels of estimated cost coefficients are more acceptable which suggests that parameter estimates are more precise.

It is interesting to compare the results of the departmental cost analysis with those of the service cost analysis presented next. Results for newborn nursery; laboratory and ECG; radiology, EEG and nuclear medicine; and physical medicine and rehabilitation are necessarily similar as a result of the (justifiable) use of the output indices in the service cost analysis. Results for other services are also highly comparable with one exception-. preferred accommodation has an impact on service costs but none on reported ward service costs. Since reported service costs include housekeeping and depreciation costs, whereas reported ward service costs do not, this finding is easily reconciled.

c. <u>Service Cost Analysis</u>

The results of the service cost analysis are presented in Table 3. Total patient-days is the regression weight, implying that the service cost equation is estimated in per diem form. Although rated patient-days and admissions also appear to be suitable as regression weights (in so far as heteroscedasticity is avoided), their use seems to result in less precise estimation (assuming an intermediate principal components stage of estimation). In particular, the per rated patient-day and per admission variation in costs and in the independent variables is such that 13 and 16 principal components must be included in regressions (whereas only 9 are needed in the per diem cost equation) if they are to adequately explain hospital costs.

The use of the principal components technique seems to be especially fruitful in the service cost regression. The nine principal components included in the regression account for 87 percent[12] of the per diem variation among the independent variables. They permit an $\bar{R}^2$ which is as large as that achieved by including all 18 principal components (or independent variables) in the regression. Not surprisingly, F-tests of the incremental R^2 captured by the excluded principal components (introduced by groups in descending order of eigenvalue) confirmed the insignificance of excluded principal components at the 5 percent level. The effects of including all principal components (which is equivalent to using straightforward OLS without the intermediate principal components stage) are twofold: first, to reduce the T-statistics (and precision) of estimated coefficients; and second, to result in little change in estimated parameters other than offsetting changes in the parameters of two sets of highly collinear independent variables - deliveries and newborn patient-days, and admissions and rated patient-days.

Estimated coefficients are quite robust. They fail to change by more than a small fraction of their standard errors of estimate in any of the following cases: when the regression weight is rated patient-days or admissions; when the service cost variable is modified to include one or more of the initially excluded cost categories such as interest expense; when the service cost variable and independent variables are modified to encompass in-patient services only; and when the sample excludes teaching hospitals.

The major sources of inter-hospital variation in per diem costs can be determined by an examination of the coefficients, significance levels and factor loadings of the principal components corresponding to the largest eigenvalues of the correlation matrix of the per diem values of the independent variables. The first three principal components explain 80 percent of the variation in per diem costs and capture 25, 20 and 10 percent respectively of the per diem variation among the independent variables. Their coefficients are relatively large and their T-statistics are impressive - 20.1, 5.3 and 4.4 respectively. Their factor loadings (correlations with independent variables) suggest that hospitals have per diem costs above the mean ($57.90) if they provide above-average amounts of radiology, laboratory, surgery or obstetric care, if they are large or involved in medical teaching, if they have below-average occupancy rates, or if they are minimally involved in long-term chronic and convalescent care.

In general, the signs and magnitudes of the service cost coefficients are readily defensible and their significance levels quite reasonable. In particular, the coefficients of all resource-consuming activities are positive (including the <u>net</u> cost coefficient, $16.07 + $9.20 - $8.73, of long-term patient-days). The magnitudes of the cost coefficients of all output variables, including the output indices, are moderately larger than those found in the departmental cost analysis, which simply reflects the incidence of indirect or overhead costs among services. Finally, all cost coefficients are significant at the 5 percent level, except those for deliveries and out-patient operations which are significant at the 10 percent level.

The results of the service and departmental cost equations are highly comparable with the exception of the large and statistically significant coefficients for private and semi-private patient-days and medical teaching expense, which again simply reflect the inclusion of overhead costs in the service cost variable. In all other cases explanatory variables have the same relative importance in explaining service costs as they have in explaining departmental costs.

In analyzing the structure of hospital costs it is extremely useful for policy purposes to determine the relative importance of stay-, admission- and capacity-related costs. In order to accomplish this, independent variables have to be appropriately categorized and then assigned typical values. Stay-related costs would then be calculated as the sum of the products of typical values of the stay-related variables and their estimated cost coefficients. Admission- and capacity-related costs would be similarly calculated. Each cost category could then be compared with the total costs of the typical hospital.

In the service cost equation, purely stay-related variables include total and newborn patient-days (the latter is a component of the newborn nursery output index). Admission-related variables include admissions, trips off-ward, in- and out-patient operations, deliveries and emergency visits. Capacity-related variables include rated patient-days (which also appears as a component of the physical medicine and rehabilitation output index), bassinets set up (a further component of the newborn nursery output index) and the number of operating rooms. Since appropriate capacity measures are unavailable for the remaining services (laboratory, radiology, and long-term, intensive care, private and semi-private ward services), the corresponding output variables have to be interpreted as including a capacity dimension. In such cases it is assumed that 40 percent of output costs are capacity-related (which is the same proportion as for the most important hospital activity, ward service). Hence, 60 percent of laboratory and radiology costs are admission-related and 40 percent capacity-related, while 60 percent of non-standard ward service costs are stay-related and 40 percent capacity-related.

Using typical values for independent variables, stay-, admission- and capacity-related costs are 35, 25 and 40 percent respectively of all hospital costs. In the study done by Evans (1971), they are 46, 40 and 14 percent of all costs; in the study done by Lave, Lave and Silverman (1972), they are 50, 18 and 32 percent of all costs. In addition, the capacity costs associated with ward service are 20 percent of all hospital costs. The values found in this study suggest that a 10 percent reduction in length of stay would reduce hospital costs by 3.5 percent, or by 5.5 percent if 10 percent of beds were also closed. A 10 percent reduction in admissions would reduce both stay- and admission-related costs by 10 percent, so hospital costs would fall by 6 percent, or by almost the full 10 percent if 10 percent of capacity were also closed. The closure of an entire hospital, assuming that patients were subsequently treated in previously under-utilized hospitals, would permit cost savings in the form of the capacity costs of the closed hospital, or 40 percent of its total costs.

It is true that the economies of scale for ECG and nuclear medicine (which are implicit in the laboratory and radiology output indices) are important at the departmental level. It is also true that the minimum fixed costs for nuclear medicine are important at the departmental level. However, because both ECG and nuclear medicine are small departments, overall standardized per diem costs in hospitals providing large amounts of ECG and nuclear medicine are at most 2 percent lower than in typical small output hospitals.

By way of contrast, overall diseconomies of scale are implied by the positive coefficient for squared rated patient-days. It must be remembered, however, that the squared rated patient-days variable proxies for unremoved product heterogeneity in the operating room cost equation as well as indicates diseconomies of scale in the ward service cost equation. If the coefficient of the squared rated patient-days variable in the service cost equation (.0000166) is apportioned among ward service and operating room in accordance with its departmental values (.00000768 and .00000148), 84 percent of its value, or .0000140, is attributable to diseconomies of scale. Diseconomies of scale are such that standardized per diem costs are about 8 percent higher in large (1,000-bed) than in small (100-bed) hospitals. Assuming that large hospitals provide large amounts of ECG and nuclear medicine while small hospitals do not, net diseconomies are such that standardized per diem costs are perhaps 7 percent higher in large than in small hospitals. They compare in magnitude and exceed in statistical significance those found in the case-type cost analysis of Evans.[13]

Substantial indirect teaching costs are indicated by the positive coefficient (.780) for medical teaching expense in the service cost equation, since only a small portion of the coefficient is reasonably attributable to the unremoved product heterogeneity indicated in the departmental cost equations (total coefficient values for medical teaching expense were .157). They result in standardized per diem costs which are about 5 percent higher in teaching than in non-teaching hospitals, which compares with the results of Lave, Lave and Silverman.[14]

The results of the service cost analysis are generally quite satisfactory. The service cost variable is carefully calculated so as to exclude the remuneration of attending physicians. Specification is systematic and sensitive to the forms in which key variables are included. Estimation is made precise through both a priori aggregations and the principal components technique. The degree of explanation of costs is the highest yet achieved (the $\bar{R}^2$ equals .82, compared with .56 in the service cost analysis of Ingbar and Taylor and .62 in the case-type cost analyses of Evans). Clearly an improved methodology is attained, along with powerful insights into the structure of hospital costs which are important for policy purposes.

