THE DESIGN OF A BUDGET-BASED CONTRACT AS A TOOL FOR INCENTIVE MOTIVATION TO IMPROVE EFFICIENCY IN THE ALLOCATION OF RESOURCES IN THE HEALTH CARE SECTOR, WITH SPECIAL REFERENCE TO THE PUBLIC HOSPITAL SECTOR IN SPAIN.

By

GUILLERMO LOPEZ-CASASNOVAS

Thesis submitted for the degree of D. Phil.

University of York

Department of Economics and Related Studies

17 September 1984

C O N T E N T S

LIST OF TABLES AND FIGURES.

TABLES

FIGURES

ACKNOWLEDGEMENTS

The present thesis is the result of four years of intensive research in the field of Public Economics at the University of York and my interest on the subject, arised from the start of my MSc course.

I am indebted to my supervisors Ron Akehurst, Bernard Stafford and Professor Alan Williams, who with a blend of friendliness and toughness have not only been a great help in the realization of this thesis, but, more importantly, have shown me what research in public economics is about.

I am grateful too to some of my British friends who have gone through various chapters of this thesis and helped me to produce a more readable text.

I would like to thank, for financial support, the Spanish Ministry of Education and the Fundación Oriol-Urquijo, and, in the final stages of my stay in England, the Fundació de Cultura Catalana and the Department of Health and Social Security of the Generalitat de Catalunya.

My special thanks are also due to Professors Alexandre Pedrós-Abelló and Francisco Domínguez-Del Brío, who made the first suggestions for this research and continued to encourage me in it, and who have accepted me as a member of the Department of Public Finance in the Faculty of Economics in the University of Barcelona.

ABSTRACT

The following research develops from the seminal studies by M. S. Feldstein on Hospital Efficiency and by R. Evans on Incentive Reimbursement in the Hospital Sector.

Within this field, the thesis aims to analyse the present and potential role of public hospital funding policies for incentive promotion and the implementation of efficiency guides in allocating health resources.

Regarding present budget mechanisms in the Spanish Hospital Sector, at no stage of the decision-making process can we find any rationale favouring an efficient allocation. Indeed, actual financial arrangements appear to maintain 'system-inefficiency', which the currently-debated 'poor' performance of the sector appears to reflect.

Our analysis shows, that, in addition to the fact that the system as a whole seems to operate in an inefficient manner, some parts seem to behave more inefficiently than others. This observed issue is examined in the context of 'variable productivity' and its behavioural implications on the institutional performance are then expounded.

Thus we conclude that present policies of financial restraint, which follow a previously expansionary era, are not only well addressed to tackle both former problems - failing to insert efficiency motivation targets - but they actually create 'perverse' effects in terms of vertical equity, as a result of applying spending limits across the board on individual past hospital expenditure. Therefore, and confliciting with the principles postulated by most of these policies, no rational elements of efficiency motivation (holding 'output' constant) exist in reality to achieve cost-saving targets.

This thesis recognises in detail the actual setting of system-inefficiency and of 'variable productivity' in the Hospital Sector and provides an operational approach, tackling the aforesaid problems, aimed to correct them through a change in the present budget arrangements.

A prospective Revenue Allocation Policy, (PRAP) is thus designed which:

(i) establishes the budget relationship involved between the Central Funding Authority and the Hospital Board, in terms of an 'Agency' relationship from which an 'optimal' budget-based contract is defined;

(ii) recognises differential levels of performance, on which differentiated reimbursement levels can be advanced, under objectively verifiable conditions by both parts of the relation, for the prior commitment of the total amount of resources available for distribution within the sector;

(iii) proves its effects in promoting incentives for the efficient performance of the institution over time;

(iv) finally, the actual construction of the PRAP enables us moreover to test some hypothesis put forward in the thesis, such as the absence of self-correcting mechanisms and a cummulative causation effect on hospital performance over time.

INTRODUCTION

Our interest in the field of the promotion of incentives which would introduce efficiency-guides in the allocation of resources in public sector activities goes back to the initial reading for an MSc project. Indeed, at that time I became very involved in the literature on the private-versus-public provision debate, concerning certain service activities for which an 'appealing' picture of the poor public sector performance was presented.

Having returned for a brief period to Spain, I realized that what seemed initially to me to be literature which was merely jumping on the political band-wagon in Great Britain, with the revival of the old Smithian views on national wealth, was being imported wholesale by some Spanish colleagues in almost identical terms.

At first it appeared that these views on the 'catastrophic' performance of the Public Sector for most of its activities could be considered extrapolable everywhere at different periods and under quite different conditions. But obviously, this must be over-simplification. In fact, any analysis of this whole area must be undertaken with care and precision, the case must be supported by the detailed study of particular bits of public activity, and centred on space and time coordinates.

The second issue which came to mind was that the conclusions derived from this debate should prove the legitimacy of any comparison between two different sectors in producing identical outputs, under, at the very least, similar conditions of provision.

After detailed analysis of local public services, it had to be considered that, again, such premises were grounded in over-simplification: too many terms of the comparison were postulated rather than tested, remaining,in our view, in unrealistic domains.

It then became evident that an effort to reverse the issues of controversy for analysis on pure economic grounds should lead us to some other more solid arena. We therefore focussed our interest towards those aspects related to the performance of public sector activities which could be treated in a positive way. We redressed attention towards some intra-sector public activities which were being produced by statutorily designed identical units on decentralized bases, under explicitly regulated conditions of provision. At a first glance, the obvious feature which appeared from an exploration of empirical analyses was that, in this more restricted set, a wide range of variation in most of the performance indicators could be observed, even after adjusting for various exogenous factors which might make for some differences in the conditions of provision.

In searching for the reasons which could explain this phenomenon of so-called 'variable productivity', we felt that preferential attention should be devoted to the way in which revenues are allocated to those decentralized institutions, ultimately encharged of provision of the service.

Examination of those budget arrangements suggested then, a first explanation: lack of incentives at any stage of the decision-making process, which run parallel to a guarantee of reimbursement for all claimed spending. Thus, no intrinsic motivation was installed in an effort to improve performance over time. On the contrary, because of the utilization of the historical pattern of expenditure for revenue allocation purposes, single 'decisional' variables are allowed to affect all levels of the management performance of the institution.

If this had been the case in an expansionary age, in which any budget request of the institution was bound, finally, to be reimbursed, the result, once financial restraints have started to be imposed on the sector, usually under the form of cash-limits with across-the-board increases, may be particularly 'perverse'. Although the rationale of these policies is justified particularly in grounds of 'efficiency', our conclusion in the former context is just the opposite: These policies are, in fact, penalising those previously 'cost-conscious' units and rewarding, in relative terms, the high spenders. In addition, these policies once more encouraged indulgence in strategic budgetary games and the search for ways to breach the budget iteration process at its weaker points.

The question to which this thesis is adressed then became obvious: <u>How can an alternative budget policy, which takes into account the phenomenon of differential performance, be designed for purposes of providing incentive motivation and inserting efficiency-guides in the allocation of public resources?</u>

The context for the operationalisation of our research is henceforth straightforward. Given our purpose of undertaking the analysis of some area within public sector activity, and, optimally, to a single branch of activity, and given our stay in York, we decided to take profit of this 'locational rent' and apply our research to the Public Health Care Sector, in particular to the Hospital inpatient care.

In addition, we were committed to seek a contribution to the field of Health Economics in Spain, despite the constant threat of inadequate statistical data.

Hospital health care and health care expenditure otherwise fitted well with our initial objectives. In point of fact, this sector enjoyed very rapid expansion in Spain until the end of the seventies, a sort of 'black hole' which has constantly claimed additional resources and which has been the subject of a large amount of literature concerning its very poor performance record.

This sector has, as have other parts of the public sector in Spain, been recently stringently tightened under policies of financial restraint and the threat of expenditure cuts. This has increased the disruption in hospital activity and the dissatisfaction of the users; all of which comes, apparently, to validate arguments in favour of alternative systems of provision on behalf of 'efficiency' targets.

CHAPTER I THE GROWTH AND RESTRAINT OF HEALTH CARE EXPENDITURE IN SPAIN: RATIONALE AND CONSEQUENCES

CHAPTER I THE GROWTH AND RESTRAINT OF HEALTH CARE EXPENDITURE IN SPAIN: RATIONALE AND CONSEQUENCES.

1.1. INTRODUCTION

During this last decade there has been a growing concern about the rise of health care expenditure in advanced economies. In Spain, as in many other European countries, this concern has been, in practice, put in the context of a more general issue: the increase in the share of the Gross Domestic Product absorbed by public expenditure in the economy. Upon this matter we must (i) distinguish the inferences raised from value judgements and the tested features, on purely economic grounds, of observed reality (see, in this sense, the analysis presented in the Appendix to this Chapter), and (ii) we must recognize that, strictly speaking, the former phenomenon has appeared under different systems of health care provision in various countries.

In relation to this concern, a set of control policies for expenditure restraint has been recently proposed. This Chapter deals with the contribution that past and present policies for the control of health care expenditure in Spain have exerted and can exert on the efficiency motivation, which is the most commonly cited justification for a more cost-conscious allocation of resources in the Health Care Sector. (1)

In introducing the contents of this Chapter, we will first present a picture of the Spanish Health Care system. In fact, while much of the work that follows may be of interest to other countries, attention is here focussed on the study of the specific budgetary policies for controlling health care expenditure in the Spanish context. Thus we will examine the system of resource allocation from the Central Health Authority mainly in the hospital sector - our research is limited to this area - and its potential contribution to the efficiency motivation in the allocation of health resources in the sector.

1.2. THE GROWTH OF HEALTH CARE EXPENDITURE IN SPAIN AND THE RATIONALITY FOR ITS CONTROL

As it can be seen in Table I.1, health care expenditure has increased rapidly in Spain over the course of the last decade. This has been the case although it is not clear that today health policy objectives have come any closer to being achieved as a result of this growth.
For instance, the observed levels of geographical inequalities in health care provision have not diminished (see Table I.2.) and at the same time the cost-quality relationship achieved is still seriously questioned.[2]

TABLE I.1

Social Sec. Health Care Expenditure [a]

Year	Public Health Care Expenditure -1-	%	Pharmeceutical Expenditure -2-	%	Govern. Public Expendit. (*)-3-	%	GDP -4-	%	Implicit price deflator (**)	Gral. index GDP (mkt.pr.)	Ratios 1/3	Ratios 1/4
1967	20.78	-	10.57	-	213.04	-	1817.8	-	72.50	72.50	9.7	1.143
68	25.55	23	12.86	27	238.68	12	2037.5	12	78.63	89.69	10.7	1.253
69	32.65	28	18.44	37	275.46	16	2317.0	14	55.85	93.61	11.8	1.409
70	44.06	35	23.43	27	307.24	11	2576.2	11	100.00	100.00	14.3	17.10
71	56.74	29	30.73	31	374.48	22	2920.0	13	112.35	108.00	15.2	1.943
72	76.33	35	37.46	22	418.40	12	3432.3	17	127.64	117.39	18.2	2.223
73	105.09	38	49.28	32	483.44	15	4139.7	21	143.00	131.27	21.7	2.538
74	133.70	27	57.21	16	589.14	22	5102.0	23	167.90	153.03	22.7	2.620
75	177.11	32	68.05	19	689.47	27	6018.3	18	185.32	178.56	25.7	2.942
76	221.93	25	81.36	20	824.88	20	7234.2	20	225.19	208.36	26.9	3.067
77	319.65	44	94.87	17	1068.12	29	9178.4	27	273.77	255.92	29.4	3.482
78	410.78	29	104.67	10	1377.87	29	11332.8	23	322.29	307.10	29.9	3.624
79	441.86	8	105.11	0.4	1668.89	21	13226.6	17	370.47	358.52	26.5	3.340
80	500.96	13	116.39	11	2047.23	23	14932.1	13	415.41	404.75	24.5	3.354
81	591.07	12	136.55	17	2360.45	15	16918.0	13	474.90	458.58	25.0	3.493
82	673.82	14	161.12	18	2902.30	23	19736.0	17	550.88	521.41	23.3	3.414
83	844.80	25	164.96	1.	3460.00	12	22778.0	15	628.01	582.291	24.4	3.709
84[b]	875.22	4	149.95	-9.	3638.40	5	25934.7	13	691.44	651.72	24.1	3.375

(*) - PE excludes here transfers to other Public Administrations.
(**) - Index of consumption prices of Health and Medical Services from the Spanish National Accounts.
(Soc. Sec. Health Care Expenditure does include pharmeceutical expenditure).
a -Soc. Sec. Health Care Expenditure represented in 1980 and 90.3% of total Public Health Care Expenditure and 68.3 % of Total Health Care Expenditure according to the Spanish National Accounts.
b - Estimated figures.

TABLE 1.2

Provincial inequalities in Health Care provision (1978-1979)

	Social Security coverage (%)	Total no. of beds per thousand population	Total no. of physicians (per thousand population)	Life expectancy (years)	Morbidity rate (per thousand)	Child mortality rate (per thousand)
Max.	93	8.22	3.01	73.8	9.74	22.68
Min.	68	4.66	1.29	67.7	5.83	13.46
Max./Min.	1.37	1.76	2.33	1.08	1.67	1.68

Source: INSALUD and own calculations.

An analysis of the recent growth in health care expenditure illustrates that this has been one of the most dynamic items in public expenditure in Spain during the last two decades (1967-79). This past record, and the persistent claim for additional resources in the present age of financial restraint, is perhaps at the source of the general concern about health care expenditure and the interest in various features surrounding its control.

Moreover, the following aspects give rise to our interest from a pure public economics point of view:

(i) The nature of the output "health", and the internal characteristics of the health care industry in general (and the character of the technological advances in particular) which appear to be mixed with "output" and are usually mistaken for some of the other quality related aspects of its provision.

(ii) The way in which the public provision of health care is actually organized in Spain on territorial and institutional grounds (as described in the following pages).

Among the factors that contribute to (i) and (ii), we would point out here:

a) the nature of the output provided, with its fundamental aspects of a public good nature(3), and its income elasticity of demand generally greater than one.

In fact, health care possesses characteristics of 'merit goods' based on the view that a citizen has an analienable right to health, but features other aspects:

- its contribution to the economic productivity and community welfare;
- its provision in a context of market failure, given the degree of externalities, lack of consumer information, etc;
- its ultimate effects, strongly related to distribution goals.

b) The characteristics of the output produced, with some of the features of the 'service' activities, such as:

- its relatively labour intensive nature and its 'quality' proxies;
- its limited range of variation in the elasticity of substitution between labour and capital inputs: generally technical innovation tends to move both factors in the same direction;
- its inclusion in the category of service activities which manifest a general contemporaneous flow of consumption and provision.

c) Its actual provision through decentralized units under the domain of professional influences, with the supply influenced by demand and marked by the presence of vested monopolistic interests in the sector. Combined they make for frequent assertions about behavioural inefficiency, unnecessary utilization, treatments of dubious value and waste.

In order to put our principal interest in the appropriate context, we will provide in the following section an introductory picture of the Spanish Health Care system.

1.3. AN OVERALL DESCRIPTION OF THE SPANISH HEALTH CARE SECTOR.

The Spanish Health Care Sector is predominantly publicly financed with the public sector accounting for a little over 70% of total health care consumption.

TABLE 1.3

A picture of the Spanish Health Care Sector

Total Health Care provision

- from which public consumption on Health services represent. . . 72%
  - of which Social Security institutions 60%
  - of which Central Administration (Army and Education Departments). 6%
  - of which Local Public Sector 6%

Social Security institutions

- Revenues financed by employers and employees contribution .. . .85% (this component is being reduced)
- Revenues financed from direct State transfers (increasing) . . .15%
- Expenditure on Health Care .30%
  - from which INSALUD - Central Authority represents 87%

Hospital Health Care provision

- P S involvement in total inpatient care (number of beds) 64%
  - from which INSALUD owns directly 42%
  - and under private-public agreement 22%
- Social Security Health care expenditure on inpatient care (over total) 54%
  - from which INSALUD own hospitals. 40%
  - and under private-public arangement 14%
- Population Coverage (total) 85%
  - in terms of inpatient cases (over total) 73%

Hospital Care Units

- "Health Residences": standard health care across the country
  involve 50% of total INSALUD Hospital expenditure
  imply 53% of total INSALUD own beds
- "Health Cities": more sophisticated health care at a regional level
- "National Centres": some few specialized hospitals (Tuberculosis,etc.)

Source: Spanish National Accounts; National Institute of Statistics (I
and INSALUD data.