Despite the shortcomings of restricting the case-type cost analysis to the hospital component of case-type costs, the results of the case-type cost analysis presented next confirm all of the insights into the structure of hospital costs which are provided by the service cost analysis. In particular, they suggest the same relative importance for stay-, admission- and capacity-related costs, returns to scale and indirect teaching costs. In addition, they indicate a similar magnitude for the unexplained portion of case-type and service costs, suggesting that an evaluation of the relative cost performance or efficiency of hospitals in this sample is somewhat insensitive to the choice of service or case-type output definitions.

d. Case-Type Cost Analysis

The results of an analysis of the hospital component of case-type costs for in-patients are presented in Table 4. The case-type cost variable is computed using two alternative procedures. In-patient case-type costs are computed as the costs of all patients less either an estimate of out-patient costs (derived from the service cost analysis) or out-patient income. The values of the alternative in-patient case-type cost variables are highly comparable. The case-type cost coefficients and T-statistics corresponding to the first value of in-patient costs are indicated by the superscript "a" and to the second value of in-patient costs by the superscript "b".

Total short-term admissions is the regression weight in both case-type cost equations, implying that equations are estimated in per admission form. Total patient-days or rated patient-days would also have been suitable regression weights in dealing with heteroscedasticity, but their use would have created a redundancy for the excess patient-days or unused rated patient-days variables already included in the regression.

The principal components technique seems less powerful in dealing with multicollinearity than it is in the service cost analysis. Because the case-type cost analysis makes few empirically justified a priori aggregations, the number of explanatory variables is much larger (49 compared to 19). At least 24 of the maximum of 48 principal components (representing 91 percent of the per admission variation among the independent variables) must be included in the regression (in keeping with an F-test of the incremental R^2 associated with groups of principal components introduced into the regression in descending order of eigenvalues). The retention of such a large number of principal components is bound to reintroduce into the regression a large amount of the multicollinearity among the independent variables, as can be verified by an examination of the low values of the T-statistics of many of the 40 non-standard case-types.

An examination of the major sources of inter-hospital variation in per admission case-type costs proceeds in a manner similar to that outlined for the service cost analysis. Hospitals have per admission costs above the mean (about $484) if they treat disproportionate number of patients diagnosed as suffering from such costly ailments as malignant neoplasms (cancer) or diseases of the central nervous system, if they are large or involved in medical teaching, if they have below-average occupancy rates, or if they are minimally involved in long-term convalescent care.

Estimated coefficients for those independent variables other than the case-type variables are highly significant. All but one, rural rated patient-days, are significant at the one percent level. They are also robust with respect to the alternative definitions of in-patient costs, to the inclusion in case-type costs of such initially excluded cost categories as interest expense, and to the exclusion from the sample of teaching hospitals. Also, as will be seen below, they imply a structure of hospital costs which is similar to that suggested by the service cost equation, which adds to the plausibility of the results of both cost analyses.

The estimated coefficients of the case-type variables are in general not highly significant. This is probably the result of residual collinearity, since the alternative explanation of extremely variable marginal costs for individual case-types seems implausible. However, the case-type cost coefficients do seem to be reasonably robust with respect to the alternative definitions of in-patient costs. The marginal cost of a particular case-type, for example NBMISC (40) is the algebraic sum of its coefficient and that of the short-term admissions variable, i.e., - $116.23 + $290.23 = $174.00 (using the first definition of in-patient costs). The marginal costs of most case-types, especially those that have a reasonable statistical significance, appear to be plausible with few exceptions.

The structure of case-type costs is determinate, but in a less direct manner than for service costs, since the case-type cost coefficients reflect stay- and capacity-related costs as well as admission-related costs. Capacity-related costs can nonetheless be computed by first noting that the costs of (rated-total) patient-days are 6 percent of all costs and that one hospital bed in six was unoccupied during 1971. This suggests that capacity costs are about 36 percent of all costs (assuming standard accommodation for all patients), or 40 percent of all costs (assuming the typical mix of standard and preferred accommodation). Stay-related costs can be computed by first noting that one day of over-stay by a patient costs approximately $25. Since the typical patient stays about 9 days, then the stay-related costs of a patient might be $225 of the total costs of $484, or 46 percent of all costs. However, if patients receive any admission-related services during over-stays, stay-related costs are less than 46 percent of all costs. Admission-related costs are computed residually as being <u>at least</u> 14 percent of all case-type costs.

Net diseconomies of scale are indicated by the coefficient (approximately .0000150) of squared rated patient-days. They imply that standardized per admission costs are 8 percent higher in large than in small hospitals.

Substantial indirect costs for medical teaching programmes are undoubtedly indicated by the positive coefficient (about .60) for medical teaching expense. They imply that standardized per admission costs are about 6 percent higher in teaching than in non-teaching hospitals.

The negative coefficient of rural rated patient-days indicates that standardized per admission costs are about 2 percent lower in rural than in urban hospitals. It reflects differences in illness severity and treatment patterns rather than lower factor prices, since land costs are excluded from the analysis and rural location is insignificant in the service cost equation.

The case-type cost equation is reasonably satisfactory except for the imprecision of the case-type cost coefficients. The degree to which costs are explained, as reflected in the $\bar{R}^2$, is marginally higher than those achieved by Evans (and without the use of age-sex variables) and Lave, Lave and Silverman (and without the use of either age or surgical variables). Clearly an improved methodology in formulating the cost variable, in specification, and in estimation technique is again achieved.

A comparison of the service and case-type cost analysis is made in the next section. As the reader might guess, the alternative cost analyses perform equally well in most respects, although the service cost analysis does seem to permit greater overall precision in estimating cost coefficients.

4. A Comparison of Service and Case-Type Cost Analysis

The service definition of hospital output is advocated in Chapter II as being the output definition appropriate to an examination of a cost variable computed as observed hospital costs, exclusive of the remuneration of attending physicians. The case-type definition of hospital output is, strictly speaking, inappropriate to an examination of this cost variable, since case-type costs are properly computed as hospital costs inclusive of the remuneration of attending physicians.

In spite of this objection to using the case-type definition of hospital output, a case-type cost analysis was undertaken in which case-type output was used to explain the "hospital component" of case-type costs, or observed hospital costs exclusive of the remuneration of attending physicians. Such a case-type cost analysis leads to downwardly biased estimates of case-type cost coefficients since although the total cost variable includes the preponderance of case-type costs (the hospital component incurred in service production), it excludes the remaining case-type costs (the physician component incurred in service demand generation). Moreover, it may result in biased insights into the structure of hospital costs (the relative cost impacts of the independent variables) and biased inter-hospital comparisons of standardized costs (as reflected in cost residuals). It does this to the extent that the inputs of hospitals (employed in service production) and of physicians (employed in service demand generation and the performance of surgical and obstetrical procedures) are imperfect complements and to the extent that the behaviour patterns of attending physicians (in service demand generation) differ among hospitals.

Whether the various types of bias in case-type cost analysis actually occur must to a great extent depend upon sample properties. Although the downward bias in case-type cost coefficients is inevitable as a result of the exclusion (from the hospital component of case-type costs) of the remuneration of attending physicians, the bias in structural insights and in inter-hospital comparisons of standardized costs may be negligible given certain sample properties. In particular, bias will be mitigated if hospital and physician inputs are highly complementary, or any substitution between these inputs small and randomly distributed among hospitals. Bias will be further reduced if the behaviour patterns of physicians in service demand generation either are similar or any differences in them are small and randomly distributed among hospitals.

The case-type cost analysis does seem to be vindicated by sample properties. It suggests a structure of hospital costs which is similar to that implied by the service cost analysis for the relative importance of stay-, admission- and capacity-related costs, diseconomies of scale and the indirect costs of teaching programmes. In addition, it produces an explanation of costs which is as good as that of the service cost analysis (estimated per admission cost equations, corrected for either case-types or services, both have $\bar{R}^2$ of .92). Moreover, it generates per admission cost residuals for individual hospitals (indicators of poor relative cost performance or inefficiency if positive) which are highly correlated (ρ = .60) with those of the service cost analysis.

Case-types are advocated by Evans (1971, 207-208) as the definition of output most appropriate to hospital cost analysis. On the basis of <u>this</u> cost analysis, they are no more appropriate <u>in an empirical sense</u> than are services. In fact, they may be less appropriate to the extent that the precision of estimated cost coefficients is lower as a result of the large numbers of case-type output variables (and principal components) needed to explain costs. In addition, they do not permit the locational insights needed for the exploitation of returns to scale in <u>specific</u> production activities (ECG, nuclear medicine and ward service). Finally, in a conceptual sense, case-types are less appropriate than are services in an examination of hospital costs exclusive of the remuneration of attending physicians.

5. Summary

A statistical cost analysis of a sample of 101 Ontario general hospitals during 1971 incorporates improvements in output definition and the specification and estimation of hospital cost equations. Departmental, service and case-type costs are analysed (although, as a result of data deficiencies, departmental costs exclude any related overhead costs and case-type costs exclude the remuneration of attending physicians). Cost equations are specified in a total cost form which, with the exception of squared capacity terms designed to capture any economies of scale, include all output and characteristic variables in linear form. All cost equations are estimated in an average cost form (in order to avoid heteroscedasticity) using the principal components technique as an intermediate step (in order to increase the precision of estimation while still maintaining interpretability). In addition, estimated departmental cost equations are used to impose aggregative restrictions on the explanatory variables entering into the service cost equation.