Some International comparisons (1978)

(1978)	Public Expenditure	Total Health Consumption	Public Health Consumption	Private Health Consumption
EE.UU	33.1	8.62	1.14	7.48
France	45.2	7.10	-	-
Italy	46.1	5.75	3.30	2.45
Denmark	50.6	6.70	-	-
U.K.	43.7	5.10	4.50	0.60
Spain	29.3	5.20	4.00	1.20

Source: National Accounts O E C D countries (1978).

From Table I.3 it can be seen that the Spanish Health Care sector is dominated largely by the hospital sector. In fact, hospital expenditure absorbs a large proportion of public and private health care spending and, as in many other industrilized countries, expenditure on inpatient care has been the fastest growing component of total health care expenditure over the last ten to twenty years. Its expansionary evolution in Spain can be seen in Table I.4.

TABLE I.4

The evolution of Hospital Health Care Services in Spain (1963-1981)

year	number of centres	number of beds	number of personel	ratio health personel/beds	number of physicians	number of nurses	number of ancilary workers
1963	1561	141259	58616	0.41	14416	25370	18830
1971	1332	157150	117123	0.74	2387	47989	45347
1976	1229	194097	214881	1.11	40939	104465	67255
1981*	1054	193895	278582	1.44	52185	140382	81912

year	ratio beds/ physcicians	ratio beds/ nurses
1963	8.8	5.2
1971	6.6	3.3
1976	4.7	1.9
1981*	3.7	1.4

Source: National Institute of Statistics (INE) and own composition.

* - Last year available ("Encuesta de Establecimientos sanitarios en régimen de internado")

- The hospital sector.

The hospital sector represents half of total Social Security Health Care exepnditure. Public sector involvement in inpatient care represents almost two thirds of the existing beds, three quarters of which are owned by the National Institute of Health - (INSALUD) (4) , the rest being managed under agreements (or "conciertos") with other hospitals, privately owned or dependent upon other public institutions. The private sector is thus reasonably prominent in a system of health care provision predominantly of the public type. Indeed, the private sector accounts almost for 30% of total health care consumption, a little over 45% of the hospitals, and about 17% of all beds and 16% of the total manpower in the sector.

In short, the public and private sectors co-exist in the Spanish Health Care system. One remarkable behavioural feature of this structure is the phenomenon of "pluriempleo": i.e. Spanish physicians can hold appointments simultaneously in the public and private sectors. This is not unusual, and holds true for most physicians in Spain. Statistics from INSALUD (see Table 5) indicate that Spanish doctors hold on average 2.4 appointments, 1.3 of which are in the public sector. In addition, a recent survey showed that, on average, 62.1% of doctors have more than one job in the health care sector. (5)

TABLE 1.5

The Distribution of Spanish physicians according to the place of work and working time (hours/day).

type of job	Proportion of physicians	Working time per physician	Number of jobs	Percentage physicians
Social Security	81%	4.3	one	35.8%
Private practice	58%	3.4	two	34.7%
Private insurance	42%	2.2	three	19.7%
Public Medical Serv.	24%	3.0	four	6.1%
Other public institutions	18%	4.3	five or more	1.6%
Other private institutions	17%	3.0	(no answer)	2.1%
	240%	8.4		

Source: "Investigación sobre la Asistencia Farmacéutica en España", INP 1977.

This phenomenon of so called "pluriempleo" has never been fully challenged (at least until the present proposed reforms)[6] by the public Health Authorities. This aspect of the Spanish Health Care Sector - whether it really is a cause of the relative importance of the private sector or whether the consequence of its existence are open to debate - will be subsequently considered when we examine the rationalisation of the health care sector in Spain, but it must be borne in mind in the ensuing discussion concerning hospital efficiency.

In addition to this behavioural aspect of the health care system, some other aspects of the hospital organization and funding of the system need to be considered, which may prove to be relevant to the management of the whole sector. These are explored below.

1.3.1. The Hospital Health Care Sector in Spain: its organisation and funding.

The organisation of the hospital public sector in Spain can be described by the following stylised facts:

(i) Its organisation is strongly centralized not otherwise well supported by an efficient hierarchical network for its management.[7] Indeed, INSALUD is a highly centralized Institute. Rules are executed at the provincial and institutional level by politically responsible bureaucrats (e.g. Provincial Delegates and Hospital Directors) and under a hierarchical set of formal authorisations and innocuous controls. However, political changes have traditionally interfered with the running of its structures. Financial policies depend separately on the funding Authority (the Ministry of Labour and Social Security) and the managing Authority (the Ministry of Health and Consumption).[8]

(ii) Obsolete regulation rules the internal organization of the hospitals.[9] Most of these rules refer to the nineteen sixties and they reflect an authoritarian regime and a non-participative organization. In addition, the workings of its structures are hindered by corporatist attitudes, facing a more flexible management. This regulation gives, moreover, absolute predominance to the self-interested 'agents' in the sector with regard to the decentralized stages of the decision-making process.

(iii) In more general terms, eligibility for health care coverage of the Spanish population is still restricted.[10] In addition, as explained before, the 'compatibility' principle between the private and medical practices enjoys general acceptance amongst Spanish physicians.

(iv) Some changes have been proposed during the last year (1983) in accordance with the desire to implement a National Health Service type of health care provision. These system changes try to tackle in the former field, the obsolete regulation concerning working conditions in the sector by allocating greater autonomy at the regional and institutional level, to face the singularity of the semi-administrative contractual nature of the physicians' working status, and to reinforce greater discipline in the fulfilment of their contractual duties. These changes are being assisted by rising unemployment amongst young medical professionals, which seems to have broken the formerly compact corporative behaviour of the medical organization - with respect, in particular, to the existing level of "pluriempleo" amongst some professionals in the sector.

(v) In addition to the organizational-behavioural factors described, the nature of funding in the hospital sphere has supported the absence of concern for a more cost-conscious management of the resources in the sector. In fact, as we will see in the next sub-section, the current method of revenue allocation of resources from the Central Health Authority to the hospital units has not provided an explicit and objective financial constraint with which questions of operational efficiency in the hospital sector should be in principle confronted.

1.3.2. Outline of the revenue allocation process in the Health Care Sector in general.

The mechanism of revenue allocation to the decentralized health care units (the providers of services) can be analysed in three stages:

a) in relation to the determination of the global sum of expenditure available to the Health sector in general;

b) in relation to the distribution of monies to each health care subsector in particular (say curative vs. preventive care, hospital vs. ambulatory care, inpatient vs. outpatient treatment); and, finally, c) in relation to the distribution amongst the specific units charged with the provision of that particular form of health care.

The first decision is usually faced at the Governmental level after inter-ministerial negotiation between the Ministry of Finance and the Ministry of Health (i.e. equivalent to the role of the British 'Public Expenditure Survey Committee'), taking into account the aggregate expenditure figures of the joint decisions from all Departments. This planning or forecasting exercise can be carried out in accordance with some explicit factor (existing public pressures, political lobbies, etc.), or fixed guidelines (such as the percentage rise of the GDP). The control problem here can obviously be located in the need, finally, to adjust the total amount of resources available to the set of aggregate decentralized claims by the subsidiary units charged with service provision.

The second decision is solved at the Department of Health level, and is influenced by the politics of the Party Manifesto, so that health care system as a whole becomes in some way a reflection of the philosophy of the party in power.

Finally, the _third_ step in the decision-making process, and the one in which we are mostly interested here, refers to the way in which the previously settled amount is to be distributed amongst the decentralized units-providers of the services within each particular subsector of health care.

In countries with a certain degree of decentralization, an initial distribution in favour of the States or Regional Authorities usually precedes the final allocation of resources, although the criteria for the manner of assignment has to be determined in principle in a fashion consistent with the implementation of globally-stated planning policies.

Demographic variables - such as population, number of persons affilliated, or beneficiaries in the region, are mostly considered in this intermediate stage, with account taken of some territorial or social inequality factors.(12) However, in practice, other implicit criteria may also be present, such as those based on pure incrementalistic patterns, political lobbies, public minimization of complaints, or some other "passing money around" element - which offer a generous treatment one year and restrictive the next.

Whenever no explicit criteria are set for the latter re-distribution, the previous years' expenditure or what has already been actually spent in the current year take on an unsystematic nature: one which in reality leads to some form of cost-based re-imbursement. This may fail to translate the financial environment of resource limitation into the internal behaviour of health organizations in order to give the agents an incentive for improving performance over time.

Most of these features will be found in the system of budget 'coverage' for hospital expenses in Spain.

1.4. THE SYSTEM OF BUDGET COVERAGE FOR HOSPITAL EXPENSES IN SPAIN.

The hospital sector in Spain is centrally funded. No decentralized stages for the planning and allocation of health resources on a regional basis exist (although this system is bound to disappear in the future as the decentralization process becomes more wide spread).[13]

In this context, the decentralised hospital unit network has been managed under a type of bureaucratic pattern, typical of a hierarchical administration. INSALUD Central Authority, built as a separate institute for management within the Department of Health and Consumption, has ruled the provision of inpatient care with a set of instructions, reinforced by the support of the Provincial Delegated Authorities.

The central funding of the decentralised units encharged of inpatient care provision is set up, at least theoretically, within an interacting framework, upon which a prospective system of revenue allocation of their budgets is meant to be based. In actual fact, hospital budget requests follow in outline an administrative sequence from preliminary proceedings, central approval and decentralized execution, up to the final budget settlement.

However, as we will see in the next pages, the results have come to be based, in practice, on a retrospective system, or cost-based reimbursement, on the basis of the reported actual expenses, which ultimately results in already-incurred expenditure being reimbursed.

1.4.1. The INSALUD funding of the public hospitals.

For the sake of a brief description of the INSALUD Central funding system, three different stages can be distinguished regarding the final determination of the allocation of revenue for hospital expenses.

1. The determination of the wage settlement and hence the hospital wage bill;
2. the allocation of the overall hospital revenues; and,
3. the final settlement of the executed budget, as the actual expression of the financial constraint.

1. The wage settlement

Since almost 67% of total hospital expenses are manpower costs, the prior wage settlement is bound to have a crucial importance.

This step could be expected to be framed, in principle, as a bargaining exercise under the postulated terms of a bilateral monopoly between INSALUD (the main employer in the hospital sector) and the representative organisations of the hospital staff: mainly the Medical Association and the respective branches of the two biggest Spanish Trade Unions.

For historical and institutional reasons not pursued further, no explicit bargaining has occurred in the sector for these purposes until recent years.(14) In fact, hospital agents have accepted rather passively an unfavourable central wage determination which has left wage increases far behind the general rise in the cost of living (as shown in the second chapter).(15) In order to understand the passive attitude of physicians, until recently absent from a bargaining table, we have to consider the whole system of budget reimbursement. In fact, the system has allowed, ultimately, a strategic 'retreat' from the physicians' main interest towards the domain of influence of other variables which have also decisively affected real wages, in precisely those fields of hospital activity where central audits and monitoring controls have been especially weak: (16) i.e. the real time devoted to public activity and its compatibility with private sector practice. This point will be further analysed in the next Chapter.

As a result, this type of income policy approach to hospital expenditure restraint has achieved some short term results, relatively satisfactory as far as the Central Board of Hospitals is concerned, although it has left the real problems of the sector unfaced.

In sum, the process for the determination of the wage bill starts with the settlement at the beginning of each budgetary year by the Health Department of the level of wages and other related payments per each hospital staff category. Wage rates are mainly attached to jobs rather than to workers, with no relation in principle to the external wage level in the private sector nor the agents' levels of performance, but to a complex set of administrative procedures, based mostly on job allocation, which is to some extent dependent on seniority in the buraucratic hierarchy and experience, together with some other formal qualifications.

If one assumes significant job skills are acquired simply by experience on the job, this solution in favour of a wage hierarchy can, in theory, reduce potential bargaining costs, particularly in those sectors with high information assymmetries.

However, it also provides a constant perverse incentive structure, since it is based on job promotion for life (mostly achieved by not "blotting ones copybook"), and, in simple terms, this may perhaps explain the general lack of incentives for efficient performance over time.

2. The allocation of hospital revenues.

A second step in the process of budget reimbursement is given by the allocation of the overall revenue to the hospital units, once the wage bill is known.

Basic proceedings start with requests put forward by each Hospital Board, and usually submitted six months in advance of the beginning of the fiscal year, to INSALUD - Provincial Authorities. These 'bids' or 'Anteproyectos' are budget forecasts statutorily based on the probable expenditure in the current year, adjusted for cost increases as experienced during recent years, or as anticipated from expected hospital activity. In reality, they contain much pure guesswork.

For each item of expenditure, the forecasted figure has to be accompanied by the 'probable expenditure in the present year'. It is then easy to understand how these amounts are going to be seen as the basis for an incrementalistic pattern instead of the result of a full expenditure reappraisal.

Formally, the Hospital Board has to assess whether the current evolution of hospital expenses will continue in a similar direction and volume, allocating the expected cost increases correspondingly.

A statement of proposals (or'Memoria') must justify these considerations for each of the so called 'spending elements', e.g. in terms of the relative price increases, output or work-load changes, or any related quality improvement with respect to the previous year's amounts. But, in practice, most of these 'Memorias' reflect no more than a general guess about the fore-casted evolution of past expenditures, with little or no justification concerning their increases.

In general terms the forecasting exercise has two parts: one relates to current expenditure (or Economic Plan) and one to new investment (or Plan of Needs). Standarized procedures for the presentation of information (up to 13 Appendices) are requested for these purposes. They are essentially constructed according to certain administrative formalities rather than forming a specific functional design for budget evaluation objectives. A set of guidelines are issued annually for this, providing codes, definitions, delimiting the scope of the particular items to be included, certain criteria for arithmetical calculations (e.g. fixing the coefficients and estimated values for the depreciation policies), and pointing out priorities for the replacement of investments, rebuilding and equipment.(17)

Manpower expenses are determined according to the staff levels approved by the Health Department and the announced wage settlement. The rest of the budgetary items (medical and non-medical inputs) are left to the hospital's discretion. Hospital expenditure on outpatient services are formally separated from overall hospital spending. Since no clear-cut criteria exist for ascertaining this service, it is left to each hospital administrator to establish the size of this service by the use of ratios such as the share of manpower time devoted to outpatient activities, inputed utilization of certain material inputs to specific departments, etc.. This then introduces the question of data reliability for analysis in this field.

In summary, the budget request is formally adopted in May by the Governing Board of the Hospital and transmitted by the end of May at the latest to the Provincial Delegation of INSALUD for the initial proceedings. From this moment, the rest of the budget iteration involves no more than some administrative controls at the provincial level. Although INSALUD Provincial Boards are in theory charged with the responsibility of examining those assessed criteria, forecasting the evolution of current expenditure (e.g. with regard to the utilization of human and material inputs) and checking whether the bases of the distribution of planned hospital expenditure are consistent with respect to the priorities assigned, nothing seems, up to the present, to validate the fulfilment of these duties. In fact, the Provincial Services of General Accountancy (or 'Intervención'), with few resources and lacking explicit standards for evaluation purposes, have supervised the fulfillment of the budget instructions through a rather cursory glance, rubber-stamping the set of budget proposals and passing them up to the INSALUD Provincial Director for formal approval and, to some extent, further political support for the rest of the budget iteration. In this sense, it should be noted that until quite recently, Provincial Directors, although appointed by the Central INSALUD Authority, have more or less been direct representatives of the predominant sectorial interests in the province. In this way, they have acted as supporters of any spending proposal raised from the hospital health units in the province.

Once the 'Anteproyectos' are signed by the INSALUD Provincial Authority, they become budget projections and are sent to the Sub-General Direction of Administration of INSALUD, and 'Service of Budget and Expenditure Evaluation'. Subsequently, they are considered for approval.

Some minor legal administrative controls and statistical accountancy procedures have been employed up to this stage for purposes of expenditure control, prior to the final aggregation, in order to confront the budget proposals with the overall figures available for spending in the Hospital Health Care Sector. In this field, the discretion of INSALUD Authorities has been almost complete. Until recent years (1978-79) all these expenditure proposals seem to have been approved quasi-automatically. However, increasing concern about health expenditure growth has, more recently, made that approval conditional upon reaching certain targets, mainly linked to the global growth of the economy[18] or to some other general 'cash limit' stated for anti-inflationary purposes.

The absence of a more complete budget appraisal and expenditure evaluation on which to link those general restraints might well have led to chaos. But as we explain in the next few paragraphs, no external disruption has apparently occurred. In fact, the method has, overall, sufficient mechanisms to weaken any direct effect of those administratively set "prospective" constraints on the hospital activity.

3. The budget settlement

The approval of the budget has not been felt as the final expression of the financial constraint according to which the hospital activity has to be carried out.

Although there have been constant warnings from the INSALUD Direction Board that 'no expenditure can be made above the figures approved and credited for each specific budgetary item' and that 'prudence and balance have to be paramount in conducting the budget execution' [19], the overrunning of the budget has been a rather common practice in most Spanish hospitals.

No specific pattern can be demonstrated at this stage in the sequential process of budget submission, approval and execution, with reference to their initial and final outcomes in terms of the level of reduction, degree of overrun and the future use of any of the former budget bases for new planned expenditure.

However, some tentative hypotheses will be formulated in the course of this research about the nature of this incrementalistic pattern.

1.4.2. <u>Tentative changes and the need of the hospital budgetary reform.</u>

A proof of the fragile situation which involves the budget procedures today in Spain can be seen in the continuous changes in the technical formalities required for the elaboration of the budget requests, rather than the establishment of a more solid basis in which a consistent evaluation policy should be permanently grounded.

For example, following some tentative reforms in 1980 which took steps towards Program Budgeting, INSALUD Authorities today request, in an optional way, three different levels of expenditure forecasts: for each elementary "program unit", a 'minimum', 'current' and 'additional' level of effort can be set down according to the planned expenditure targets.[20]

The <u>minimum</u> level is defined as the amount "strictly indispensible" to run the hospital programs as statutorily regulated; <u>normal</u> levels allow for additional amounts, once it is proved the possibility of "reaching a more favourable combination of means for a greater output"; and, finally, the <u>maximum</u> level requires the consequences of the additional allocation of resources in terms of the quantity and quality effects achieved with respect to the initial levels of provision to be shown. Corresponding to this approach, a table of priorities amongst programs and levels of effort has to be put forward in a consistent fashion with the targets already stated in the "Memorandum of Intent". But given the nature of the exercise, to the best of our knowledge, few Hospital Boards seem to have opted for this system. They argue, in fact, that "the level of effort required for the completion of a request for each of those budget levels is very high and the expected probability assigned today to something other than the allocation of the minimum amounts does not make it worth-while".[21]

In the absence of a complete reform of the hospital budgetary methods, the system runs under the influence of the pressures exerted by hospital Directors on the basis of current levels of expenditure. In theory, in order initially to prevent such results, Hospital Boards are ordered to follow the budget execution monthly. But if under"exceptional circumstances", problems of overspending appear, the Board is entitled, "exceptionally", to present, once for all, additional financial proposals. Although the justification of the causes involved must be "clear, concise and sufficient" and it is reminded that "no pressumption should exist of automatical approval", this stipulation recognizes and provides a formal procedure to contend with common budgetary overruns.

Lobbies and public pressures are thereafter exercised for the full coverage of the over-expenses incurred and for the recognition of the current expenditure levels for future requests. The final outcome depends on subjective factors such as the forecasting-mistakes of the hospital administrators, the personal abilities of the hospital directors to achieve additional funds, the risk-aversiveness of the hospital staff when commiting additional amount to those initially budgeted, the influencial power of some local Authorities, etc.

In conclusion, the system comes to operate, in practice, as a retrospective reimbursement. In this context it is hard to think of any general motivation for a cost-conscious management, other than that derived from certain variables "ad hoc".

The efficiency implications of the process will be analysed in the next section.

1.5. EFFICIENCY IMPLICATIONS OF THE CURRENT SYSTEM AND ALTERNATIVES.

In brief, this description of the allocation of revenues to the Spanish public hospitals illustrates, in our view, the core of the problems involved in the present system of health care delivery. That is, the absence of a proper budget restraint from which operational efficiency and a cost-conscious management of resources could be expected.

In other words, the way in which the revenue allocation process operates in the Spanish hospital sector would explain the lack of systematic incentives for an efficient performance in the sector.

In this context, the research described in subsequent chapters is explicitly orientated towards answering the following relevant questions with regard to the analysis of the operational characteristics of the hospital health care sector in Spain, with the ultimate aim of providing an outline for its reform. They are:

a) Why should the patient covered by public assistance be concerned about utilization costs instead of the "best" quality care?

b) What reasons has the physician to select a less costly alternative to what he considers a proper (but expensive) mode of hospital treatment?

c) Why should a hospital be managed cost-consciously considering its revenues are obtained from a central reimburser who pays without regard to costs?

1.5.1. The pressures for change in the budget financial arrangements and existing alternatives.

We have seen so far that as the budget revenue allocation methods have been operated in the past in Spain, no effort for efficiency can be expected from rational self-interested agents in the sector.

However, the current cost-based retrospective reimbursement system has been distorted during the last five years. In fact, external pressures from the general economic situation have imposed a tough financial restraint on public expenditure and on health care expenditure in particular. This is, moreover, being pursued at a time when the political aim is that of extending the coverage of health care provision to the entire Spanish population by the construction of the Spanish National Health Service. There seems little doubt that in order to succeed on both fronts, a serious effort for expenditure rationalisation will have to be made.[22]

This effort for expenditure rationalization that accompanies policies of financial restraint is, by no means, a feature unique to the Spanish Health Care System. Identical proposals can be found in several cost-containment policies put forward by different countries.

In general terms, these policies try to tackle any of the following three elements involved in cost increases in health care:

(i) consumers' behaviour. A common view is that the presence of consumers' full insurance and the absence of prices are the "causa generandi" in a demand-pushing model leading to cost increases. In this context, policies which insert financial motivation - together, in some cases, with others that influence the supply side - and consumer education for a behavioural change for a more appropriate use of the Health Care System, have been proposed under the form of cost-sharing propositions (coinsurance, charges and deductibles) of the type of the United States 'Consumer Choice Health Plan' of 1978.

(ii) Demand induced supply of the agents providing health care. Here the concern is with the absence of financial motivation at the micro levels of the decision-making process and the lack of training of health agents as effective managers of health resources as the main factor in a cost-pulling model of cost increases. The cost-containment policies proposed usually try to relate performance, and the allocation of overall or specific responsibility in resource utilization, with a fixed financial constraint, and they introduce some type of incentive behavioural change of health care agents.

We can note here the present policies in the U.K. on departmental and clinical budgeting, and the implementation of basic or intensive training on the economics of health resources. Other policies in this field try to control directly the utilization of the inputs available: for instance, labour policies on manpower increases of the French type and programs of wage control (US 'Economic Stabilization Program', or the Spanish type of incomes policy), non-labour inputs and major investment (planning equipment and the US 'Certificate of Needs', control of technological innovations of the West Germany type); and with reference to some other factors in an operationally-related approach, regarding, for instance, input utilization, appropriateness of admissions and continued stay through Utilization Review Policies (such as the US Professional Standard Review Organization,Second Surgical Opinion Programs, etc.)

(iii) finally, and more generally, expenditure rationalization policies attempt to tackle the financial environment in which Health Care delivery is organized and, specifically, the particular system of provision. These policies try to implement at the institutional level some major changes in the way in which resources are allocated; in the direction of a clearer revenue planning, prospective funding or any other initially commited reimbursement policy. They pursue ultimately the goal of translating the external financial constraints into the institutional performance by forcing the implementation of managerial techniques (whichever they are at the disaggregated level) in the decision-making process. In this vein, we can point out some revenue allocation strategies such as vertical budgeting (i.e. pre-paid group practices, pre-fixed rate settings and target ceilings) and horizontal budgeting methods (from intra-group performance comparisons); in both cases, at clinical, departmental or global budgetary levels, among others.

It is still difficult at this stage to offer a full validity test with regard to the mechanisms through which the former factors are supposed to operate in each of those policies. Probably, these sort of tests are impossible if our aim is to derive extrapolative conclusions to apply to systems other than those for which they were designed.

In the next section, we will examine how the general financial restraint on public expenditure has been translated in practice into the Spanish Hospital sector, which elements of rationality have been put forward, how this restraint has operated up to the present, which effects can be forecasted from it, and which alternative methods are most consistent with the aims for which these policies are designed.

1.5.2. Current policies for financial restraint in the Spanish hospital sector and the efficiency motivation target.

The historical record illustrated in Table 1 shows that Health care expenditure has been expanding rapidly over and above the general evolution of the other economic indicators up to the end of the last decade. In fact, the total consumption of health services as a proportion of the GDP rose from 3.76 in 1976 to 5.93 in 1978. However, 1978 seems to point to the end of this expansionary age, and the beginning of the imposition of strong financial restraints in hospital sector expansion.

In Spain, macroeconomic constraints on public expenditure have been translated into the health care field mostly through the implementation of general cash limits and "across-the-board" increases with respect to past expenditure amounts.

These financial policies for expenditure restraint have been justified initially by "efficiency principles". This rationale has been adopted from the assumption that the resulting imposed cost-savings must imply parellel improvements in resource utilization. By holding "output" constant, this leads to a rather automatic result in "efficiency terms", which follows from the new balance achieved between expenditure and the constrained revenue reimbursed.[23]

Whether these policies imposing a general "across-the-board" increase can actually help achieve improvements in efficiency is a different matter.

Firstly, what is to be understood for efficiency in this context is not always clear. If the former conventional approach is considered - roughly speaking that which consists of doing "the same but cheaper" - it needs to be proved that the same outcome is actually achieved, and for this purpose we would be best advised to account for the varying circumstances from which this result can be plausibly expected. If the "same" refers to what someone else does, it needs to be proved that similar circumstances are involved in the comparison, especially if, from the alternative provisions compared, we expect identical performance.[24] Although the main attention has shifted from the assumption rather than the proof of such results,these questions still remain relatively unexplored in the literature.

Secondly, what this policy for expenditure restraint in fact does is to validate the past level of hospital expenses, since it comes to use them as the bases for future projections. We have seen that those bases are the result of factors such a budget revenue allocation grounded in the current retrospective system. Therefore, this is a very weak basis for expenditure rationalisation, since it reflects the consequences of a perverse set of institutional arrangements. Past levels of expenditure cannot provide a sound present base for an expenditure rationalisation policy projected to the future.

Thirdly, the implementation of a general cash-limit with "across-the-board" increases tends to deal with potential "unequals" as "equals". In this sense, it penalizes former cost-conscious spenders and rewards high spenders, which are left relatively better-off after the implementation of the policy, given the different basis of application of those general increases.

As a result, this policy introduces questions of comparative justice and a certain dislocation effect in favouring the return to some strategic budgetary games in order that the initially prospective financial constraints can be evaded.

Therefore, whereas past reimbursement policies have not provided any effective incentive motivation for a cost-conscious management of health resources, present policies for expenditure restraint in the hospital health care field fail equally in addressing the efficiency problems in whose name they are usually justified. In short, "across-the-board" increases based on self-reported actual costs, in a context of inadequate monitoring and expenditure evaluation, cannot intrinsically contribute to the expenditure-rationalization targets, required if a serious and constant effort for improvement in the allocation of public resources is to be pursued in the hospital health care sector. Moreover, this policy introduces a set of distortions against a more rational basis for the development of hospital activity.

1.6. SUMMARY OF THE CHAPTER AND CONCLUSION

This first Chapter has been devoted to an introduction of the system of health care provision in Spain and to focussing attention on the objectives of this thesis. For this purpose, at this initial stage, we have pointed out a set of behavioural and organizational factors related to the provision of inpatient care in Spain, which are postulated to be at the root of the poor performance of the sector.

As far as the former is concerned, we have focussed on the influence of the budget funding method on the allocation of hospital revenues, emphasizing its efficiency implications from a hospital-operational view. We could conclude from this that, in the presence of a cost-based reimbursement system, no incentive for a cost-conscious management of health resources can appear in any stage of the decision-making process involved in the utilization of the hospital resources. On the contrary, 'system inefficiency' becomes established in hospital performance and in current levels of expenditure, since they collect the record of the operational efficiency from the management of the hospital resources over time.

As we will analyse in the Second Chapter, in designing alternative budgetary techniques as a means of promoting efficiency in the hospital sector we need to account not only for a different basis to that which results from the evolution of past expenditure, but also for the issue of "variable productivity". That is, we need to account for the existing levels of differential performance within the hospital sector in order that we may provide an efficiency-related benchmark for reimbursement.

On both grounds, an alternative budget policy will be constructed in this thesis, providing a more sensible extension of the global financial constraint imposed on the sector, and it is one that can be translated into the hospital decision-making process in a more cost-conscious management of health resources.

FOOTNOTES.

(1) Here we should distinguish between the arguments for control of Public expenditure and those propositions for a pure expenditure restraint, or its simple reduction in real terms, given a submission to some more wider anti-inflationary policy targets. (See for these purposes the Appendix to this first Chapter).

(2) See on this issue, "Seguridad Social". Papeles de Economía Española, no. 12/13. 1982 (page 442). Obra Social de la Confederación de Cajas de Ahorro; the so called "pink book" on Health Care in Spain: "La asistencia farmacéutica en España. Estudio Socio-económico sobre el conjunto de la asistencia sanitaria española". Instituto Nacional de Previsión, Ministerio de Trabajo, Madrid 1977; or the small booklet of critical reports collected under the title of the 'Black book of the Spanish Social Security': "Apuntes para un libro negro de la Seguridad Social" YUSTE GRIJALBA, Javier. Akal edit. 1982.

(3) See in this line of argument A. CULYER, A. MAYNARD and A. WILLIAMS in "Alternative Systems of Health Care Provision: An Essay on motes and beams" in M. Olson (ed) "A New Approach to the Economics of Health Care." A. E. I., Washington, 1981.

(4) Organically, the "Instituto Nacional de la Salud" is a body of management depending on the Department of Health, but funded by the Department of Labour and Social Security - since, as noted before, the Social Security Scheme is still the main financial funder of public health care provision in Spain (although an 'out of taxation' coverage is today being proposed).

(5) See further comments on the results of this survey in "Pluriempleo y conciencia de paro en la profesión médica." MARTIN LOPEZ Enrique. Revista Española de la Seguridad Social. 1982.

(6) In this sense, what has been called the new 'Law of incompatibilities' (Law 20/ 1982 of 9th June) views exclusively the cases of pluri-empleo within the Public Sector in which an identical professional possesses two jobs under identical time schedules (!).

(7) The National Institute of Health (INSALUD) was created in 1977 as a Board of Management of the Social Security Health Care. In most of its functions, it replaces the old "Instituto Nacional de Previsión", framed up to that date within the Department of Labour and Social Security.

The specific position of the INSALUD can be seen in the following organization tree.

National level:	Department of Health and Consumption	
	Subsecretary of Consumption	Subsecretary of Health
	Institutional Administration of Health	INSALUD - - General Board - Execut.Commissi
Provincial Level:	Territorial Delegations of the Department of Health and Consumption	Provincial Office - - - Execut. Commiss
Local Level:	Local Authorities.............	Local Health Authorit

(8) The setting of the Health Care Administration has suffered several changes since the implementation of the first Social Security public Schemes (1942). In fact, most of the public health Administration was initially structured as a branch of the Home affairs Department. Later, it was attached as a Section to the Department of Labour and Social Security. In 1977 it was reorganized within the Department of 'Labour, Health and Social Security', to become, finally, a new and separate Ministry of Health and Consumption (Dec. 1981). Only during the Spanish Second Republique (1931-1936) existed a modern Department of Health as such.

(9) In fact, the internal operational regulation (for the "Régimen, Gobierno y Servicios") of the Social Security Health Institutions was established in 1972.

At the beginning of the same decade, the working conditions of physicians were regulated ("Estatuto Jurídico del Personal Médico". 1971), of the Auxiliar Personnel (1973) and Ancillary workers (1971). All of these regulations result from the development of the Social Security Law of 1966.

(10) In this sense is formulated the first draft of the "Anteproyecto de Ley General de Sanidad", January 84, as published in '<u>Tribuna Médica</u>' no. 1018 (page 18-20).

This new law tries to introduce openly a National Health Service system (art 1.2) under clear influences of the British and Italian experiences.

(11) In this direction, the particular organization of the proposed National Health Service, according to the recently announced bill of Reform of the Hospital System (1984), is asserted.

With regard to the proposed changes, Pedro Sabando, Subsecretary of Health, declared to an Spanish magazine quite recently: "We are committed to have hospital physicians working exclusively in only one job. We understand however that for doing this we need to remunerate properly those professionals according to their status in our society" ('Cambio 16no. 595). The proposed 'Law of Hospital Reform' follows similar lines.

(12) See for this purpose the RAWP formula for England (1976) and the comments on its workings in the study of Townsend and Davidson (1982) on the Black Report (pages 146-151).

(13) Only the Catalan autonomous Community today enjoys a full decentralization of health services. However, the final generalization of the process for the rest of the Spanish regions is purportedly to be concluded around the middle of 1986.

(14) In principle, labour relations between INSALUD and the professionals in the public health care sector do not statutorily allow for explicit bargaining formulae for the determination of wages. In fact, only since 1982 has a certain bargaining procedure been set up for these effects, although as far as legal obligations go, its strength is one merely of moral pressure. According to the Social Security Law of 1966, it is an internal ruling of INSALUD which entails the final wage settlement, with or without the agreement between the parties.