A higher degree of explanation of costs and precision of estimation are obtained than were achieved in comparable earlier studies. The insights into the structure of hospital costs which are provided by alternative departmental, service and case-type cost equations are remarkably consistent. Since stay-, admission- and capacity-related costs are estimated to be about 35, 25 and 40 percent respectively of all costs, a larger portion of all hospital costs are capacity-related or quasi-fixed than suggested by other studies. Reductions of capacity are therefore more beneficial and reductions of output less beneficial than previously thought. Although economies of scale exist in two small departments (ECG and nuclear medicine), overall diseconomies of scale (probably in intra-hospital patient transportation and coordination) are of sufficient magnitude to result in standardized per diem costs about 8 percent higher in large (1,000-bed) than in small (100-bed) hospitals.

Although it was necessary (as a result of data deficiencies) to exclude the remuneration of attending physicians from the case-type cost variable, alternative case-type and service cost analyses explain hospital costs equally well and suggest a similar structure of hospital costs. However, the service cost analysis permits more precise estimation since output corrections are fewer as a result of a greater number of available <u>a priori</u> aggregations. In addition, sample properties may not always empirically vindicate the use of case-types in explaining a cost variable more closely approximating service than case-type costs. For these two reasons, service cost analysis is generally preferred to case-type cost analysis for North American hospitals.

Footnotes

1. Homogeneity is suggested by both casual empiricism and other considerations. A number of dummy variables controlling for hospital age had only small and statistically insignificant coefficients in cost equations, and so differences in technology and vintage of capital seem unimportant. Also see Kushner (1969) for confirmation of the inter-hospital similarity of nursing salaries in Ontario.
2. An explanation of these forms can be found in <u>1971 Instructions and Definitions for the Annual Return of Hospitals, Form HS-1 Facilities and Services and Form HS-2 Financial</u> (Ottawa: Statistics Canada - Department of National Health and Welfare).
3. See <u>1971 Edition, the Relative Stay Index Report</u> (Toronto: Ontario Hospital Services Commission).
4. These are presented in <u>Canadian Hospital Accounting Manual: Accounting and Statistical Procedures for Canadian Hospitals</u> (Toronto: Canadian Hospital Association, 1968).
5. See Chapter II, in particular the discussion of the manner in which the activities of service production (by hospitals) and service demand generation (by attending physicians) result in the illness remission of case-types.
6. See Mincer (1962) for further discussion.
7. See Johnston (1972, 218-221). The latter test permits

greater use of available degrees of freedom and so is preferred in hospital cost analysis.

8. Both Kushner (1969) and Cohen (1970) fail to correct for heteroscedasticity and therefore regress total cost equations.

9. All statistical tests are conducted at the 5 percent level unless otherwise stated.

10. The number of standard units to be assigned to each different laboratory test is established by the Statistics Canada publication, Schedule of Unit Values for Clinical Laboratory Procedures.

11. Starkweather (1973) and Williamson (1967) provide empirical and theoretical reasons for this belief.

12. If the weighted independent variables had been orthogonal, the first nine pricipal components (one-half of the total) would have captured 50 percent of the variation among the weighted independent variables.

13. Lave, Lave and Silverman (1972) found no significant returns to scale.

14. Evans (1971) did not control for teaching activity in his cost equation.

VI. POLICY IMPLICATIONS OF STATISTICAL COST ANALYSIS

1. Introduction

The policy implications of the structure of hospital costs which is indicated by statistical cost analysis should be formulated within a benefit-cost framework. Benefit-cost analysis suggests that a particular policy (for example a reduction in the average length of stay of hospital patients) should be pursued to the point that marginal social benefits are just equal to marginal social costs (for example the combined benefits to the patient and society of an extra day of stay should be just as large as its costs). In order to formulate policy recommendations it is therefore important to identify and quantify the benefits and costs associated with practicable policy programmes. The greater the divergence between marginal social benefits and costs in the pre-policy situation, the more highly recommended are policies aimed at transforming the situation into one in which such divergences are eliminated.

It is argued that the situation in the Ontario hospital industry is one in which hospital costs are excessive as a result of both inefficiency and output levels which are too large (in the sense that marginal social costs exceed benefits). It is therefore the case that net social gain can be derived from policies which promote efficiency <u>and</u> optimal output reductions.

Inefficiency in the hospital industry takes many forms. Allocative and X-inefficiency are suggested in studies done for comparable hospital industries by Kaitz (1968), Greenfield (1973), Ogur (1974) and Granfield (1975). They are also suggested in this study by a standard error of estimate which is sufficiently large (5 percent) in relation to standardized hospital costs as to suggest inter-hospital differences in X- and allocative efficiency in addition to any deficiencies of output definition. Scale inefficiency in the form of diseconomies of scale is suggested by Starkweather (1973) and by the observation that over 50 percent of hospital capacity in this study is provided by hospitals which are larger than the scale-efficient 100-300 bed size. Excess capacity costs are suggested by a standard deviation which is large (9 percent) in relation to the mean (83 percent) of occupancy rates and by the observation that a significant number of sample hospitals (about 10 percent) operate at or near full capacity without significantly higher standardized costs.

Excessive output levels are suggested by several considerations. The absence of coinsurance - a requirement that patients insured under the Ontario Health Insurance Plan (OHIP) pay a portion of the costs of their hospital treatment - suggests that (in the absence of externalities and inelastic demand) output of Ontario hospitals is excessive. Beck (1974) asserts that the introduction in Saskatchewan (during 1968) of user charges which were about one-third of all (per diem) costs, reduced the demand for health services in general (and hospital admission in particular) by an estimated 7 percent. Hence the demand for hospital admission is by no means entirely price-inelastic and zero-pricing does in fact encourage hospitalization.

Excessive admissions, length of stay and special services are the result of physician as well as patient behaviour. They permit income and utility augmentation for physicians who wish to increase their billings, reduce malpractice insurance premiums and fulfill their humanitarian aspirations (perhaps somewhat questionably in light of recent media discussions of the alarming increase in certain types of elective surgery such as hysterectomy). Ample excess capacity among Ontario hospitals ensures that the hospitalization plans of patients and physicians are immunized from any serious non-price rationing imposed by queuing.

Strong evidence of excessive length of stay is provided by the introduction of the Relative Stay Index Studies by the OHSC. Some hospitals have a standardized length of stay (correcting for case-mix and age-sex composition, even among hospitals of similar size and teaching status) which is 15 percent higher than the average. Such hospitals are encouraged by the Studies to avoid any unnecessary overstays.

Hospital cost reductions which are achieved through increased efficiency (undertaken in such a manner as to leave quality and accessibility unchanged while ignoring any loss of satisfaction to the hospital decision-maker) are of unambiguous social benefit. In comparison, cost reductions which are achieved through reduced output levels produce on a dollar-for-dollar basis a smaller net social benefit, since patient health status will fall (and/or the output and costs of health sector institutions providing substitute treatment might rise). Assuming a linear demand schedule for the health status provided by hospitals, the same reduction in hospital costs is only one-half as beneficial when achieved through optimal output reductions (those that would be achieved by moving from zero to full user charges) rather than increased efficiency (using the reduction in consumer surplus arising from fewer hospitalizations as a measure of the social welfare loss associated with the expected reduction of health status). In order to ensure the maximum social benefit from output reductions whenever the optimal output reductions are unknown, the hospital regulator should reduce output in small decrements while constantly monitoring and evaluating any increases in other health sector costs and any reductions in patient health status. The last decrement undertaken should produce cost savings which are just equal to the combined dollar value of corresponding changes in other health sector costs and patient health status.

The discussion of the next two sections recommends that hospital regulators pursue policy programmes which reduce hospital costs through increased efficiency and decrements in hospital output. Increased efficiency would permit cost savings equal to about 10 percent of total hospital costs. It would be achieved through a restructuring of the hospital industry in terms of the number, size distribution and allocation of special service facilities and the number and size distribution of hospitals which would best exploit economies of scale (through pooling or centralization of facilities for ECG and nuclear medicine), avoid diseconomies of scale in intra-hospital patient transportation and coordination (through a small hospital construction policy), and more fully utilize both special service and general bed capacity (through closure of as much as one-sixth of existing capacity, perhaps in the form of entire hospitals). Reductions in output should be undertaken whenever the marginal social (hospital) costs avoided are thought to be large in relation to any marginal social benefits foregone (especially in the form of patient health status). Modest reductions (of perhaps 7 percent) in average length of stay and in the number of elective admissions, especially if accompanied by the closure of hospital capacity, would result in proportionate cost savings and proportionately smaller reductions in patient health status. Administrative fiat, hospital reimbursement incentives, user charges or capacity closure could be used to achieve desirable output and cost reductions.

The discussion of an additional section deals with the increased allocative and X-efficiency (and resultant cost savings) and other advantages (especially greater horizontal equity or more equal treatment) which might derive from an incentive reimbursement scheme for hospitals which uses statistical cost estimates (for services or case-types) as the basis for multi-product hospital reimbursement. Although a number of authors, for example Evans (1971), Lave, Lave and Silverman (1973) and Walker (1974) suggest this additional policy application of statistical cost analysis, only Lave et. al. explicitly outline an operational reimbursement scheme. The merits of this scheme in promoting increased allocative and X-efficiency are recognized, and an additional recommendation is made that the hospital regulator monitor and counteract through rate revision any undesirable output charges precipitated by even small inaccuracies in statistical cost analysis. For example, he should downwardly revise the reimbursement rate of an output for which costs are sufficiently overestimated (and initial reimbursement rate correspondingly inflated) as to result in undesirable increases in output levels.

2. Increasing Hospital Efficiency

Increases in hospital efficiency are of unambiguous net social benefit since they reduce the costs but not the benefits provided by the hospital industry. They can be achieved through measures designed to promote scale-efficiency and to eliminate excess capacity. In particular, the hospital regulator can employ administrative fiat to reduce and restructure special service and bed capacity (in terms of the number, size distribution and allocation of special service facilities among hospitals and the number and size distribution of hospitals). The next subsection suggests that the centralization or pooling of ECG and nuclear medicine facilities would exploit economies of scale and permit cost savings which would be 1 or 2 percent of all hospital costs. Reductions in special service capacity by one-sixth would result in cost savings which would be about 3 percent of all hospital costs. A concluding subsection suggests that restrictions on additions to existing hospitals and a construction policy which would require that any new hospitals be of the scale-efficient 100-300 bed size would realize cost savings of perhaps 3 percent of all hospital costs. The elimination of excess bed capacity (perhaps one-sixth of all bed capacity), especially through hospital closure, would result in savings of perhaps 4 percent of all hospital costs. Combined savings from scale-efficiency and maximum capacity utilization in the provision of special and ward services are therefore as much as 10 percent of all hospital costs.

a. A Recommended Reorganization of Special Service Capacity

In order to determine the cost-minimizing number, size distribution and allocation of special service facilities among hospitals, assuming a pre-specified overall requirement for special services, the hospital regulator must first determine the magnitudes of any economies of scale, minimum fixed costs, capacity costs and excess capacity. The economies of scale which are indicated by the ECG and nuclear medicine cost equations are such that unit costs for both ECG and nuclear medicine are about 40 percent lower in typical large output hospitals (500,000 units ECG or 10,000 visits to nuclear medicine per annum) than in small output hospitals (50,000 units ECG or 1,000 visits to nuclear medicine per annum). In addition, minimum fixed costs of $5700 per annum are indicated by the nuclear medicine cost equation. They are about 25 (2.5) percent of all nuclear medicine costs in small (large) output hospitals, which implies that costs per nuclear medicine visit are a further 22.5 percent less in large than in small output hospitals.

It is readily apparent that the total costs of providing a pre-specified requirement for ECG and nuclear medicine would be reduced if output levels of large output hospitals were increased while those of small output hospitals were decreased. It is therefore recommended that ECG and nuclear medicine facilities should be pooled or centralized so that a minority of hospitals within a geographically proximate group of hospitals (e.g. those in Metropolitan Toronto) would specialize in the provision of ECG and nuclear medicine to routine out-patients and to those routine in-patients for whom there is anticipated (prior to admission) a need for ECG (e.g. patients having heart disorders) and nuclear medicine (e.g. patients having cancer). In order to ensure the continuing accessibility of other patients to ECG (which is provided by all sample hospitals) and to nuclear medicine (which is provided by one-half of all sample hospitals), it is further recommended that the remaining hospitals continue to provide ECG and nuclear medicine but at lower output levels than previously. Such a programme would result in savings which would be as much as 2 percent of all hospital costs.

It is not possible from this study to accurately determine the savings which might be realized from the elimination of excess special service capacity, since estimates of excess capacity and related costs are not available for most departments. However, if as an approximation it is assumed that excess capacity is one-sixth of all capacity and capacity costs are 40 percent of all costs (as is true of ward service) and noting that 45 percent of all hospital costs are directly attributable to special services, then the elimination of excess special service capacity would reduce hospital costs by a further 3 percent. It is therefore recommended that the hospital regulator reduce special service capacity by as much as one-sixth (somewhat less if emergency capacity is desired) either by ensuring that selected hospitals reduce staff and equipment requirements, perhaps eliminating them altogether through a pooling arrangement which would divert affected patients to hospitals specializing in particular services, or by closing entire hospitals.

Other authors have investigated the increased efficiency and cost savings which might be realized through a reorganization of special service capacity. However, none has investigated nuclear medicine and only one has investigated ECG. Cohen (1973) finds no economies of scale in ECG. Since his measure of output is tests rather than units, however, his results do not detract from our finding of economies of scale and related recommendations for pooling/centralization arrangements. In addition, Ingbar and Taylor (168, 115) support the recommendation that excess capacity costs in the provision of special services should be eliminated through pooling arrangements.

b. A Recommended Reorganization of Bed Capacity

In order to determine the cost-minimizing number and size distribution (in terms of bed capacity) of hospitals, assuming a pre-specified need for generalized hospital care, the hospital regulator must determine the magnitudes of any economies of scale, minimum fixed costs, bed capacity costs and excess bed capacity. Diseconomies of scale in intra-hospital patient transportation and coordination are suggested by the ward, service and case-type cost equations. Standardized unit costs in the smallest sample hospitals (100 beds) are 8 percent less than those of the largest sample hospitals (1,000 beds). Even if all hospitals provided proportionate amounts of ECG and nuclear medicine (about 1.2 ECG units and .03 visits to nuclear medicine per rated patient-day), small (100-bed) hospitals would still have (otherwise) standardized unit costs almost 8 percent less than those of large (1,000-bed) hospitals.[1] If there are substantial economies of scale in the excluded cost categories (especially teaching and research), hospitals having a capacity of 100 to 300 beds (the smallest size of sample hospitals involved in medical teaching and research), and providing ECG, nuclear medicine, medical teaching programmes and research facilities would appear most efficient.

It is therefore recommended that the hospital regulator discourage additions to existing large hospitals. He should also direct that, wherever possible, new hospitals be relatively small (100-300 beds), thereby ensuring that diseconomies of scale are avoided and accessibility enhanced. Savings equal to about 3 percent of all hospital costs would result.[2] In addition, accessibility of the population to hospital care could be enhanced through the geographical dispersion made possible by a greater number of small hospitals.

Capacity costs are relatively large among the sample hospitals since they are about 40 percent of all hospital costs. Excess capacity is also substantial, since occupancy rates of sample hospitals had a mean of only 83 percent (five-sixths), a large standard deviation of 9 percent, and a considerable range (70 to 100 percent), and a significant number (10 percent) of hospitals operated at or near full capacity (occupancy rate of 100 percent) without significantly higher standardized costs.

It is therefore further recommended that the hospital regulator pursue policies which would ensure that selected hospitals reduce staff and equipment by as much as one-sixth (by a lesser amount if emergency capacity is desired), perhaps closing entire wards. He should also continue with the current programme of closure of one or more hospitals within a group of geographically proximate hospitals while diverting affected patients to previously underutilized hospitals. Savings equal to 4 percent of all hospital costs could be achieved.[3] Combined savings from the elimination of excess bed and special service capacity, the pooling of ECG and nuclear medicine facilities and a small hospital construction policy would be as much as 10 percent of all hospital costs, or $50 million per annum.

Other authors have investigated the benefits which could be derived from the reorganization of hospital bed capacity. Ingbar and Taylor (1968, 114-115) recommend the elimination of excess bed capacity while suggesting that hospital construction policies favouring particular bed capacities are of little benefit (since they detect no significant economies of scale). In the absence of any indication of returns of scale, Ro (1968) recommends a large hospital construction policy (since he feels that the inadequacy of his corrections for product heterogeneity masks underlying economies of scale). Evans (1971, 211-212) suggests that neither the elimination of excess capacity nor hospital construction policies favoring specific bed capacities would be of benefit. None of these results, however, presents a serious challenge to our [illegible]ndations for at least two reasons. First, the cost [illegible] upon which any contrary policy recommendations are b[illegible]uffer from the weaknesses discussed in Chapters II - [illegible] Second, the interpretation of the regression results of these cost studies is suspect in at least one case. Evans suggests that he finds no large or significant diseconomies of scale, despite coefficients for rated bed capacity (in the three versions of his per admission cost equation) which have T-statistics ranging in value from 1.6 to 4.1 and values such that standardized per admission costs are 15 percent higher in 1,000-bed than in 100-bed hospitals. It is therefore the case that other cost studies fail to weaken or in fact strengthen our policy recommendations.