(15) As we will further analyse the Second Chapter, the central wage settlement has imposed a tight approach to physician's wage increases which has left wages far behind the general rise in the cost of living. In contrast, the auxiliary personnel reached higher increases, although from relatively lower basic levels. Ancillary workers, on the other hand, achieved a closer adjustment to the general increase in the cost of living.

(16) In this way can be explained, for instance, the exaggerated growth of expenditure on overnight duties, which have been viewed by the physicians as a relatively easy way of complementing the basic wage level. This feature is widely recognized by the Administration (see "Medicina Amarga". 'Cambio 16' no.595).

(17) For instance, the set of instructions for the 1982 budget fixes the rate of depreciation at 2.5% for buildings (such as valued at 31-XII-1980 by the proper central services) and at 15% for equipment and clinical material over the value of the second year prior to that of the budget forecast.

(18) Without apparent justiciation, the set of instructions from the INSALUD Central Board for the 1978 'Plan of Needs' budget allowed for a general maximum increase of 15% with respect to last year amounts.

Otherwise it established that the total capital expenditure of 1979 should be equal to that approved for 1978 adjusted for the general price increase.

A general cash limit was openly put forward by the general Director of INSALUD for 1983 budgets. The purpose of limiting the growth of the health care expenditure to the general growth of the economy was also formally announced by the former Minister of Labour and Social Security (see "El Pais" 30th of January 1982).

(19) From INSALUD internal instructions to the Hospital Boards (1979); in the same direction, the 'Real Decreto Ley' no. 24/1982 29 Deciembre, Disposición Adicional 4ª.

(20) The differences given as examples in the INSALUD initial instructions for 1983 offered the following illustrative benchmarks; Max/minim. level = 1.15; Normal/minim. level = 1.09; and Max/normal level = 1.06.

(21) From different conversations with hospital administrators in Spain.

(22) Thus we can describe the present budget process in Spain: "The experience shows, on one hand, the systematic utilization of a coefficient with an homogeneous increase for all the cases ...., and, on the other hand, that it is the persistence of some bureaucrats which makes for a final settlement over the amounts stipulated. In consequence, the so called 'budgetary rationalization' is still today an abstract concept in Spain" (in Alfonso GONZALEZ FINET and Angel DIEZ GONZALEZ. "Racionalización Presupuestaria: Algunas reflexiones". Revista de Hacienda Pública Española, 1982.

(23) Very germane, in our view, to this purpose is the paper of P. M. JACKSON "The public expenditure cuts: Rationale and consequences". Institute of Fiscal Studies. 1979.

In the managerial applied field this approach can be seen in R. BANYARD "Efficiency savings in the NHS". Hospital and Health Services Review. Nov./Dec. 1982, or in D. POYNTON, R. ALLAN and M. RIDLEY "Managing in a time of financial restraint" in Public Finance and Accountancy, Oct. 1983.

(24) In the Spanish Health care context, see the communcation of A. DURAN "Comparando lo incomparable", presented to the first meeting of the Spanish Health Economists' Group, Barcelona 1980.

(25) It could however be argued that any form of tight overall financial restraint may reinforce the necessity to evaluate decisions by sharpening the debate about the need of choice. But this would not change the strength of our argument, since this form of constraint cannot sustain a long term effort for an efficiency improvement in the sector. In some way or another it is based on the actual performance and it does not intrinsically remove the perverse institutional incentives which surround the provision of the service.

APPENDIX.

The increasing concern about the 'opportunity cost' of Public Sector involvement on the Spanish economy.

Most of the literature supporting the need to impose a tighter restraint on public expenditure has been outlined from the ultimate argument that a rising 'opportunity cost' results from the Public sector involvement in the economic activity.

This assertion has been mostly derived from some over simplifield models (*) which postulate some type of economic interference between both private and public sectors in the economy.

The question of the 'rising opportunity cost' and the derived need to restrain the relative increase in Public sector activities has to be initially distinguished from that more general view of the need to control public expenditure. For this second proposition, theoretical support can be found from several economic views. In short, the neoliberal and monetarist approach develop those prepositions from some type of financial 'crowding-out' effect, whereas Keynesian and post-Keynesian theories view the need to control public expenditure as an important part of the management of the overall demand in the economy, or on the grounds of improving the productive basis of the general economic performance. Therefore 'control' , and not 'pure restraint', make for the relevance of the distinction, with regard to their consequences in terms of the policies to be instrumented.

It seems otherwise difficult to validate the specific mechanisms of the operation of those propositions on pure economic grounds and, therefore, any of the inferences derived on them about the potential rise of an opportunity cost from the relative 'overdimension' of the Spanish Health sector in particular.

(*) - In this sense, those models which derive a real or financial

Looking at the evidence available - the historical evolution of the Spanish public expenditure, the approach to its real size, its manning policies and their corresponding effects on the revenue side of the public sector budgets (e.g. taxation elasticities with respect to GDP, charging policies, etc.), as surveyed in the following pages - those assertions do not seem to be proved unambiguously; indeed such suggestions would actually run counter to most of the data we have. In fact, the Spanish Public Sector appears to be relatively small and takes a smaller share than could be predicted on population, per capita income basis and type of economic system amongst the OECD countries (TAIT, 1982) - always assuming that these factors alone could be considered satisfactory guidelines.

In explaining a potential rise of concern around some notion of 'opportunity cost' of the Public Sector involvement in the Spanish economy, we need to go back to some other fiscal-sociological factors which may have helped to engender this concern. Among these factors, we would probably find: (i) the implementation of some necessary taxation reforms in a period of economic decline (such as that of 1974, 1976 and 1980), facing mostly the set of unsatisfied social claims which have accompanied the new democratic process; and (ii) the argument of the traditional inefficiency of the Spanish bureaucracy.

Nevertheless, and independent of the empirical support which could be argued for each of the former assertions, this phenomenon is likely to persist in an age of general economic recession (such as the present) and hence these inferences about the relative size of the Public Sector will arise.

(...) crowding-out, either from the interference of two sectors with different 'tradeability characteristics' or their more or less 'industrial', 'marketable', 'progressive' or 'productive' components are illustrative (see for example R. BACON and W. ELTIS (1976) and, originally, W. BAUMOL (1966)).

Some available evidence.

The growth of Public expenditure in Spain has followed a broadly similar trend to that of other developing economies.

Historically, public expenditure in current prices during the XXth century can be recorded under three main stages:

a) A rather moderate expansion between 1900 and 1935, with an average increase of 6% although under some evolving fluctuations (particularly around the period of the First World War);

b) a more generalized rise after the Spanish Civil War, up to the beginning of the 'Stabilization Program' (1959), with an average rate of growth close to 13%; and, finally,

c) a more resolute increase, with rates of growth close to an average value of 17%, during the most recent period (1959-80).

In constant prices, this general view changes little, except during the First War period and the post Spanish Civil War period, with no increases in real terms given the more rapid general increase of prices.

The most relevant periods for our analytical purpose are the one which follows the 1957 Taxation Reform and lasts until 1974, and from the latter year to the present, since they accompany the main structural changes of the modern Spanish economy.

As it can be seen in Table 1A, total expenditure increased relatively faster in Spain than in most other European countries, although obviously from lower initial levels. Certainly, this does not say very much about the concerned 'opportunity costs' of the Public sector involvement in the economy, since no account is made about the levels of services provided and the extended benefits from such activity.

BEST COPY

AVAILABLE

Variable print quality

This picture may give then a first indication of a relative 'catching up' effort, raising revenue against the increasing public expenditure needs, particularly notorious in the period between 1972 and 1976. Data availability does not allow, however, for a comparable study for a more recent period.

However, the global evalution of Public Expenditure and Taxation ratios to G.D.P. can be pointed out:

TABLE 3-A

	Public Expenditure/G.D.P. (%)				Taxation/G.D.P.			
	1977	1978	1979	1980	1977	1978	1979	1980
Spain	27.5	29.3	30.5	32.4	36.5	27.1	28.3	28.6
U.K.	44.1	43.7	43.5	44.6	39.4	38.2	39.0	40.4
O.E.C.D.	37.7	38.2	38.6	39.4	35.0	35.1	35.9	36.8

Source: O.E.C.D. Economic Outlook no. 32 December 1982. Paris O.E.C.D.

The question of the real size of the Public expenditure growth has been raised recently by M. BECK (1976), accounting specifically for the notion of the relative price effect in public service activities. This effect may be defined here as the ratio between the average price increase of goods and services in the Public Sector over the price increase in the rest of the economy.

Our results obtained by adopting Becks's approach to the real growth of the Spanish public expenditure are presented in Table 4A. Contrary to his conclusion with respect to the most developed economies (which indicates for the latest years recorded an inflexion point in their real expenditure growth -hypothesis which was refined later (BECK, 1979) in assessing a real decline for the final consumption although with a real increase for the transfers, once a differential deflation is considered) - the Spanish case does not seem to support either a real decrease for the total figure nor for the final consumption. This result is otherwise more similar to PLUTA's findings (1981) for most of the developing countries studied by him.

TABLE 4A

Nominal and real perspective of Public Expenditure growth in Spain

Total government expenditure (including transfers from Social Security)		1968	1978
in current prices	a) Index of amount	100	971
	b) as % of GDP	15.7	27.5
in constant prices (1970)			
	a) Index of volume	100	286
	b) as % of GDP	15.7	22.8
weighted deflator		100	469.4
Government Consumption Expenditure (compensation of employees and intermediate consumption)			
in current prices	a) Index of amount	100	791
	b) as % of GDP	8.6	12.5
in constant prices (1970)			
	a) Index of volume	100	238
	b) as % of GDP	8.6	10.4
Deflator (government consumption expenditure)		100	496
Government Transfer Expenditure (transfers, subsidies and Social Security benefits)			
in current prices	a) Index of amount	100	1146
	b) as % of GDP	7.1	12.4
in constant prices (1970)			
	a) Index of volume	100	1146
	b) as % of GDP	7.1	12.4
Deflator (private final Consumption expenditure)		100	443•5
Gross Domestic Product			
in current prices	a) Index of amount	100	364
in constant prices (1970)			
	a) Index of volume	100	161
Deflator GDP		100	381

However, the evidence provides greater support to BECK's proposition about a differential trend in the rate of growth between the general Government final consumption and the current transfers.

The present level of aggregation would not, however, make it an unambiguous conclusion to state that this differential trend in their growth already collects a supposed reaction to an adverse relative price effect with respect to the more controlable parts of Government consumption. In this sense, any potential verification in the terms of the proposition about a 'taxpayer revolution' may be, in fact, misplaced.

The exercise provides, above all, the conclusion that any macroeconomic analysis of the governmental budget constraint which tries to account for total expenditure with respect to some notion of a 'tolerable level of taxation' needs to focus attention on the different categories of the public expenditure included: (i) the resource-using side of public expenditure (general) government final consumption and investment; (ii) debt interest payments and net lending; and (iii) current grants, subsidies and capital transfers, -more correlated to the exogenous general economic circumstances.

Therefore, the question 'How justified is the argument about a rise of concern in the opportunity cost of the public sector involvement in the economy', is a question which cannot be justified today on economic grounds only. In the absence of appropriate measures in the overall incidence of the budget, we have tried to put that claim in context. We have seen, from the revenue side, some factors which may have affected, at least from a social-psychological point of view, potential changes in the state of this concern such as those resulting from some taxation changes carried out in a period of economic decline. However, no definitive answer seems today sensible concerning the accuracy of the bases of this concern, without accounting for the expenditure side of the public services provided and the degree of progressivity of the fiscal system as a whole. (See on this point J. K. GALBRAITH, 1981).

If the current pressures on Government expenditure in Spain, resulting from a greater attention to the unsatisfied social needs and from the negative effect due to the present general economic situation, are to be faced, it can be forseen that claims about an increasing opportunity cost from the 'public sector activity' seem bound to continue, at least for those economic groups mostly affected by the tax redistribution policies which will be required in order to finance that expenditure.

CHAPTER II THE ISSUE OF VARIBLE PRODUCTIVITY IN THE SPANISH HOSPITAL HEALTH CARE SECTOR: DEFINITION AND SOURCES

C H A P T E R I I

THE ISSUE OF VARIABLE PRODUCTIVITY IN THE SPANISH HOSPITAL HEALTH CARE SECTOR: DEFINITION AND SOURCES.

2.1. INTRODUCTION.

In the first Chapter we studied the general setting of the question of system inefficiency in the performance of the Spanish hospital sector. In this second Chapter we will explore, the question of why, some hospitals appear to be more inefficient than others.

The line of argument in this chapter starts from the recognition of some stylised factors concerning hospital variations in performance, its importance and the problems and caveats which appear in its analysis. Given our interest in studying the extent of their relevance against the current implementation of some general 'across-the-board' strategies for financial restraint, we will review the existing literature on the subject, providing some basic definitions for a more consistent approach to the phenomenon of variable productivity.

In the line of other behavioural studies in productivity in Public Sector analysis, we will focus on the role of the budget financial restraint in explaining the motivation for efficiency, and hence on its differential effects in terms of institutional performance, according to some basic behavioural postulates.

A model of hospital behaviour suited to the Spanish context is presented in this chapter and some behavioural factors of potential relevance in explaining the existence of variable productivity are correspondingly pointed out. These factors, and their mechanisms of operation, are then analysed and to a certain extent validated according to the existing data available. In the conclusion of the chapter we argue the need for introducing major changes in the organisational framework and the financial arrangements existing today in the hospital sector, focussing on the direction of the research described in subsequent chapters.

2.2. STYLISED FACTORS CONCERNING VARIATIONS IN PERFORMANCE IN THE SPANISH HOSPITAL SECTOR.

It was argued in the previous chapter that under a system of hospital revenue allocation based on retrospective spending, agents in the hospital sector have no incentive to manage resources in a cost-conscious manner. Furthermore, it was pointed out that existing policies for expenditure restraint which provide 'across-the-board' increases in hospital budgets, according to some externally imposed cash limits on the basis of hospital past expenses, cannot support a 'rational' constraint aimed at the improvement in the allocation of hospital resources.

One strategy for encouraging cost-conscious management would be to undertake experiments to examine how efficient the hospital system could be made and, therefore, to provide a benchmark against which the present system could be compared. However, such a strategy, which involves the estimation of what the hospital could do under a set of trial-controlled circumstances - which are deemed to reflect ideal operating conditions - is likely to be highly costly and not practical in the present context (see Dillon, 1979, and Lusk and Lusk, 1979). Instead, attention will be focussed, in the present thesis, on the potential efficiency motivation raised from instrumenting a budget control policy, aimed to reduce variations in the relative efficiency of hospitals, through a prospective revenue allocation system.

Variations in performance in the Spanish Hospital Sector.

It is clear that as a result of the set of 'perverse' financial arrangements which surround hospital reimbursement in Spain, one could expect that the system in its entirety to be inefficient. But it is also becoming increasingly recognized that some parts of the system seem to behave in a less efficient way than others.

The considerable variation which can be observed in the performance of comparable hospitals may indicate that such variations cannot be justified by the single notion of 'hospital uniqueness'. Therefore, some mechanisms for its control are required.

A crude indication of the existence of these variations in the Spanish Hospital Sector can be seen in Table II.1 below.(1) Similar variations have been recently studied by Rothberg (1982) for the United States and illustrated by Akehurst and Haycox (1982), among others, for the United Kingdom.

TABLE II. 1

Some variations of 'performance' in the Spanish Hospital Sector (1979)

Subsample selected under criteria for increasing homogeneity

Values of the Coefficient of variation: (Standard Deviation/Mean)

Average length of stay	17.11	Average cost per bed	32.0
Bed Occupancy rate	18.60	Average cost per patient day	20.3
Number of cases per -total personnel	27.1	Average cost per case -total	23.2
-per physician	28.3	-pharmacy inputs	63.0
-per nurse	27.7	-laboratory inputs	58.5
no. of beds available per ancillary worker	20.6	-radiology inputs	54.3
		-surgical inputs	95.3

Subsamples with greater within group homogeneity.(*) (Coef. of variation)

	Subsample selected (total)	Only teaching hospitals	Non teaching hospitals	Hospitals with outpatient services	Hospitals without outpatient services
Beds/Physic.	27.6	21.6	30.6	24.0	45.5
Nurses/physic.	27.4	19.3	31.3	24.9	38.9
Cost per case	23.2	19.8	25.9	20.8	23.3
Cost per patient-day	20.3	16.3	23.3	17.2	18.6

(*) The description of this selection can be found in chapter V.

Large degrees of variation are to be found for most of the indicators of hospital performance utilized. This can be seen for the Spanish hospitals in the tables included in the appendix to this chapter. Moreover, this substantial variation remains when one confines one's attention to comparatively small groups or relatively homogeneous hospitals (see table 2 in the Appendix).