3. <u>Reducing Hospital Output</u>

Net social benefits from hospital output reductions are achieved whenever the decrement in hospital costs exceeds the sum of any increase in the costs of other health sector institutions and the monetary value of any reduction in patient health status. Although the welfare-maximizing number, size and timing of output reductions are theoretically found by optimal control techniques, the output reductions undertaken in practice (recognizing the uncertainty of health status and cost information) should take the form of small decrements in order to avoid a politically unwise overshooting of the optimal output reduction. They can be achieved through administrative fiat, capacity closure or restrictions on new construction (as a means of non-price rationing), reimbursement disincentives for hospitals, or user charges for patients. The magnitudes of both the marginal and capacity costs of specific outputs are key indicators of the desirability of output decrements. Reductions in both output <u>and</u> related capacity result in maximum cost reductions and are especially desirable in view of the relative importance of capacity-related costs.

Noting that output and capacity decrements should be small during each step of output reductions, the next two subsections suggest that the strongest case for output reductions (in the form of fewer admissions, especially in such recently controversial categories as elective hysterectomy, and shorter length of stay) requires capacity closure (perhaps in the form of selected hospital closure). Output reductions of the magnitude achieved in Saskatchewan in the late 1960's (7 percent) seem not unreasonable (since their ultimate reversal through the elimination of user charges in 1971 seemed motivated more by recognition of the regressiveness of user charges rather than a concern with health status).[4] Assuming a linear demand for hospital services, they would produce social benefits of about 5.5

percent of all hospital costs.

a. Reduced Admissions

The hospital cost savings associated with a reduction of admissions, assuming a proportionate capacity closure, take the form of stay-, admission- and capacity-related costs. Since these costs constitute almost all hospital costs (excluding only the minimum fixed costs of providing nuclear medicine), hospital cost savings are approximately proportionate to any reduction in admissions. However, overall cost savings in the health sector would be less to the extent that out-patient care were provided by hospitals and other health sector institutions as a substitute for in-patient care. Moreover, the net social benefit would be reduced even further to the extent that affected patients suffered a loss of health status.

It is therefore recommended that the hospital regulator undertake only those reductions in admissions which yield a net social benefit. Moreover, he should effect these reductions through capacity closure. Capacity closure might take the form of a reduction of capacity in most or all hospitals, or the closing of designated hospitals among groups of geographically proximate hospitals. The latter form of capacity closure would probably be more workable than the former and would perhaps be more effective in terms of discouraging elective admissions while minimizing hardship for other admissions.

Alternative means of reducing admissions include administrative fiat, reimbursement disincentives for hospitals or user charges for patients. However they would not be as effective in reducing hospital costs for an obvious reason - they would not necessarily produce a proportionate reduction in capacity, which would in turn imply a less than proportionate reduction in capacity costs. Total cost savings could therefore be as little as 60 percent of those savings associated with capacity closure (since capacity costs are 40 percent of all hospital costs). Moreover, a less than proportionate reduction in capacity might encourage increased length of stay and therefore result in increased stay-related costs per admission which would further offset any cost savings.

Regardless of the means by which they are achieved, reductions in admissions should be undertaken in small decrements. After each such reduction, the hospital regulator would evaluate health sector cost savings in relation to the monetary value of any loss in patient health status, and proceed with further reductions only if a net marginal social benefit were indicated (i.e. the former exceeds the latter). Optimal overall reductions in admissions might be of the order (7 percent) investigated by Beck (1971, 1).

Some authors suggest that the hospital cost savings which would be foregone as a result of a failure to ensure a proportionate reduction in capacity when reducing admissions would be even greater than suggested here, while other authors suggest they would be less. Ingbar and Taylor (1968, 107-108) suggest that a reduction in admissions without capacity closure would be of little benefit since they detect substantial capacity costs. Evans (1971), 206-207) and Lave, Lave and Silverman (1972, 178) envisage a more substantial benefit from reducing admissions without capacity closure, with savings being 86 and 68 percent respectively of the maximum potential savings associated with capacity closure. Since the methodology of our cost analysis is superior and since no concensus exists among other studies, our recommendations for reduced admissions via capacity closure are made with considerable confidence.

b. Reduced Length of Stay

The hospital cost savings associated with a reduction in average length of stay, assuming a proportionate reduction in ward (bed) capacity, take the form of stay-related costs and that portion of capacity costs associated with ward services. Since these costs are 35 and 20 percent respectively of all hospital costs, hospital cost savings from a 10 percent (one day) reduction in average length of stay would be about 5.5 percent of all hospital costs. However, hospital cost savings would be partially offset to the extent that additional special services were provided during shorter stays. In addition, overall cost savings in the health sector would be further reduced to the extent that out-patient care were provided by hospitals and other health sector institutions as a substitute for in-patient care. Finally, any net social benefit would be further reduced to the extent that patients suffered a loss of health status.

It is therefore recommended that the hospital regulator undertake only those reductions in average length of stay which yield a net social benefit. In addition he should effect decreased length of stay through closure of ward (bed) capacity or even selected hospitals.

Alternative means of reducing average length of stay include administrative fiat, reimbursement disincentives for hospitals and user charges for patients. However, they would not be as effective as capacity closure in reducing costs since they would not necessarily result in a proportionate reduction in bed capacity. Total cost savings might therefore be as little as 35 percent of all hospital costs. Moreover, a less than proportionate decrease in bed capacity might encourage increased admissions, and hence both stay- and admission- related costs, which would further offset any cost savings (and net social benefit if admissions were already at socially optimal levels).

Regardless of the means by which it is achieved, a reduction in average length of stay should be undertaken in small decrements. Each such reduction should be evaluated on a benefit-cost basis (comparing cost savings with any loss of patient health status), and further reductions pursued only if a net social benefit is indicated.

Optimal overall reductions in average length of stay might also be of the order (7 percent) investigated by Beck. Net social benefits from 7 percent reductions in both admissions and length of stay might be equivalent to 5.5 percent of all hospital costs.[5]

Some authors suggest that a reduction in average length of stay would generate considerable cost savings even without capacity closure. Both Evans (1971, 207) and Lave, Lave and Silverman (1972, 178) suggest that a 10 percent reduction in average length of stay would result in at least a 5 percent reduction in hospital costs even without capacity closure, and a 6 to 8 percent reduction with capacity closure. Other authors, notably Ingbar and Taylor (1968, 107-108), suggest that reduced length of stay without capacity closure would result in little cost savings. Since the methodology of our study is superior, particularly in its specification of a reciprocal rather than linear form for length of stay in per diem cost equations, and since no consensus exists among other studies, our recommendation that a (moderately beneficial) reduction in average length of stay be pursued via capacity closure is also made with considerable confidence.

4. Statistical Cost Analysis in the Incentive Reimbursement of Hospitals

A further policy application of hospital cost analysis arises from the increased third-party and governmental involvement in the incentive reimbursement -

especially prospective reimbursement - of hospitals. As Dowling (1974, 163) describes it, "Prospective reimbursement - or more accurately, prospective budget or rate-setting and reimbursement - is a method of paying hospitals in which 1) amount or rates of payment are established in advance for the coming year; and 2) hospitals are paid these amounts or rates regardless of the costs they actually incur". By way of contrast, retrospective reimbursement of hospitals results in payment to hospitals of an amount equivalent to costs actually incurred.

In their discussion of hospital reimbursement, Lave, Lave and Silverman (1973, 85) assert that the ideal reimbursement scheme must promote efficiency and horizontal equity among hospitals as well as the continued economic viability of efficient hospitals.[6] In order to achieve these ends, they suggest a prospective reimbursement scheme in which the reimbursement of an individual hospital is computed as the costs which it would have incurred had it produced its output at the same level of efficiency or unit costs (adjusted for inflation) experienced on average by all hospitals in some previous (base) year. The scheme would promote horizontal equity (equal treatment) and standardize efficiency levels among hospitals (the latter varied ±5 percent in the Ontario sample). It would also through "regulatory lag" - the choice of a base year which would provide the standard for unit costs for several subsequent years - provide an incentive for increased levels of allocative and X-efficiency.[7]

Although absolute efficiency is best served by a policy which allows hospitals to retain a large portion (if not all) of surpluses and requires them to absorb a large portion of deficits, horizontal equity may not be. Since inter-hospital differences in output quality are to be expected, hospitals producing disproportionate amounts of low quality outputs may achieve unearned surpluses while others will incur undeserved deficits. Therefore hospitals would be allowed to retain only a portion of surpluses and be required to absorb only part of any deficits in order to mitigate the horizontal inequity which arises from inter-hospital differences in output quality.

Statistical cost analysis can play a key role in the prospective reimbursement of hospitals. It can provide estimates of the base-year unit costs of individual hospital outputs which are in turn used as guidelines for their re-imbursement. However, because these statistical cost estimates are sensitive to the deficiencies of the specification and estimation of the hospital cost equation, statistical cost analysis may not fulfill its role in an entirely satisfactory manner.

One consequence of inaccuracies in statistical cost estimates is immediately obvious. If overestimates of the unit costs of some hospital outputs and underestimates of the unit costs of others occur, the former outputs will be over-reimbursed and the latter under-reimbursed.

Hospitals providing disproportionate amounts of over-reimbursed services will achieve unearned surpluses while other hospitals incur undeserved deficits. Reimbursement schemes might therefore allow the flexibility in the treatment of surpluses and deficits advocated by Lave, Lave and Silverman because there exists not only inter-hospital variation in output quality but also the possibility of a combination of inaccuracies in statistical cost estimates and inter-hospital variation in output mix.

A second consequence of inaccuracies in statistical cost estimation has not yet been appreciated. Outputs which are over-reimbursed as a result of over-estimates of their unit costs may be favoured by hospitals and their levels increased while other outputs which are under-reimbursed as a result of underestimates of their unit costs may be disfavoured by hospitals and their levels decreased. These predictions are straightforwardly obtained from the theory of profit-maximization by price-taking firms under conditions of upwardly-sloping marginal cost curves. Since firms produce at an output level where price equals marginal costs, they increase the levels of those outputs for which prices rise and decrease levels of outputs for which prices fall. These "side-effects" of reimbursement schemes which employ statistical cost analysis are therefore comparable to the often-discussed side-effects of policies which questionably move from a multiple- to a single-output basis for hospital reimbursement. In order to better understand the former it will be instructive to look at the latter side-effects.

Dowling (1974) informally discusses the impact on hospital behaviour of prospective reimbursement schemes employing alternative single or multiple "payment units" or rate bases. Abstracting from any inaccuracies in formulating reimbusement rates, he catalogues the probable effects of using such payment units as the patient-day, the admission and the vector of individual hospital services. He suggests that hospital reimbursement schemes can affect both output levels (length of stay, admissions and special services) as well as quality and efficiency in a qualitatively predictable manner. For example, he anticipates that the prospective per diem reimbursement of hospitals encourages increased length of stay[8] and provides no disincentive to inefficiency. An earlier study by Pauly and Drake (1970) provides empirical support for the latter but not the former prediction.

The expectation of increased output levels for the adopted rate base, for example total patient-days through increased length of stay under per diem reimbursement, and reduced output levels for other outputs, is consistent with a model in which the hospital is a profit-maximizing multi-product firm (employing separable production processes) which experiences product price changes. The effect of adopting a single rate base is to increase the price of the rate base variable above initial marginal costs and to decrease the price of other output variables below initial marginal costs. The hospital is therefore encouraged to expand production of the rate base variable and reduce production of other outputs.

The quantitative impact ~~on output~~ levels of adopting a single rate base reimbursement scheme may be affected if interactions in production are admitted into the model. For example, if increased length of stay (or a *ceteris paribus* reduction in admissions) facilitates the production of stay-related services, the hospital decision-maker may to a greater extent react to per diem reimbursement by encouraging increased length of stay (and less than proportionately encouraging or even discouraging admissions). However, the relevance of such interactions for hospital behaviour is probably slight in view of the finding of no large or statistically significant interactions in the cost analysis of the previous chapter.

The quantitative impact on hospital output of adopting a single rate base reimbursement scheme may also be affected by the introduction of a degree of not-for-profit behaviour into the model. For example, hospital decision-makers might derive satisfaction not only from profits but also from the levels of various outputs and inputs. If total patient-days (high occupancy rates) are a direct source of utility to the hospital decision-maker, one anticipates that his response to a move from a multiple-output to per diem reimbursement will greatly exceed that occurring in the profit-maximizing situation. However, if it is remembered that the analysis requires that the rates of reimbursement under the multiple-output and per diem reimbursement schemes are contrived to yield the same overall level of hospital reimbursement, the length of stay response of the hospital decision-maker may not be substantially greater in the not-for-profit than in the for-profit situation.

Inaccuracies in the estimates of the unit costs of individual hospital outputs may, even if a multiple-output basis for reimbursement is pursued, result in incentives for output changes similar to those that occur when a move is made away from a multiple- to a single-output basis for reimbursement. For example, if the unit costs of stay-related services are overestimated while those of other services are underestimated, hospitals may increase total patient-days, primarily through the low-cost route of increased length of stay rather than increased admissions.

The reimbursement authority is encouraged to use statistical cost estimates in the incentive reimbursement of hospitals, but it must anticipate the possibility and effects of inaccurate estimation. It must monitor output levels, especially such key indicators as occupancy rate and length of stay. It can in subsequent periods downwardly adjust reimbursement rates for outputs experiencing disproportionate increases in volume while upwardly adjusting reimbursement rates for other outputs. It will thereby not only promote efficiency and a greater degree of horizontal equity but also mitigate undesirable output changes.

5. Summary

The policy implications of statistical cost analysis must be formulated within a benefit-cost framework. The reduction of any inefficiencies (allocative, X- and scale-inefficiency as well as excess capacity costs) unambiguously increases the net social benefit provided by the hospital industry. Output and related capacity reductions are less powerful in increasing the net social benefit provided by the hospital industry since the benefits resulting from hospital cost savings are to some extent offset by increases in other health sector costs and reductions in patient health status. In order to ensure a social welfare maximum, output reductions should be undertaken in small decrements as long as related benefits exceed costs.

Cost savings of approximately 10 percent of all hospital costs can be achieved through the scale-efficiency associated with the centralization of ECG and nuclear medicine facilities among hospitals and a small hospital construction policy, and through the elimination of up to one-sixth of existing special service and bed capacity (perhaps by hospital closure). Cost savings (of less unambiguous social benefit) can be achieved through measures designed to reduce admissions and length of stay (and related capacity), especially capacity closure and to a lesser extent administrative fiat, reimbursement disincentives for hospitals and user charges for patients. Output reductions of about 7 percent would result in cost savings of a further 11 percent of hospital costs, but a somewhat smaller increase in the net social benefit provided by the hospital industry. A benefit-cost analysis of small decrements in output may, however, suggest a different value for optimal output reductions.

Statistical cost analysis can be used to provide guidelines for the formulation of reimbursement rates for individual hospital outputs (services or case-types) in an incentive reimbursement scheme for hospitals. Such a scheme should incorporate regulatory lag (in order to promote increased efficiency) as well as some flexibility in the treatment of deficits and surpluses (in order to preserve horizontal equity). It should also counteract through rate revision any undesirable output changes which result from inaccuracies in statistical cost analysis.

Footnotes

1. The cost-minimizing hospital size is readily derived from the service cost equation. The discussion of Chapter V suggests that the coefficient of the squared rated patient-day variable is attributable to both diseconomies of scale and product heterogeneity. That portion of the coefficient attributable to diseconomies of scale is .0000140. The coefficients of the squared units of ECG and the nuclear medicine squared visits and dummy variables (which are implied by the laboratory and radiology output indices and coefficients) are -.000000202, -.00125 and 8527.66 respectively.

The total costs of providing hospital services (TC_s) depend upon the squared terms in rated patient-days (Q), units of ECG (ECG) and nuclear medicine vists (NMV); the capability of providing any nuclear medicine; and other variables listed in Table 3 such as total patient-days, rated patient-days and newborn nursery service (X_i, i = 1, 2...I). In symbols,

$$(1) \quad TC_s = .0000140Q^2 - .000000202ECG^2 - .00125NMV^2 + 8527.66 + \sum_{i=1}^{I} a_i X_i + \text{error}.$$

For the typical hospital providing ECG and nuclear medicine,

$$(2) \quad ECG = 1.34Q$$

and
$$(3) \quad NMV = .0265Q$$

It is also true that for the typical hospital,

$$(4) \quad X_i = k_i Q \qquad i = 1, 2...I$$

Substituting equations (2), (3) and (4) into (1) gives the result:

$$(5) \quad TC_s = .0000128Q^2 + 8527.66 + \sum_{i=1}^{I} a_i k_i Q + \text{error}.$$

Now costs per rated patient-day (AC_s) are found by dividing equation (5) by Q:

$$(6) \quad AC_s = .0000128Q + 8527.66/Q + \sum_{i=1}^{I} a_i k_i + \text{error}'$$

For minimum costs per rated patient-day (or minimum total costs for a pre-specified aggregate rated bed capacity), $dAC/dQ = 0$. Substitution of equation (6) yields the result:

$$(7) \quad .0000128 - 8527.66/Q^2 = 0$$

Solving gives $Q = 25,811$ rated patient-days. Dividing Q by 365 (days per year) gives the cost-minimizing bed capacity for an individual hospital as 70.7 beds (which has a 95 percent confidence interval of 57 - 93 beds).