Some existing problems and caveats.

There are those who may argue, however, that because of the peculiarities of hospital output such comparisons are not valid. In chapter five of the present thesis it is shown that even after controlling for a large number of variables which might be construed as having an effect on the nature of the hospital output, a good deal of variation in hospital performance remains. It seems unlikely therefore, that one can account for these apparent differentials in performance only on the basis of differences in the nature of hospital output.

Having said this, it should be recognised that these are crude indicators of hospital performance: For instance, the ultimate output of hospitals, namely 'health', is a multidimensional concept and one that is extraordinarily difficult to measure, even in principle, let alone in practice when one considers all the limitations of available hospital statistics.

In relation to this issue, we would like to make clear the following points:

Firstly, the true definition of hospital output would be related to some notion of patients' health states. In fact, as Fuchs (1973) has pointed out, the contribution to health of health services, provided by the health industry in general, should be ultimately referred to the utility derived from the status of being "healthy" or from having improved the individuals' health state. In this sense, "health" appears as an objective, and the production of health care is just one method of attaining that objective (Williams, 1983).

But 'health' as an ideal state, such as that defined by the WHO from a broad welfare notion, involves substantial difficulties for its valuation and measurement.

A relative concept has been preferently instrumented in which "health" has been defined "in relation to the essential possibility and existencial reality of disease" (Stacey, 1977).

Thus, 'health' has been defined with regard to some social, physical and mental dimensions such as "functional fitness", "absence of pain and distress", "capacity to engage roles and tasks in social life", "self-care", etc.. In this sense, health appears as a relative concept which may vary among different societies over time and even between different social classes. Illustrating this, Culyer (1981) has pointed out the fact that lower back pain appears as a part of every day existence of women in low social classes. Consequently, and to certain extent, the definition of health becomes a part of a "social norm" within the larger concept of a group and cultural definition (Twaddle, 1974).

In sum, the dominant concept of health is the one of 'ill-health'. According to this notion, a conceptual framework for the measurement of health services output has been given by Culyer, Lavers and Williams (1971), and some related methods have been developed and used. (A full review of the issues in relation to the measurement of health indicators - with regard to the different purposes for which they are designed - can be seen in Rosser (1981)). For our purposes - the empirical approach to hospital output - the issue is discussed at some length in chapter four of the present thesis.

Secondly, the apportionment of the set of relevant inputs to be considered (and the external quasi-inputs in some cases), and its attribution to those multidimensional aspects which in practice constitute the output of the health care industry(2) is not an easy task, due, largely, to the relevance of joint production in the hospital sector.

Thirdly, the types of performance indicators mentioned so far can even be missleading because they may fail to reflect real improvements in patients' health states. This is particularly true in the context of technological change. For example, medical innovations which permit, for the first time, postponment of death or the compensation for the impairments of particular diseases - rather than curing or preventing such diseases - will be reflected in the conventional performance indicators as reductions in hospital productivity. It should also be born in mind that such statistics, unless used carefully, may concile real increases in output associated with increased case-mix complexity, brought about by the introduction of new medical treatments (Wadley, 1972).

Bearing these points in mind, the fact remains that there do appear to be considerable differentials in hospital performance, which need to be explained.

The remaining of the chapter is concerned with the issue of how such apparent variations in efficiency might be analysed. Prior to this, however, some definitions concerning 'efficiency' and 'variable productivity' are offered with a view to clarifying the ensuing discussion.

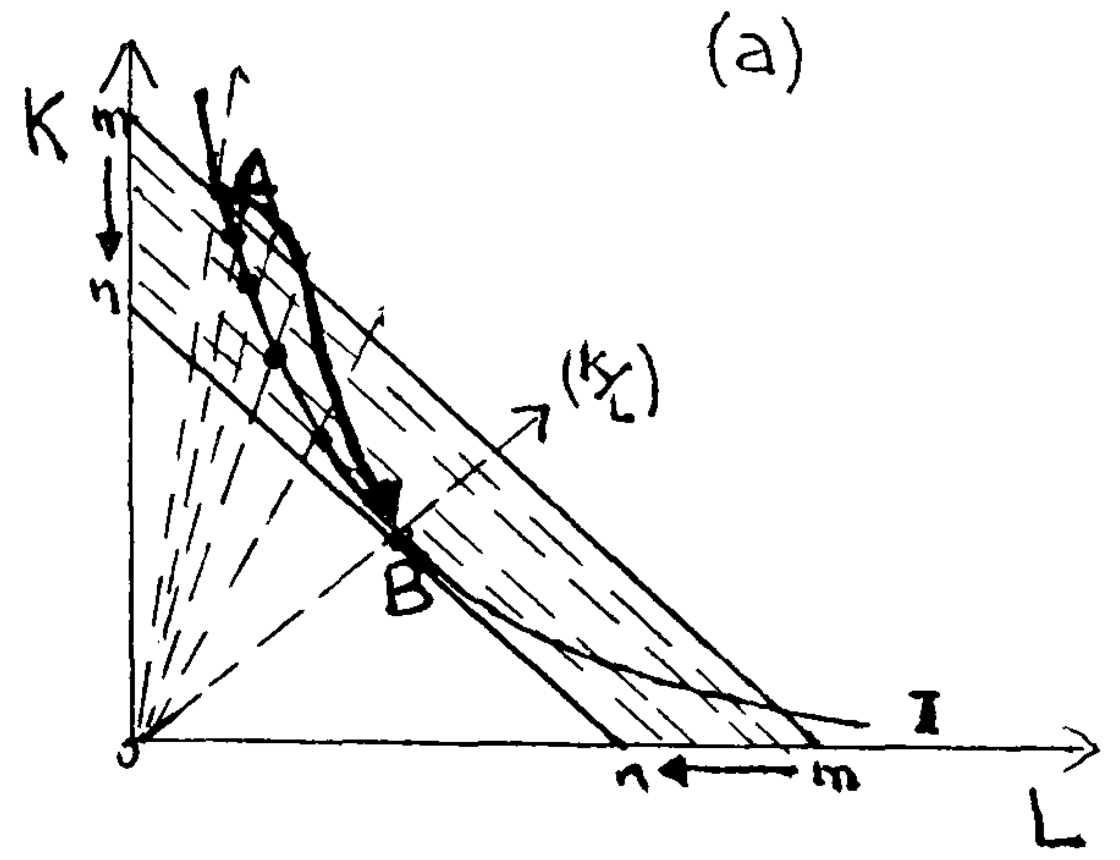
2.3. DEFINITION OF EFFICIENCY AND VARIABLE PRODUCTIVITY.

In the present thesis, 'efficiency' is defined in the sense of 'production efficiency' as used in standard microeconomic theory (see Layard and Walters, page 12).(3) Consequently, efficiency in performance is defined in terms of a production unit producing the maximum output with a given set of inputs, or equally, producing a given level of output at minimum cost. This notion of efficiency can be used to provide the basis for definitions of *relative efficiency and improvements in efficiency.*

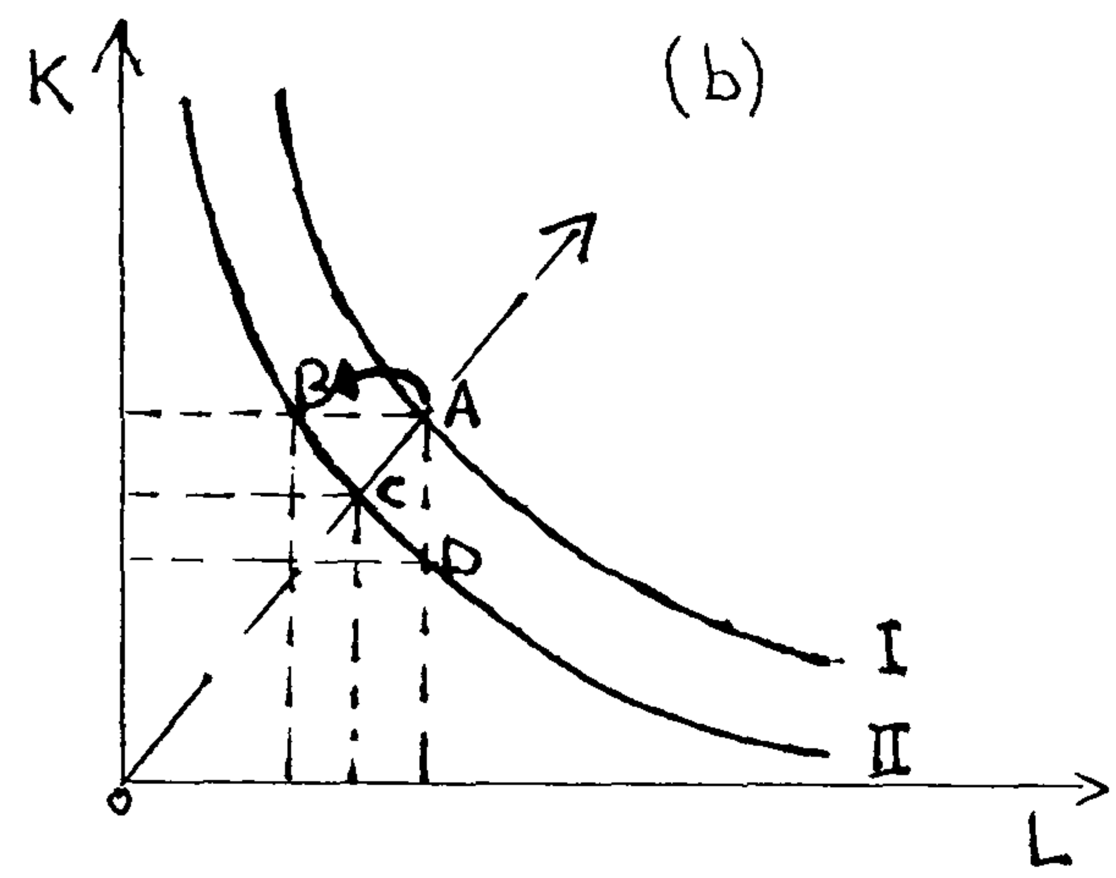
A production unit is considered to be relatively efficient if it produces the same output as another production unit but with a smaller set of inputs, or if it produces the same level of output as another production unit but with a lower cost. The existence of variations in efficiency (i.e. some producers being more efficient than others) can be termed a situation in which *variable productivity* exists.

Efficiency improvements (productivity gains) occur when a producer increases her output without increasing correspondingly the level of inputs utilized, or equivalently, produces the same level of output at a lower cost. The notion of efficiency and one factor productivity improvements can be seen in figure II on the following page:

FIGURE II. 1



Efficiency improvement: Movement from A to B. This is, identical output with less cost through a different input combination. (I isoquant; mm, nn, isocost lines, from higher to lower levels, holding identical relative input prices): Avoidance of costs incurred by operating on an inferior isocost line.



One factor productivity improvement: Movement from A to B. This is, identical output with a lesser amount of one of the input utilised holding the amount of the other input constant. (I and II represent isoquants of an identical level of output: K and L Capital and Labour inputs). Factorial Productivity Improvement: Movement from A to C, or any other point on B D.

The terms 'efficiency improvements' and 'productivity gains' will, in the following pages, be taken as identical concepts. This is particularly acceptable in a comparative static approach, although in more rigorous applications, specific assumptions about perfect price competition and constant returns to scale are required (Denny and Fuss, 1980). (4)

In a practical sense, 'productivity' is a relative concept with respect to which 'efficiency' constitutes a simple guideline for analysing performance.(5) According to this view, three variants of these notions can be accounted for stating performance comparisons. These are those derived from:-

a) Achieving the same as before under a different set of circumstances.

b) Doing the same as someone else in a similar position under similar circumstances, but outside this sector.

c) Doing the same as someone else in a similar position within the sector under similar circumstances.

Efficiency comparisons on a cross-section basis in the past have mostly referred to the public and private institutions debate (e.g. Borcherding, 1975 and Spann, 1977 et al.). In general, this exercise has proved useless in our view, since it involves attempting to compare incomparables, due to (i) the failure to estimate accurately all the dimensions in the value of the public output, for instance, in terms of 'equality' and not just of 'efficiency' targets, and from the analysis of both expenditure and income sides of the Public sector activity; (ii) the conditions imposed in the public context by outside regulators, such as minimum standards, extent of geographical and social coverage, etc. This in practice distorts the comparison (Lindsay, 1976); (iii) the lack of realism in the proposed alternative systems of provision: one based mainly on the principles of market competition and Pareto efficiency postulates, against another built in a set of prejudiced assumptions about the behaviour of bureaucracy in public administrations. (6)

The public-versus-private provision debate can also be refuted however 'ab initio',(7) on the grounds that the question of the entitlement of a public service must be solved on grounds of principles, independently of features such as the actual production characteristics (features which should be controlled by other measures, for example, through internal policies or direct regulation).

But as we will see in the next section, less attention has been devoted to the analysis of the relative efficiency of an _activity in one unit as compared with the same activity in another similar unit within one industry or branch of activity_. The reasons for this will be explored in the next section. A review of the literature of potential relevance to the issue of variable productivity will be for these purposes presented.

2.3.1. A review of the literature of potential relevance to the issue of variable productivity.

Neoclassical economists have not pursued the analysis of variable productivity very vigorously. This has been probably due to the fact that, according to neoclassical theory, factors leading to differential productivity levels within the industry - such as those derived from the level of technological implementation and rate of input utilization - should be regarded as momentary distortions in market resource allocation, which sooner or later would be corrected by the operation of markets. In this sense,either the production technological posibility frontier is reached, or market prices will result by pushing the firms that failed to attain it out of business. This logic is at the core of the neoclassical argument stemmed from the notion of a production 'black box' in the context of fully efficient market mechanisms, in which output appears as an automatic and direct result of the inputs available. Although some sort of productivity gap could rise due to the induced bias of innovation (see to this respect the Ahmad-Kennedy controversy, quoted from Ferguson, 1967), any recurrent feature of variable productivity was bound to be inconsistent with the workings of the market, according to neoclassical theory.

This view has been correspondingly translated into practical policy terms by devoting attention primarily to the design of policy instruments derived from first principles and under the assumption that they would have the desired impact. They have neglected therefore both the issues associated with the implementation of policies and the question of whether they would be automatically addressed in the manner intended. As Heald (1983) and other authors have pointed out, such implementation would require altrustic public officials, maximizers of social welfare and untainted by self-interest, who are omniniscent about cost and demand conditions, skilled in the use of analytical techniques and committed to their underlying value judgements. Only the homogeneity of all these factors could then make, in practice, for similar levels of performance.

This lack of recognition of the issues of variable productivity[8] has consequently left unexplored the potential interrelationship between the static and dynamic components of productivity gains. Except in the case of the Theory of Cumulative Causation[9], the effects that different levels of productivity can have on their sequential rate of growth, over and above the single intensity change element stemming from the increase in the amount of inputs available or in their rate of utilization, have remained unexplored.

In this context, intersectorial studies of productivity have been analysed on some spatial references - across nations, across regions- and over time. They have focussed mostly on the accountancy rather than on the behavioural aspects in explaining productivity difference, with some minor exceptions (Jorgenson and Nishimizu, 1978). These studies do not provide, however, much insight into the issue of relative efficiency or variable productivity from an intra-sector intra-industry approach to the analysis of the institutional activity. This is mainly due to the degree of sectorial aggregation of this type of studies and the extent of the problems involved in comparing different sectors in practice, such as the lack of input and output homogeneity in the data compared and, in the case of the international studies, due to the problems which affect the vector of prices in order to reflect the relative purchasing capacities (Kravis, 1976).

Leibenstein's X-efficiency Theory stands aside of the former view and, questioning some of those basic neoclassical postulates, comes to recognise the problem of relative efficiency in most of the economic activities in reality, over the pure issues of theoretical allocative efficiency. In this sense, he argues for the possibility of 'living with' an eventually persistent failure to reach the production possibility frontier in some economic activities due, for instance, to social and psychological factors which surround their production processes. Generally, this would appear to be the case in those productive activities which take place through real time, involving individuals who try to shift constraints in their own interest, who 'learn' about processes and goals, and perform in an environment where production takes place in a climate of conflict, uncertainty and information assymmetries. But other organisational factors may have a role too: This has been pointed out by some theories of industrial sociology and job design (Lawler et al.), structure of organisations and networks within bureaux (Breton and Wintrobe, 1982), and workers' participation and industrial democracy (Hodgson, 1982).