Equation (6) suggests that hospitals of 100 beds have standardized per rated patient-day costs which are only $.04 above the cost-minimizing 70-bed hospital, while hospitals of 1,000 beds have costs which are $4.04 (7 percent above those of a 70-bed hospital).

2. The per annum cost savings which would arise if all hospitals were of 300 bed capacity (CS) are computed as follows:

$$CS = \sum_{i=1}^{101} .0000140(Q_i^2 - (300 \times 365)^2),$$

where Q_i is the rated patient-day capacity of the ith hospital. These cost savings can be expressed as percentage of all hospital costs, $\sum_{i=1}^{101} TC_i$.

3. The relative cost saving of 4 percent per annum is found by noting that one bed in six is typically unoccupied, 40 percent of all ward costs are capacity-related, and ward costs are 55 percent of all hospital costs. Relative cost savings (RCS) are computed as follows:

$$RCS = \frac{1}{6} \times .40 \times .55 \times 100\% \text{ or about } 4\%.$$

4. See Beck (1974, 1). By regressiveness it is meant that user charges resulted in a greater relative reduction of real disposable income (in the form of greater direct outlays on health care or reduced health status) for the poor than the rich.

5. A 7 percent reduction in admissions and related capacity reduces hospital costs by 7 percent. A 7 percent reduction in length of stay and related capacity reduces hospital costs by 4 percent (stay-related costs only). Net social gain is one-half of cost savings, assuming a linear demand schedule for health status and optimality of the 7 percent output reduction.

6. Efficiency is interpreted in this context as allocative and X-efficiency. Horizontal equity or equal treatment (i.e. equally efficient hospitals are equally rewarded by the reimbursement scheme) is desirable for reasons of both inherent fairness and strengthening incentives for efficiency.

7. Lave, Lave and Silverman do not actually use the term "regulatory lag".

8. He also predicts increased admissions. This prediction, however, seems to assume medically or administratively imposed upper limits on length of stay which are sufficiently restrictive as to require hospitals to pursue increased total patient-days through the relatively costly route of increased length of stay and admissions rather than increased length of stay with constant or even decreased admissions.

VII. SUMMARY AND CONCLUSIONS

The study has been successful in attaining its threefold objective. It developed an improved methodology for the statistical cost analysis used in determining the structure of hospital costs. It was successful in determining the structure of a sample of (Ontario) hospitals with a greater degree of explanation and precision of estimated cost coefficients for the departmental, service and case-type cost categories analysed. Finally it utilized the insights into the structure of hospital costs in the examination of current hospital policy issues.

An improved methodology for the statistical cost analysis of hospitals requires more careful definition of hospital output and more careful specification and estimation of hospital cost equations. The definition of hospital output should take full account of institutional arrangements in the hospital industry. If attending physicians are salaried or fee-for-service hospital staff members, hospital output is case-types. If attending physicians are private practitioners, hospital output is services.

A more careful specification of the hospital cost equation proceeds without overdue regard for anticipated problems in estimation arising from the inefficiency and imprecision associated with heteroscedasticity and multicollinearity respectively. Although the total cost equation must usually be reformulated in an average form for efficient estimation, it better focusses the attention on the analyst on the form of output variables which is most consistent with prior information on the nature of the underlying hospital production process. For example, a total cost equation which is linear in rated and total patient-days and admissions clearly assumes the absence of returns to scale and interactions in production (other than those captured by polynomial and cross-product output terms). Assuming such a total cost equation is plausible, the equivalent per diem cost equation includes reciprocal rather than the linear terms in occupancy rate and length of stay which have characterized the cost literature. In addition, although a cost equation including all plausible output corrections will often have to be subjected to aggregation in order to ensure precise estimation, it will usually yield greater product homogeneity in its final form than a cost equation which eliminates one or ~~more~~ plausible output corrections before fully investigating alternative methods of achieving precise estimation.

A synthesis of aggregation techniques permits estimation more precise than that achieved in other hospital cost studies. A priori information should be generated and used to aggregate output variables. For example, previously estimated departmental cost equations provide estimates of service cost coefficients which can be used to aggregate groups of services into output indices to be used in the explanation of all hospital costs. The principal components technique can be used in an intermediate manner in order to yield estimates of cost parameters which are more precise (have smaller mean square errors) than those achieved by ordinary least squares alone. Stepwise regression, used in conjunction with the principal components technique, can also be used to eliminate theoretically unjustifiable and statistically insignificant output corrections in an orthogonal manner (thereby avoiding selectivity bias).

A statistical cost analysis of a sample of 101 Ontario general hospitals during 1971 was highly successful. Departmental, service and case-type cost equations were estimated. The degree of explanation of hospital costs was higher (for example, $\bar{R}^2$ = .82 for per diem costs) than that achieved in comparable studies done by Evans (1971) and Evans and Walker (1972) for Ontario and by other authors such as Ingbar and Taylor (1968) for U.S. hospitals. The robustness and significance levels of cost coefficients were also higher. Services and case-types appear to explain the costs of Ontario hospitals with equal success and to indicate a similar structure of hospital costs in terms of stay-, admission- and capacity-related costs (35, 25 and 40 percent respectively of all hospital costs) and net diseconomies of scale (standardized per diem costs were 8 percent higher in large 1,000-bed than in small 100-bed hospitals). In addition, the departmental and service cost equations indicated significant economies of scale (although globally unimportant due to smallness of departmental budgets) for nuclear medicine and ECG.

The policy implications of the indicated structure of hospital costs are explored in considerable detail. Although their divergence from the recommendations of other studies is primarily one of emphasis, they are contrasting in at least one respect - they recommend a small hospital construction policy.

It is recommended that the hospital regulator restructure special service and bed capacity in order to achieve scale-efficiency and eliminate excess capacity. A pooling of ECG and nuclear medicine facilities within groups of geographically proximate hospitals, a small hospital (100-300 beds) construction policy, and the elimination of excess capacity (perhaps through closure of some hospitals and increased utilization of others) would result in cost savings (10 percent of all hospital costs) of unambiguous net social benefit. It is further recommended that the hospital regulator pursue capacity closure (rather than administrative fiat, reimbursement disincentives for hospitals or user charges for patients) in order to reduce (elective) admissions and average length of stay (by perhaps 7 percent) to socially optimal levels. Finally, it is recommended that the hospital regulator employ statistical cost estimates of the unit costs of services or case-types as a basis for the incentive reimbursement of hospitals. However, it is recognized that he must be prepared to monitor the output levels and decrease (increase) the reimbursement rates of those services or case-types for which unit costs are sufficiently overestimated (underestimated) as to result in undesirably large increases (decreases) in output levels.

The need for further research is apparent. Better standardized output measures are needed for individual hospital departments such as the operating and delivery rooms. Capacity measures are needed for many services: laboratory, radiology and the delivery room to name but a few. Finally, further investigation of the sources of economies of scale in special service departments and overall diseconomies of scale, perhaps through engineering cost studies, is not only desirable as a means of corroborating the findings of this hospital cost analysis but also as a further guide to policy.

BIBLIOGRAPHY

Anderson, D. L. (1974) 'Economic Aspects of Clinical Laboratory Service', paper presented at the Symposium for Health Care Economists, Queen's University.

Anderson, R. and J. T. Hull (1969) 'Hospital Utilization and Trends in Canada and the United States', *Health Services Research* 4, 198-222.

Bain, J. S. (1954) 'Economies of Scale, Concentration and the Condition of Entry in Twenty Manufacturing Industries', *American Economic Review* 44, 15-39.

Beck, R. G. (1974) 'Some Dynamic Effects of Copayment on Utilization of Medical Services in Saskatchewan' paper presented at the Symposium for Health Care Economists, Queen's University.

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

Berry, R. (1970) 'Product Heterogeneity and Hospital Cost Analysis', *Inquiry* 7, 67-75.

Bijnen, E. J. (1973) *Cluster Analysis* (Netherlands: Tilburg University Press)

Burgess, D. F. (1975) 'Duality Theory and the Pitfalls in the Specification of Technologies', *Journal of Econometrics* 3, 105-121.

Carr, W. J. and P. J. Feldstein (1967) 'The Relationship of Cost to Hospital Size', *Inquiry* 4, 45-65.

Cheng, D. C. and H. J. Iglarsh (1976) 'Principal Component Estimators in Regression Analysis', *Review of Economics and Statistics* 58, 229-234.

Cohen, H. A. (1967) 'Variations in Cost Among Hospitals of Different Sizes', *Southern Economic Journal* 33, 113-122.

__________ (1970) 'Hospital Cost Curves with Emphasis on Measuring Patient Care Output', in H. E. Klarman (ed.) *Empirical Studies in Health Economics* (Baltimore: The Johns Hopkins Press), 279-293.

__________ (1973) 'Cost Functions of Hospital Diagnostic Procedures: A Possible Argument for Diagnostic Centres', *Journal of Economics and Business* 26, 83-88.

Davis, K. (1972) 'Economic Theories of Behaviour in Non-profit Private Hospitals', Economic and Business Bulletin 24, 1-13.

Davis, K. and L. B. Russell (1972) 'The Substitution of Hospital Outpatient for Inpatient Care', Review of Economics and Statistics 54, 109-120.

Dowling, W. L. (1974) 'Prospective Reimbursement of Hospitals', Inquiry 11, 163-180.

Dusansky, R. and P. J. Kalman (1974) 'Toward an Economic Model of the Teaching Hospital', Journal of Economic Theory 7, 210-223.

Evans, R. G. (1971) '"Behavioural" Cost Functions for Hospitals', Canadian Journal of Economics 4, 198-215.

Evans, R. G. and H. D. Walker (1972) 'Information Theory and the Analysis of Hospital Cost Structure', Canadian Journal of Economics 5, 398-418.

Feldstein, M. (1968) Economic Analysis for Health Service Efficiency (Chicago: Markham Publishing Company)

Feldstein, P. J. (1961) An Empirical Investigation of the Marginal Cost of Hospital Services (Chicago: University of Chicago Press).

_______________ (1968) 'An Analysis of Reimbursement Plans' in Reimbursement Incentives for Hospital and Medical Care (Washington: U.S. Department of Health and Welfare)

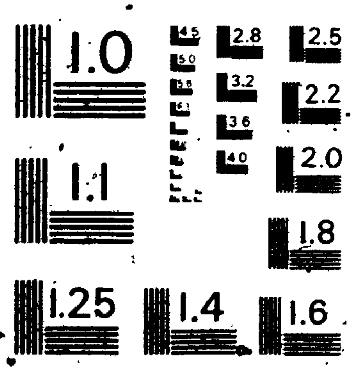
Francisco, E. W. (1970) 'Analysis of Cost Variations among Short-term General Hospitals' in H. E. Klarman (ed.) Empirical Studies in Health Economics (Baltimore: Johns Hopkins Press)

Fraser, R. D. (1971) Canadian Hospital Costs and Efficiency (Ottawa: Economic Council of Canada Special Study No. 13)

Fuchs, V. R. (1972) Essays in the Economics of Health and Medical Care (New York: National Bureau of Economic Research)

Granfield, M. E. (1975) 'Resource Allocation Within Hospitals: An Unambiguous Analytical Test of the A-J Hypothesis', Applied Economics 7, 241-249.

3 OF/DE 3


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6
PY RESOLUTION TEST CHART
REAU OF STANDARDS - 1963 - A

Greenfield, H. I. (1973) Hospital Efficiency and Public Policy (New York: Praeger Publishers)

Grossman, M. (1972) The Demand for Health: A Theoretical and Empirical Investigation (New York: National Bureau of Economic Research)

Grundy, R. A. (1971) 'Hospital Accounting', Management Accounting 53, 45-47,50.

Haldi, J. and D. Whitcomb (1967) 'Economies of Scale in Industrial Plants', Journal of Political Economy 75, 373-385.

Harris, S. E. (1964) The Economics of American Medicine (New York: Macmillan)

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

Ingbar, M. L. and L. D. Taylor (1968) Hospital Costs in Massachusetts (Cambridge, Massachusetts: Harvard University Press)

Jacobs, P. (1974) 'A Survey of Economic Models of Hospitals', Inquiry 11, 83-97.

Jaszi, H. H. (1964) The Economics of Health and Medical Care (Ann Arbor: University of Michigan Press)

Johnston, J. (1960) Statistical Cost Analysis (New York: McGraw-Hill)

__________ (1972) Econometric Methods (New York: McGraw-Hill)

Joseph, H. and S. Folland (1972) 'Uncertainty and Hospital Costs', Southern Economic Journal 38, 267-273.

Joseph, H. (1974) 'On Economic Theories of Hospital Behaviour', Journal of Economics and Business 27, 69-74.

Kaitz, E. M. (1968) Pricing Policy and Cost Behaviour in the Hospital Industry (New York: Praeger Publishers)

Klarman, H. E. (1965) The Economics of Health (New York: Columbia University Press)

__________ (1970) Empirical Studies in Health Economics Baltimore: Johns Hopkins Press)

Kushner, J. (1969) 'Economies of Scale in the General Hospital Industry', unpublished PhD dissertation, University of Western Ontario.

Lave, J. R. and L. B. Lave (1970) 'Hospital Cost Functions', American Economic Review 60, 379-395.

__________ (1971) 'The Extent of Role Differentiation Among Hospitals', Health Services Research 6, 15-38.

Lave, J. R., L. B. Lave and L. P. Silverman (1972) 'Hospital Cost Estimation Controlling for Case-Mix', Applied Economics 4, 165-180.

__________ (1973) 'A Proposal for Incentive Reimbursement for Hospitals', Medical Care 11, 79-90.

Lee, M. L. (1971) 'A Conspicuous Production Theory of Hospital Behaviour', Southern Economic Journal 37, 48-58.

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

Long, M. F. (1964) 'Efficient Use of Hospital', in H. H. Jaszi (ed.) The Economics of Health and Medical Care (Ann Arbor: University of Michigan Press), 211-226.

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

McCall, J. J. (1970) 'The Simple Economics of Incentive Contracting', American Economic Review 60, 837-846.

McCallum, B. T. (1970) 'Artificial Orthogonalization in Regression Analysis', Review of Economics and Statistics 52, 110-113.

McNerney, W. J. (1962) Hospital and Medical Economics (Chicago: Hospital Research and Educational Trust)

Migue, J. L. and G. Belanger (1974) The Price of Health (Toronto: Macmillan)

Mincer, J. (1962) 'On-the-Job Training: Costs, Returns, and Some Implications', Journal of Political Economy 70, 50-79.

Montacute, C. (1962) Costing and Efficiency in Hospitals (London: Oxford University Press)

Newhouse, J. P. (1970) 'Toward a Theory of Non-profit Institutions: An Economic Model of a Hospital', American Economic Review 60, 64-74.

Ogur, J. (1974) 'The Nonprofit Firm: A Test of the Theory for the Hospital Industry', Journal of Economics and Business 26, 115-120.

Pauly, M. V. (1970) 'Efficiency, Incentives and Reimbursement for Health Care', Inquiry 7.

Pauly, M. V. and D. F. Drake (1970) 'Effect of Third-Party Methods of Reimbursement on Hospital Performand', in H. E. Klarman (ed.) Empirical Studies in Health Economics (Baltimore: Johns Hopkins Press), 297-319.

Pauly, M. V. and M. Redisch (1973) 'The Not-For-Profit Hospital as a Physicians' Cooperative', American Economic Review 63, 87-99.

Peacock, A. T. and J. Wiseman (1972) The Economics of Medical Care (London: George Allen and Unwin Limited)

Perlman, M. (1974) The Economics of Health and Medical Care (London: Macmillan)

Rafferty, J. (1972) 'Hospital Output Indices', Economic and Business Bulletin 24, 21-27.

Reder, M. M. (1965) 'Some Problems in the Economics of Hospitals', American Economic Review 55, 472-480.

Ro, K. K. (1968) 'Determinants of Hospital Costs', Yale Economic Essays 8, 185-257.

________ (1969) 'Incremental Pricing Would Increase Efficiency in Hospitals', Inquiry 6, 28-36.

Salkever, D. S. (1970). 'Hospital Cost Studies and Planning Under Uncertainty: Analysis of a Simple Model', Southern Economic Journal 36, 263-267.

Scherer, F. M. (1971) Industrial Market Structure and Performance (Chicago: Rand McNally)

Scott, J. T. (1966) 'Factor Analysis and Regression', Econometrica 34, 552-562.

Shepherd, W. G. (1967) 'What Does the Survivor Technique Show about Economies of Scale', Southern Economic Journal 33, 113-122.

Shomick, W. (1970) 'A Stochastic Model for Occupancy-Related Random Variables in General Acute Hospitals', Journal of the American Statistical Association 65, 1474-1500.

Sigmond, R. (1969) 'The Notion of Hospital Incentives', Hospital Progress 50, 63-68, 97.

Starkweather, D. B. (1973) 'Hospital Organizational Performance and Size', Inquiry 10, 10-18.

Stigler, G. J. (1958) 'The Economies of Scale', Journal of Political Economy 66, 54-71.

Toussaint, J. R. (1970) 'The Problem with Hospital Accounting', Management Accounting 52, 37-40.

Walker, H. D. (1974) 'Reimbursement of Hospitals Based Upon Case-Mix', paper presented at the Symposium for Health Care Economists, Queen's University.

Williamson, O. E. (1967) 'Hierachical Control and Optimum Firm Size', Journal of Political Economy 75, 123-138.

Wonnacott, R. J. and T. H. Wonnacott (1970) Econometrics (New York: Wiley)