More fundamentally for our purposes, one of the factors pointed out by Leibenstein in explaining X-efficiency is precisely that of the absence of concern in the firms' management for financial restraint. In fact, Leibenstein analyses the influence of the budget financial concern on the agents' behaviour, and hence on the institutional performance, through the interaction between constraint concern, exerted effort, performance levels and external financial pressures. In the terms seen in the first Chapter this interrelationship is particularly relevant in causing X-inefficiency in those contexts in which agents have the power to choose the pace at which the activity has to be carried out per unit of time (effort), quality, time pattern and length. In these circumstances the predetermination of the boundaries for job specification and agents' activity are not externally imposable by reinforcing contracts on separate individual units. Information assymmetries, uncertainty, demand inducement and the nature of the output in itself come to support, and ultimately to strengthen, discretionary powers in the hands of the agents, in preference to any other institutionally assigned objectives.

Most of the features mentioned above have been seen in _the hospital care sector in Spain_: in particular, the absence of financial concern on the utilization of health resources, as a result of a system of revenue allocation based on retrospective costing which does not provide incentives for a cost-conscious management to the self interested agents in the sector; agents who are moreover vested with a set of discretional powers granted by their professional knowledge.

This view seems to be in accordance with the set of suggestions which are, it is argued, to be found in those type of contexts where the X-inefficiency postulates are present (such as a lack of evaluation feedbacks, absence of economic trade-offs in operative choices, etc.) and which are reflected in the hospital literature by a set of inferences about organisational slackness (lack of services co-ordination, duplication and misutilization of some hospital facilities, absence of monitoring and allocation or responsibilities in resource management, and an absence of scrutiny in confronting an excessive rate of technological implementation - which are biased towards product rather than process innovation,etc.).

However, the relevance of the budget financial restraint for purposes of a behavioural analysis of the performance of some public institutions has been only recently considered in the public sector literature. Thus, only in the general field of Public Choice theory, the 'median voter' model, the institutional approach to the Grants Theory and the supply theories of public government output have attention been focussed partly on this issue. In the next sub-section we will review those approaches regarding the role they assign to the budget restraint as the 'price' element towards efficiency motivation targets in the public sector activity.

2.3.2. Behavioural analysis and Productivity in Public Sector Studies.

The role of the budget restraint in any constrained optimization exercise is to reflect the resource limitation in the Agent's decision-making process. That involves, in a conventional manner, the maximization of a function $F(X_1)$, subject to the financial restraint that $M \geq (X_i)$, from which first order conditions are derived: i.e.

$$\text{(i)} \quad f_i - h_i \leq 0$$

$$X_1 \geq 0, \text{ and}$$

$$\text{(ii)} \quad M - h(X_i) \geq 0$$

$$X_i \geq 0,$$ where the budget constraint M reflects the amount of resources to be spent in order to produce a maximum return, either expressed in terms of satisfaction from consumption, output production or profit. The nature of this role does not change if the resource limitation is imposed over time, that is, whether it is considered as a revenue stream which in a planning horizon context gives the feasible set of points for a constrained economic optimization behaviour.

In the Public sector context the lack of markets and a competitive pricing system prevents, in principle, that sector from reaching a private goods type of solution. However, in some areas of the public activity, a pseudo-market interpretation has been postulated to simulate that result, once allowance has been made for some special characteristic of social goods and the revenue process involved in its provision. Two main types of decentralized public agencies could be distinguished for these purposes:

a) Agencies at least partly funded from their own revenue raising capacity.

b) Agencies under general fund financing techniques or external budget-based arrangements from a reimbursing Authority.

This distinction is important so far as the first type a) of Agencies incorporate, or can incorporate, an implict watchdog for efficiency in the revenue side of their budgets, whereas for the second type b) of Agencies, global revenues come out of general taxes on the population, without having expenditure coupled to a tax to finance it.

Public Choice theories have viewed the problem involved in a) either from the median voter type of models or from the more general view of the institutional Approach to Grants Theory.

Accounting for a tax-price notion[(10)] for some Public sector activities where the output is not sold on a market in any way, median voter models allow for the introduction of a 'price component' which can translate directly the role of the budget restraint in the organisation's performance. This component can then play the role of an indirect watchdog for efficiency on the sequential decision-making process of the public agency expenditures. In this sense, the derived price effect, even in those cases collecting primarily a productivity gap notion (Beck, 1982) may, afterall, exert a positive influence for any ex-post potential improvement in productivity levels, by forcing changes in the Agency's behaviour. In other words, the magnitude of the price elasticity of demand for some specific output services can have a direct effect on additional future expenditure and on further cost-conscious management. As Baumol foresaw in his model (1967), this price component can be in itself a regulatory mechanism for trading-off unbalanced productivity gains with unbalanced revenues.

In consequence, as far as differences in price elasticities of demand could reflect differences in the mechanisms through which the revenue of the Agency is raised, this could be in itself a source of differences in productivity levels, given its differential impact on the cost-conscious managment of resources.

No further work has been done to link the general fund financing methods with those behavioural-related productivity factors on the performance of the public institutions. The _institutional approach to the Grants Theory_ (Gramlich, McGuire, Oates et al.) represents however, a significant approach to this issue. In fact, this approach emphasizes the importance of the reimbursement processes on the decentralized behaviour of the institutions, drawing attention to those questions of policy implementation. That is, it analyses the impact of the different types of grants, or of the revenue allocation process in general, on the local decision-making process and the institutions' performance, instead of examining their effects purely in abstract.(11) As a result, questions of potential simultaneous causation between central transfers and local expenditure, the strength of the behavioural mechanisms for self-regulation(12) and the need of a prior adjustment for some interterritorial needs, etc. can be, under this approach, only then analysed.

Although similarities between the institutional approach to the Grants Theory and the issues in this section are strictly limited by the 'local resources element' which surrounds fiscal federalism, this approach provides a significant point of departure for further research on productivity and behavoural analysis in Public sector studies, so far as it searches, for modelling purposes, for the influence of the budgetary reimbursement methods on the agents' and institutions' behaviour. In other words, it searches for the analysis of their consequences in terms of agents who use inputs to produce outputs according to some assigned targets, but who try above all to maximize some sort of welfare or utility function within some boundary limits, by adjusting their position with respect to the production opportunity lines, once they are faced with different spending restraints.

On the contrary, for those _type b) of public decentralized Agencies_ with central revenue coverage of expenditure, no direct link appears between the institutional effects of the revenue and expenditure sides of their activity. In this case, the 'illusion' of a central budget reimbursement made as a 'third party' payment may lead citizens and/or bureaucrats (Niskanen, 1974) to exhaust to zero the marginal benefits derived from the demand (and/or supply) of the Agencies' output. This can happen in theory as the result of two different _patterns of behaviour_:

(i) On one hand, citizens may behave as accounting initially for some notion of the (average) tax-revenue raised on their income sources as a 'pre-paid' tax. From this initial position, an argument of the following type can be put forward: "If I were not going to pay so much for these public services, say health care services, I would be happy enough to buy this other bundle of goods (say, private health insurance). But in the current situation, my best choice consists in taking advantage of the actual free provision up to the exhaustion of the marginal benefits which they produce to me".

No predictable pattern of individual rational behaviour would be, under these circumstances, easily reconciliable with the aggregate factual behaviour. It may involve, in addition, a "composite effect trap" for increasing public expenditure throughout the chains of 'increase of output demanded' - 'need of increased revenue allocation' - 'additional revenue raised' - 'successive increases of demand', etc., assuming that the citizen constantly fails to discount the costs implied in those sequential decisions, suffering fiscal illusion from viewing the institutional expenses as a third-party payment.

(ii) It should be recognized that, obviously, no direct effect in this way is produced in reality, since some sort of control is laid down by public regulation, and central management restraints come to operate over the short-run influences in each stage of planning expenditure. However, in those contexts in which agent supplier behaviour can dominate the external regulations by shaping the citizens' demand, this type of pattern of behaviour gives to the institution some discretionary powers which may be put into play on behalf of some notion of a potential 'citizens' reservation demand' (Breton and Wintrobe, 1982). In this way, wherever no direct link appears between the revenue and expenditure side of the activity and supply influences demand, the domain for the entrance of additional 'segments of demand' can be used by the agents for reaching their own specific targets, affecting correspondingly the institutions' performance. (University and hospital admissions could be examples which illustrate the features above mentioned). Thus no single internal factor for a cost-conscious management may be able to face this strategic game leading to pressures for increasing output-activity with ultimate effects on the performance of the institutions. This can be in itself, in our context, a source of variable productivity given the characteristics of the Agency's decision-making process for resource allocation and, particularly, given the possibilities of translating the financial consequences of these decisions into the external constraints throughout the budget reimbursement process.

In sum, not many analyses have been developed in Public Economics on the issue of the influence of the revenue allocation methods on current operational behaviour of public institutions. However, from the most general view of these studies, it appears that the following point can be stressed: Whenever the budget restraint can reflect somehow a 'price' component on the institutional activity, budget revenues can have a role as implicit watchdogs for efficiency and make for some major behavioural changes in public institutions. But where this is not the case, and agents' supply factors influence demand conditions-and in this way the institutions' performance - the possibility to link budget reimbursements to pre-stated levels of performance constitute, in our view, the only control policy which may provide an efficiency motivation 'rationale' against factual pressures for sequential expenditure increases. (This possibility will in fact be studied in the third chapter of this thesis).

2.3.3. Outline of the argument.

The essence of the argument outlined so far in this section can be summarized as follows:

1. In a competitive market system, productive efficiency is ensured by the fact that all firms produce similar products, facing the same factor markets, and hence any high-cost producer will be squeezed out. In the absence of strong competition, X-inefficiency may arise and persist. The reasons for this have been analysed by Leibenstein et al., and the influence of the financial budget restraint concern on institutional behaviour has been suggested.
2. In those parts of Public sector where the output is not sold on a market and revenue is not systematically linked to performance, additional scope for the exercise of managerial discretion can be present.

3. As we have pointed out, the Spanish hospital system manifests the aforementioned properties. Hence, one could expect (i) the whole system to be inefficient, and (ii) some parts to be more inefficient than others.

The foregoing exposition has framed the context in which the budgetary process operates in general in the Spanish hospital sector. In this sense, its analysis provides a common basis for explaining the hospital activity and namely, the system inefficiency postulates which surround the provision of health care. It cannot predict in itself the reasons for a differential impact on individual performance, here collected under the notion of variable productivity. Thus in order to explain and predict 'relative efficiency' in hospital operation, and to propose some appropriate mechanisms for its control, we require a model of hospital behaviour. Some considerations for this purpose are presented below.

2.4. EXPLAINING VARIABLE PRODUCTIVITY: MODELS OF HOSPITAL BEHAVIOUR

There is a relatively extensive literature on behavioural models of hospitals, a useful survey of which is contained in Jacobs (1974). In general, we could split the existing models on hospital behaviour into the following groups:

a) Models which treat the hospital as "organism" in itself, focussing on the behavioural-organizational aspects, either from the view of hospitals as "bureaux" (Migué and Belanger, 1974), non-profit institutions (Newhouse, 1970) and monopolies (Pauly, 1974; Goldfarb et al., 1980), or from a bargaining type of approach based on the property rights of the hospital agents (Clarkson., 1972; Harris, 1977).

b) Models which focus on hospital agents' behaviour - either doctors, administrators or trustees. These models usually frame the set of variables influencing individuals' behaviour into profit, output or utility maximization functions - here with arguments such as 'sophistication' of equipment, discretion, power, etc. (Lee, 1971; Bongaerts and Carrins 1982). This maximization exercise is commonly constrained by some function defining the convex set of feasible points for equilibrium, being mainly specified in terms of:

(i) a certain level of income, exogenously given or depending on the income of some reference group (Feldstein, 1970), or some function which comes to incorporate the marginal valuation of the individuals own input time, which may also depend on income (Sloan, 1974), in order to lead to certain labour leisure trade-off;

(ii) some less-well defined notion of "professional ethics" in terms, for instance, of the "best practice" and "patient welfare" (Murray, 1974), which usually allows for a less clear-cut derivation of arithmetical results.

As it is now clear from the first chapter, model (a) fails to account for the reality of Spanish hospitals. They are, in our view mostly, dominated by doctors who feel their role is to do everything they think to be "scientifically" needed for their patients without reference to price. In Harris' view, doctors have aimed at enjoying access to hospital inputs without getting told what to do. But in our context, and contrary to what models of "property rights" assume, there has not existed an effective "supply division" in hospital management: indeed, hospital administrators have had little stake in opposing physicians and they have found their own job security tied to their ability to satisfy the demands of the medical staff, pushing thereafter for the "third party reimbursement" of total hospital expenses.

Model (b) has been splendidly discussed by Evans (1974), but in our context, it fails to account for the actual mechanisms through which hospital agents operate; this is, a situation where doctors are professional 'salaries', face a general wage settlement - taken as given for a notional working time - and "shirking" is a common feature in relation to the accepted compatibility principle in the medical practice. Indeed, this situation does not seem to offer support for the formulation of a general profit objective nor for other ouput maximization targets. As we will develop in the next section, it makes more sense to formulate a general utility function, in terms of "satisfaction" with relation to a "given" rate of earnings per effective working-time.

In conclusion, whereas most of the existing models may be of use in some institutional settings, they have little useful application in the Spanish context, given the characteristics which surround the provision of hospital care. However, their study and cross-comparison enable us to derive some general considerations in order to model hospital behaviour in a way more adjusted to reality.

2.4.1. Some Questions in modelling hospital behaviour.

In our opinion, the following questions need to be answered first in modelling institutional hospital behaviour:

a) With regard to the unit of analysis:

(i) Does the hospital as an organization involve in itself a single consistent set of behavioural features which allow for a joint analysis for the purposes of modelling?.

(ii) Can these postulates be constructed only from the behaviour of the individuals - which ultimately shape the performance of the organization - and, if this is the case, is it the full network of interrelations among the agents or is it that of some specific sub-group of them which explains the core of the resulting aggregate behaviour of the organization?

Our view in the Spanish context is that, goals are set by individuals rather than by organizations, and that the allocation of health resources within hospitals is mainly derminated by physicians, acting according to a set of dominant factors, which have their roots in the technological information and professional knowledge available to them.

This view is derived from the study of the rules which conform hospital organization and management ('Reglamento înterno de funcionamiento') and the particular status which determines the working conditions of physicians (against the 'Estatuto Laboral Unico').

b) In addition, four significant aspects have to be considered in order to achieve a better understanding of the relevant features for modelling purposes, which surround the hospital health care sector:

1. First and foremost, the ownership of the hospitals.

This is, whether the hospital is a private profit-making organisation (i.e. within the corporate sector), a private non-profit institution (i.e.) under Charity funding or Trustee) or a public Agency within, say, a decentralized Public Sector.

2. The system or remuneration of the hospital staff, and mainly that of the physicians in particular: For instance, either as share-holders of the hospital financial results or receiving fees per medical act, or having some type of profit-sharing in addition to a minimum level of earnings or as simple employees with a fixed wage.

3. The reimbursement method of the hospitals global expenses:

(i) this can either be cost based a) retrospective or prospectively; b) with or without incentive mechanisms in operation; c) with or without a budget link to the individual performance, and (ii) with the payment made a) per unit utilized (day of stay, case treated,etc.) or b) by some other broader notion which takes into account the total value reimbursed (e.g. per ward, per department or global budget).

4. Finally, the nature of the hospital in itself due to its location (e.g. whether the hospital is situated in a rural area having to provide a minimum standard of care on a geographical basis) and status (whether it has to supply teaching, medical research, etc.).

c) Furthermore, within the framework set by the former combined factors, we need to pick up those more sensible assumptions about the orientation objectives in the activity of the staff and its potential identification with the institutional goals of the hospital authorities.

Among these orientations we can point out here (See Friss and White,1977):

(i) A view of physicians purely as professionals, with ocupational targets such as the improvement of job skills, the interest for being aware of technological advances, etc..

(ii) Physicians as goal orientated individuals, pursuing their specific objectives, instead of following, for instance, a conventional set of standard procedures or "clichés" in the medical activity.

(iii) Physicians as bureaucrats who value, above all, job security, social status and a steadily increasing income.

(iv) With physicians viewing their role as an instrumental one, never regarding the work situation as a matter of realization in itself.

(v) With physicians viewing themselves as "servants" of the community, or receivers of a social command, having a strong "medical ethos" (reflected into the hippocratic oath) which requires a paramount concern for the patient as a person in need.

It can certainly be argued that reality is usually a mix of these propositions. In general, the dominance of one of the former points needs to be considered at some stage of the behavioural analysis.

d) Finally, the medical role orientation has to be balanced against the professional motivation with regard to the material rewards (salary, wage differentials, fringe benefits or any specific employer incentive plan) and non-material rewards (service and professional recognition). These components can, in practice, shape the agents' incentives in order to qualify as postulated in the modelling exercise.

In sum, accounting for the former considerations and with respect to the specific sample of Spanish General Hospitals here analysed, the particular approach outlined for the analysis of hospital behaviour is next presented.

2.4.2. A model of hospital behaviour suited to the Spanish context: The approach adopted.

The outline of a model of hospital behaviour suited to the Spanish context can be defined in the following terms:

The approximation adopted here for the analysis of the economic behaviour of our specific sample of General Hospitals, which are providers of a set of standard inpatient care according, in principle, to their status and location, and which are owned and administered by the National Health Institute through central direction, funding and decentralized management, is the following. The hospitals are neither in reality, closed organizations with their own goals, nor do they reflect a single internally consistent decision-making process which might be analysed according to some standard homogeneous postulates of the organisational theory. We believe that they may be better viewed as decentralized public agencies with a particular type of bureaucratic pattern of organisational behaviour, which is ultimately influenced by physicians as the dominant agents in the sector, given the nature of the ouput provided and the impact of the medical decisions on the utilization of the hospital resources. This pattern has seemed to allow in practice neither for a direct hierarchical authority imposed from outside the organisation, nor for the application of a set of rules widely accepted in themselves for purposes of hospital management.

This means, in fact, that hospital agents cannot be considered as 'prototypes' of bureaucrats performing within centrally stated constraints. In reality, hospital agents are able to influence these constraints in order to adjust some parameters of the hospital production function which remain under their exclusive control, and which can then be orientated towards their own advantage. They can be better defined, therefore, as 'selective rational agents' who'satisfice' rather than 'maximize' some internal goals within the Agency activity with regard to certain guidelines (such as social status, remuneration, recognition..).

The following are the <u>reasons for the adoption of this approach</u>.

(i) The particular organisation of INSALUD, highly centralized despite its proved incapacity to monitor the performance of their hospitals, and lacking an efficient communication network.

(ii) The specific system of hospital reimbursement, mostly cost-based, with the Central INSALUD Authorities struggling to control expenditure through an unrealistic incomes policy, in isolation from the monitoring of the remaining elements of the system.

(iii) The statutory nature of the Hospital Boards, primarily representative of the interests of the predominant hospital agents within the medical profession at any particular moment in time. The absence of consumers' direct representation and of any other form of citizens' presence in the Hospital Boards, have consequently enabled Medical Directors to pursue their objectives rather freely.

(iv) The statutory nature of the role of hospital administrators, which makes them civil servants, 'clerks' of the Hospital Board rather than effective managers of the hospital resources. Under the prevailing reimbursement system, Administrators have had in fact little stake in opposing the dominance of the physicians. On the contrary, they have found their own job security most closely tied to satisfying the demands of the medical staff.

(v) Finally, the generalisation of public and private medical practices made 'compatible' in the health care sector. This comes to offer an extra remuneration or a compensation in kind over the pure wage level, given the financial externalities derived, either on other public activities or on their respective private practices.

In our view, these joint factors help the understanding of the apparent neglect of the Medical Corporation as a body of open bargaining against the type of wage restraint imposed by the

Department of Health, and the search for more basic levels in the process of allocation of resources within the hospital, where physicians' particular interests better merge with the general interests of the hospital as a whole and with those of the citizens by extension. The possibility of influencing ultimately the financial constraints on these bases has been to some extent permitted, as we have seen, by the nature of the budget reimbursement process. The influence that these joint features can have in the generation of variable productivity can be consequently analysed from the observation of the two following factors: a) An increasing pressure for recruiting additional levels of staff in response to variations in the effective job remuneration; and b) the actual level of utilization of the existing medical inputs.

The examination of both factors, in accordance with their potential relevance for explaining the relative efficiency in the performance of the individual institutions in the Spanish Hospital Health Care Sector, will be introduced in the next section.

2.5. BEHAVIOURAL FACTORS OF POTENTIAL RELEVANCE IN EXPLAINING THE EXISTENCE OF VARIABLE PRODUCTIVITY IN THE SPANISH HOSPITAL HEALTH CARE SECTOR.

The literature review contained in the previous sections (together with the institutional details of the Spanish Health Care System pointed out for purposes of a behavioural analysis in the first Chapter) permit the construction of an operational hypothesis on the issue of variable productivity in the following terms: The differential impact on hospital performance can be postulated from the way in which agents proceed to adjust their effective rate of earnings through changes in the working time in principle scheduled. This can be considered, in sum, the bounded rational answer of the hospital agents to the wage settlement centrally imposed up to recent years in the sector. Hence, the adjustment is postulated to follow a guideline, given to a great extent by the opportunity cost of the public activity, in terms of the alternative earnings raised from possible given private practice. This strategic adjustment is due to the difficulty in monitoring the relationship between the individuals' work-effort and the level of output produced. This absence of monitoring makes possible the degree of differential 'shirking' required to obtain a relatively similar remuneration for a given work-effort from both fields of overlapping medical activity. Indeed, only two variables remain in practice under physicians' control:

(i) The level of effort or effective working time.

Changes in this variable are the result of pursuing the minimization of the effects of the wage policy centrally imposed and, given the characteristics of the hospital reimbursement system, the recruitment of additional levels of staff in order to hold 'output' unchanged. The adjustment follows the notion of a private opportunity cost which takes place through the labour-leisure trade-off or throughout a constrained joint utility maximization, given the postulated principles of medical compatibilit

The adjustment does not, however, allow the pursuit of maximization but just of satisfaction. This is due to the fact that maximization postulates would lead towards the provision of zero levels of effort, which are incompatible with calling public support for one's position. In this sense, there must exist a common interest in keeping some minimum levels of provision.(13) In practice, they can result from differential factors (such as the agents' risk aversion to public complaints, the level of estimated needs and the concern in avoiding unwanted direct intervention).

(ii) The level of utilization of the existing medical inputs.

Changes in this variable can be expected once it is postulated that certain levels of utility derive from the exercise of some 'property rights' in the hospitals' activity. This can be ultimately the consequence of the physicians' own consideration as part of a contract which enables them to get what they think at each moment as 'scientifically indicated' independently of its costs. In other words, this is due to the fact that once doctors become members of the 'hospital firm' they get assured access to hospital inputs without getting told what to do, when, or how to use them (Harris, 1977).

This factor goes back at times to the claim to purchase of more sophisticated, or just newer, medical equipment. Independently of their potential levels of utilization in practice and of an economic evaluation of its contribution with respect to the level of utilization of the existing alternative medical inputs. It can also be reflected in the search for overcapacity, looking for some suitable levels of slackness or safety margins against work-load fluctuations, and with respect to existing inferences about the deviation of hospital inputs for their use for private purposes.

Both factors can be consequently found in the over-stocking of resources - including manpower - in excess of those required to produce a given output stream and, therefore, behind the generation of institutional variable productivity.

2.5.1. Mechanisms of operation.

In practice, the operation of the factors mentioned above have been sustained by the nature of the developement of technologies and by the patients' demand induced by physicians. Both elements have been translated thereafter by the hospital agents into the hospital budgetary process as variations in work-loads, highlighting the potential pressures which otherwise would be created on the amount and quality of the care provided.

By postulating that throughput measures are to remain constant, the argument can be put forward in favour of additional resources (say levels of medical staff), which will allow finally for the adjustment in the utilization of the existing inputs and, particularly, in the supply of the physicians' own input time.

The rationality of this hypothesis stems from the examination of Table II.2. and the suggestion formulated by Macchia (1979) that successive increases of manpower in the hospital field have been the consequence of a reduction in the actual working-time of the existing staff, holding work-load and throughput measures constant; in our terms, they would have been the result of the effort in closing the gap between the private and public remuneration of the hospital agents by adjusting the effective working-time supplied. This can be sustained from the observation of the evolution of the wage bill of hospital registrar physicians and the growth of hospital physicians in relation to some conventional indicators of hospital activity. The causal relationship between both is the hypothesis we try to show here. This hypothesis has also been recently put forward in Spain from a sociological point of view by Pérez-Díaz (1982). He observes, in fact, the search for the proportionality between the public and private remuneration as the behavioural regulator of the physicians' actual working time in the Public sector. The notion of a minimum level of earning for a given work-effort seems to be viewed by physicians as a sort of moral imperative over any stated level of notional working time for which they might be in principle remunerated. This benchmark in fixing the amount 'to be earned' would be based in principle on social considerations about status and professional capacity, such as they are reflected from the earnings in the private practice.(14)

Pérez-Díaz has calculated through a survey of physician earnings for 1980, a gap between 60% and 53% with regard to the difference between what "should" be earned from the completion of the full notional working time established in the public practice - given what would be earned for an equal working time in the private practice - and what "is" the actual remuneration in the Public sector. Therefore, in order to equalize, in our terms, the price of the effective working time between the public and private practices, the notional working day in the hospital would have to be dropped by between 1/3 and 2/5. That is, to four or three and a half hours instead of the six hours per day fixed legally. (These figures could then provide a first indication of the size of the slackness inserted into the current hospital performance).

In formalising in simple terms the postulated mechanism described above (i.e. including only labour inputs), we can assume that hospital output (Q) is the result of the interaction between hospital physician manpower (M) and other hospital labour inputs (X) (including nurses and ancillary workers). The corresponding isoquant of production for the hospital output can then be defined from Q = Q (M, X) as:

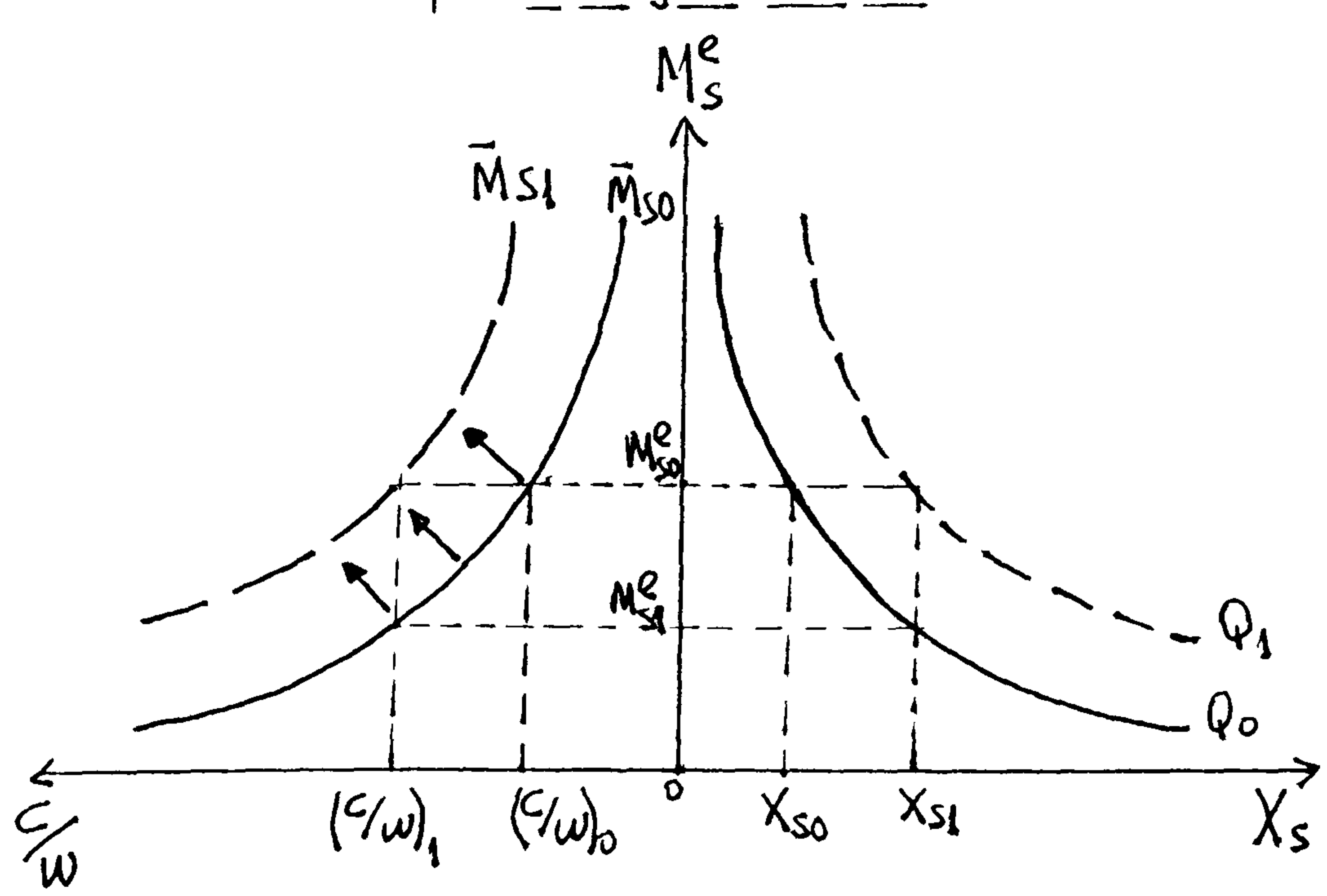
$$\partial Q / \partial M \,.\, dM + \partial Q / \partial X \,.\, dX = 0$$

However, as argued in earlier sections of this chapter, in those contexts where the pace of the work-effort remains under the individual's discretion and information assymetries do not allow for a close monitoring of the performance of the existing labour inputs, a variable which takes into account the degree of utilization of these inputs may prove relevant.

In fact, in our context, the level of medical labour input can be considered to be composed of a nominal component (i.e. the total number of physicians) $\bar{M}_s$ and the intensity with which this input is utilized M_s^e (or effective supply of physicians' working time). The marginal price of the first element is obviously the wage settled w, whereas the contribution of the second element has been assumed to follow the guideline given by the relationship between w and c, or opportunity cost of the physician's time spent in the hospital (e.g. its ratio c/w, according to our hypothesis). Hence, a given labour input in terms of 'units' of effective work-effort can be achieved either from increasing these units as a result of the incentives derived from an appropriate price remuneration, or from increasing the total amount of manpower, holding the level of work-effort per unit of labour constant.

By assuming that an equilibrium state is associated initially with a given value of the parameters w and c, say $(c/w)_0$, we can explore the comparative statics which result from an increase in the gap between c and w of the nature shown in table II.2. (e.g. from an unrealistic wage increase of the type described there).

In fact, as it can be seen in the diagram below, M_s^e, being a negative function of the ratio c/w, will drop to M_{s1}^e for a given level of physicians, $\bar{M}_{so}$, as we move to $(c/w)_1$. Figure II.2



In order to remain in an identical position on the same isoquant (Q_0), we can either increase the supply of alternative inputs up to (X_{S1}), or shift forwards the amount of hospital physician manpower available up to $\bar{M}_{S1}$ (i.e. keep the overall level of effective working-time supplied constant).

A cost-effective analysis of this choice would have to evaluate, therefore, the trade-off between the costs derived from ($\Delta \bar{M}_S$. w) and from (ΔX_S . Price of X) and, foremost, between ($\Delta \bar{M}_S$. w) and ($\bar{M}_S$.Δ w), given the initial increase in $\underline{c}$.

As argued in earlier paragraphs, this substitution does not operate in our context according to efficiency guides, say from substituting lower cost inputs for the higher costs of physicians' working time, given the existing interests involved in hospitals' management.

This result has to be put into the context of the nature of the reimbursement system and of the central policies followed on physicians' wage increases. Thus, pressures for sequential increases in the levels of medical staff have been preferently used in order to keep 'output' and 'quality' on line. These pressures have been favoured by the Hospital Boards as a way of dealing with the operational problems of the sector, and by the Medical Corporations as a way of reducing the increasing levels of medical unemployment without altering substantially the 'status quo' amongst the medical profession.

Moreover this view has been supported by the fact that once central approval of staff increases is made (probably after the threat of public complaints and patients' distress is considered at the bargaining table) expenditure can be put forward, allowing for an automatic enlargement of the external financial constraints imposed on the hospitals. Chances of job promotion and grading drifts can consequently increase, too.

Through the budget reimbursement system the process ends by transfering any conflict of resource allocation within the hospital sector onto changes in the amount and utilization of labour and material inputs. Given the boundaries within which the process operates at the institutional level, and given the workings of the budget reimbursement system in itself, the phenomenon of variable productivity can be seen then as a natural outcome.

2.5.2. Relevant Behavioural Factors.

According to the previous analysis, the following behavioural factors appear to have potential relevance for explaining the issue of variable productivity in the Spanish Hospital Health Care Sector.

a) The relative level of physicians' earnings in the particular area where the hospital is located. Although no precise figures are available in this respect, the geographical variation in the index of provincial cost of living, and the evidence available in this direction derived from the survey of physicians' earnings before mentioned, seem to support the importance of this factor.

b) The differential extent of the possibility for hospital physicians to exercise different forms of compatible medical practice.

Both former factors would account for the level of earnings from the private practice and, consequently, they would be reflected into differences in the opportunity cost from exercising public activity.

c) The initial level of slackness in hospital expenditure, on which successive budget increases come to operate, with regard to the whole set of individual responses of the Hospital Boards in facing budget restraints. Mainly, this relates to such factors as the degree of managerial habit formation, which notion prevails as the "maximum allowable request", the degree of risk-aversiveness of the Medical Director and Hospital Administrator to commit additional levels of expenditure, how strong are the bargaining powers of the territorial Health Authorities, etc. This is obviously due to the fact that the impact of those behavioural factors depend on the levels of 'discretion' allowed by current levels of expenditure on which budget increases apply.

Since an 'engineering' approach testing for the influence of these joint factors on the hospital performance is obviously difficult, and given that definitive conclusions would require very detailed analysis and managerial experimentation, we explore at this stage of our research only some suggestions in support of the former factors in a tentative way. This is,

(i) with respect to the wide range of variation in some ratios of utilization of hospital inputs, throughout measures or some simple workload indicators, as done in the Appendix to this Chapter, and which can serve to point out the extent of the issue of variable productivity;

(ii) with regard to the operational mechanisms of variable hospital productivity; this is, about some changes in manpower inputs over time with reference to the evolution of some work-load indicators, and with regard to the differential gap between wage increases[15] and the increases in the cost of living or in the implicit deflator for private health care expenditure (as alternative proxies used for collecting the actual increase in the physicians' private earnings).

The evidence provided in Table II.2, although still limited, can be used in support of this last point.

Therefore, the suggestion advanced earlier with regard to the mechanisms of the generation of variable productivity seems to be well grounded in reality. As it can be seen in Table II.2, the increase in wages of hospital registrar doctors (who constitute the most important category of hospital physicians - although the utilisation of other categories does not change the validity of these results - see table of footnote 15) - have left far behind the general increases in the index of cost of living and of the implicit deflator of health care consumption for the periods studied. At the same time, the increase in the number of hospital physicians has risen far more quickly than the total number of beds, patients or patient-days, both in the case of the total Social Security Hospitals and of the General Hospitals included in the sample studied for the periods between 1975 and 1980, and from 1977 to 1980, respectively.

TABLE II.2

Wage restrain and the evolution of some hospital indicators

	Increase in the total wage bill of hospital registrar physicians, in percentage terms, over the former year. (1)	Increase in the general cost of living. (consumption prices). (2)	Increase in the implicit deflator for health care consumption (3)	1-2	2-3
1974	11.5	15.7	17.4	-4.2	-5.2
1975	19.2	17.0	10.4	2.2	2.2
1976	13.3	17.7	21.5	-4.4	-3.0
1977	20.5	24.6	21.5	-4.1	-4.6
1978	16.0	19.8	17.7	-3.8	-4.1
1979	11.0	15.7	15.0	-4.7	-4.0
1980	12.5	15.5	12.1	-3.0	0.4
1981	12.5	14.6	14.3	-2.1	-1.8

Evolution of some hospital indicators

	1975/1980: Total Social Security -Percentage increases-	1977/1980: Social Security General Hospitals -Percentage increases-
Number of: patients	27.3	16.52
Bed-days	27.9	16.58
Beds	34.3	14.23
Hospital physicians	47.1	21.82
total hospital manpower	33.0	14.34
Total S. Security physicians	23.7	-
INSALUD hospital nurses & other clinical auxiliars	22.8*	-

* - Given data available, figures refer here to the period 1976-81

Source: INE, INSALUD, National Accounts Statistics.

2.6. THE PATTERN OF THE ORGANISATIONAL CHANGE AND SUMMARY OF THE CHAPTER.

The phenomenon of variable productivity, in the sense defined in this Chapter, is an extended feature -not always recognised in the literature- of the performance of institutions, and it appears to be particularly relevant to the Spanish hospital Health Care sector.

This Chapter has offered first a definition of 'relative efficiency' in operational terms - which provides thereafter a well grounded empirical approach for its analysis - and it has explored a set of behavioural factors which can explain the issue of variable productivity in the Spanish hospital sector and their mechanisms of operation. Although an appropriate test for the workings of these factors is still difficult at this stage, we have provided some evidence that seems to support the basic line of the argument.

In sum, we conclude that in the absence of a rational objective and an explicitly stated budget constraint imposed on hospital activity, hospital expenditure has been mostly controlled by a short-sighted incomes policy. This has left physicians' wage increases far behind the general rise in the cost of living. At the same time, the existing problems in the sector have been left untackled. A situation has existed where there is an extended lack of monitoring of the working-time schedule and disregard of any detailed evaluation of hospital expenses. In this context, the phenomenon of variable productivity can be seen to be a natural result of the agents' selective rational behaviour. In fact, hospital agents and physicians in particular may have responded to these central policies of restrain by relaxing their work-effort, following to some extent the external guidelines given by the opportunity costs of exercising the public practice in terms of the earnings achieved from private practice.

The way in which hospitals have obtained their revenues has also been decisive for this result. Thus, the actual levels of slackness have been consolidated over-time by lobbying from the current expenditure positions, and by using demand factors to influence the supply of hospital inputs by arguing for changes in the manpower levels of staff employed.

In this context, a change in the present organisational features and budget arrangements is needed in order to provide a rational framework for efficiency motivation in the decision-making process of hospital allocation of resources, and hence in hospital performance.

Some changes at the institutional level have been recently proposed in the first draft of the Spanish Bill for Hospital Reform (see footnote 11 in Chapter I) and, in short, these consist of:

(i) Facing the peculiar relationship between the financers of hospitals (INSALUD) and the hospitals themselves (mainly, because INSALUD is a highly centralized body, unable to monitor the performance of the hospitals) by:- Basing any institutional change on the new regionally decentralised INSALUD organization and on the constitution of an Area Health Council with a representative Area Health Director and an effective Area Health Manager.
- Framing the hospital management aspects into the attribution of full autonomy and responsibility to the Directive-Management-Technical Team of the Hospital (section 3.1 of the Bill).
- Allocating the responsibility for setting monitoring controls at the subcentral level, basically through a hospital information system adjusted to management orientated goals, and the implementation of analytical accountancy for cost centers.

(ii) Confronting the exclusive dominance of the physician in the hospital decision-making process (i.e. against the fact that Hospital Boards are dominated by the medical profession, with administrators as clerks of the Hospital Boards, not vested with any significant role in the decision-making process) through:
- the separation of the lines of Medical Direction and Hospital Management, assigning a supra-role to the Hospital manager responsible for the achievement of the institutional targets on the basis of the means centrally allocated.
- The restatement of the functions of the 'Chiefs of Department' in the direction of assuming administrative tasks and financial responsibilities in their Departments for limited time periods and never for life.

- The inclusion of political and citizens' representatives in the Area Health Council.
- The presentation of a Chart of the Users' rights and duties (Sept. 1983).

(iii) Proposals against the widespread existence of 'pluriempleo' amongst the medical profession (the existence of which has enabled physicians, according to our hypothesis, to vary their total remuneration by varying the hours of work in the public and private sector):
- The Bill for the Hospital Reform proposes changes in the working conditions scheduled with the inclusion of morning and afternoon duties and with the total working time in the hospital coordinated with the requirement of filling overnight duties (instead of those being instrumented as a way of achieving an extra compensatory remuneration from the public practice, as it is at present) (points 3.9.1 of the Bill proposals).
- Changes in the working status of the doctors towards; a) the exclusive or full-time exercise, non compatible with the private medical practice; b) the non-exclusive commitment or the possibility of exercising both compatible practices; c) the appointment of part-time associated or consultant physicians. Different levels of remuneration are assigned to each status (points 3.9.2.).
- The implementation at the departamental level of some standards of activity and quality for the hospital health care provided, monthly controlled by the Direction Team of the hospital.

For the second type of changes - that is, those refering to the nature of the existing budget-based arrangements - although some general points are made in the proposed Bill.(mostly in the sense of allocating budgets according to the responsibilities and the management objectives assigned) not much insight is provided in the Bill. This leaves open the problem of the efficiency motivation in which the former behavioural changes can be expected to operate in practice, and indicates the existing interest of research on the different strategies for reinforcing the budget constraint for revenue planning purposes: That is, the concern in providing an objective and prospectively stated financial constraint, capable of being inserted into the hospital decision-making process as a major watchdog for efficiency, as this thesis will pursue.

The core of this alternative policy will be explored in the next chapter, in which a budget based-output sharing system will be presented as optimal in the context of the Spanish hospital decision-making process.

FOOTNOTES.

(1) The overall sample, and subsamples selected for the initial comparison refer to 1979. For the static purpose in this context, the overall sample was scrutinized for certain observations which presented some sort of operational irregularity for that year; say, not being in full operation, suffering reforms, etc. The main results derived from the analysis of the sample were afterwards validated with respect to the rest of 'Functional-Economic Reports' of INSALUD for 1980 and 1981, already available. They in fact proved that the sample selected was fully representative and stable over the period.

(2) Similarly, pharmaceutical and hospital medical treatments could be also considered as intermediate inputs for the final outcome of improving the health states of the patients, as,for example,health education, preventive care or community life style might also be considered. Moreover, hospital institutional performance for these purposes can also be framed within certain aspects of alternative provisions, such as those medical services provided in outpatient clinics, day-surgery centers or nursing homes.

(3) 'Technical efficiency' refers to the utilization of the minimum input per unit of output for each input in the firm; 'price efficiency' can be defined with respect to the extra costs incurred by using a combination outside the optimum tangency point on the isocost line; 'allocative efficiency' refers, on the other hand, to a given economic arrangement (mix of inputs or output combination) for which there can be no rearrangement which will leave someone 'better off' without worsening the position of any one else; finally, the notion of "grand efficiency" introduces explicitly questions of equity in the allocation of resources.

(4) See for this purpose M. DENNY and M. FUSS. University of Toronto Working Papers no. 8202, 8131 and 8018.

(5) In principle, some targets can be defined (i) in terms of access to inputs; (ii) in terms of the resulting outputs (this second view would need to adjust not only the different purchasing power of resources in the different geographic areas but also the productivity of the resource utilisation throughout the different communities); (iii) in terms of the outcome achieved from the output provision, or the consequences derived from its utilization (e.g. regarding the impact on the 'client states' for different groups and regions).

In addition, they can either use (as the denominators of the indicators) 'per capita figures', a concept of 'need' or changes in individual health states. We will not explore this problem which can be seen in G. MOONEY (1980).

(6) In this line, see the article by MILLWARD and RAINEY "Don't blame the bureaucracy" in the Journal of Public Policy, 3.2. 1983.

The same argument is in P.M. JACKSON "Political economy of the bureaucracy" Philip Allan Publisher, Oxford 1982. Writing about this, Jackson points out: 'as a theoretical exercise, the comparison of market structures always permits perfect competition to win over the monopoly since the game is played with dice loaded by tautology. It will always be the case that perfect competition will win by the definition of efficiency we start-off with. However, the technical conditions required to establish a general competitive equilibrium are very stringent. In practice these conditions do not exist' (page 207).

(7) This is basically the view of A. J. CULYER, A. MAYNARD and A. WILLIAMS in "Alternative Systems of Health care provision; An Essay on Motes and Beans", in M. OLSON: "A new approach to the economics of Health Care" A. E.I.Washington, 1981; or in a more recent article of A. J. CULYER "Health Care and the Market: A British Lament", Topics in Health Economics no. 3, 1982.

(8) In "The Geography of Public Finance", R. J. BENNETT introduces tentatively the question of differences in spatial productivity "concerning the relative efficiency with which public services are provided in different jurisdictions". Productivity 'should' differ as a result of technical factors, such as community size (reflecting scale economies and density) but, in addition, differences 'may' result from the different degree of effectiveness with which some jurisdictions combine factor inputs to produce a given level of quality of public service outputs. This result can be derived in this context at least partly from the differences in incentives operating on local governments to be more efficient, from local political pressures, etc.. Productivity differences also may derive from differences in preferences for service organization or from some differential degrees of public pressures and citizens' participation, and again, differences in productivity may also differ due to the existing institutional organization.

In the applied field it should be recognized the contribution of M. DENNIS and M. FUSS in "Intertemporal and Interspatial comparisons of Cost Efficiency and Productivity". Univ. of Toronto Working Papers, 1981. The question of the need of assessing common forms of the prodution function for comparison purposes has been mathematically treated by JORGENSON and NISHIMIZU (1978) by building a "transcendental logarithmic function". In fact, its quadratic approximation requires only some few second order coefficients in the components of the comparison to be identical allowing, however, for wider differences in the constant and linear coefficients in the specification of the production functions compared.

(9) Unlike the theories which suggest that regional problems will disappear automatically because of the operation of market forces, the theory of cumulative causation maintains that the prior level of productivity is in itself an important contribution to its growth, leading to higher increases of productivity which raise thereafter the former levels and allow for further increases in their rate of growth. This theory is due originally to N. KALDOR in "Causes of the slow rate of growth of the United Kingdom. An inaugural lecture". Cambridge University Press, 1966.

(10) At the micro economic level, this tax-price component would make possible, in principle, some margins for direct market simulation: particularly in those areas in which the 'benefit principle' is applied.
The function of a pricing scheme would be to determine how the costs of financing the public good are to be shared amongst the users. However, while on efficiency grounds there is no reason to prefer the benefit pricing rule to any other financing scheme, on equity grounds these may exist if value judgements are made about the resulting distribution effects.
That market simulation is not, however, in our view, helpful from an operational point of view, given the set of specific assumptions required and which are mostly related to the notion of the marginal benefits raised from public expenditure and the marginal costs derived from the resulting tax distorsions.

(11) These authors basically note that most of the former literature on grants has derived the local response to each specific type of grant from the nominal or legal requirements involved in the structure of the grant itself. However, this is an assumption which needs to be empirically proved and it is most likely false. The effects of a grant depends critically on the structure of the local decision-making process, which ultimately determines the allocation of resources. For example, budget maximizing bureaucrats may react against the way in which grants are allocated by keeping

the citizens on their initial bundle of non-public goods, and taking any difference as a net transfer to be devoted to enlarging the amount of goods provided by the bureaus. This results in an inefficient allocation because of the change in the relative prices of the initial basket of goods up to the level of the expanded amount in the goods provided.

In a similar way, by introducing the assumption of the existence of citizens' fiscal illusion, the problem can be analyzed from the supply side, with agents capturing all the potential consumer surplus through whichever expansion in the output levels is made.

Therefore, how the state and local sector responses are influenced by the central aids of various sorts is an empirical question. What it is required is a theoretical framework that incorporates explicitly the decision-making process of the institutions which are recipients of the grants together with the own structure of the grant.

(12) As a consequence of what we expressed in the footnote 11, an adaptive response may be produced on the local fiscal behaviour, given the way in which the grant impacts on the expenditure side of the budget. This is studied particularly once the dominant approach to the determinant factors of local government expenditure is abandonned and the place is occupied by the institutional approach. The influence of that impact gives rise to the study of incentives for local governments to provide services more efficiently, stimulating additional expenditure or evaluating its current levels.

The stimulative effects of intergovernmental grants, and the potential substitution of the otherwise revenues locally raised, have been specifically studied by P. N. COURANT; E. M. GRAMLICH and H. GALPER (1973). For the British case in particular, ASHFORD et al. (1976), B. LYNCH and M. PERLMAN (1979); K. CUTHBERTSON, J. S. FOREMAN-PECK and P. GRIPAIOS (1981) and K. SCHOTT (1982) have presented some applications.

(13) At any rate, it is rather difficult to determine the precise meaning of this 'minimum level'. This problem is similar to that of fixing the notion of 'maximum earnings' in the case of the physician's profit maximization hypothesis. As R. G. EVANS (1974) has pointed out, that 'maximum' can be defined a priori, more or less reallistically, by introducing, for example, a 'medical ethics' notion into the physician objective function, defining a labour-leisure trade-off, or by redefining some levels of physician utility satisfaction.

In the same way that the profit maximization hypothesis could represent the 168 hours of work a week, the utility maximization hypothesis in our context, with the wage level being regarded as some sort of given minimum, would obviously lead to the supply of a zero level of effort.

(14) This benchmark in fixing the amount to be earned is argued to be based on social considerations about status and dignity in what is felt to be, as a whole, a 'hard work' activity for which a poor remuneration is offered during the time of the junior physician training. In fact, the survey shows that physicians view themselves as being in a continuous position of service (between 50 and 60 hours a week), and that they feel that higher incomes are a fair compensation for the time and effort the devoted yesterday to their medical training, and today to the exercise of their profession.

(15) The increase in the overall remuneration of hospital physicians (thousand pesetas) can be seen in the following table:

	1973	1974	1975	1976	1977	1978
Head of Department (6 hours/day)	-	11.7	19.6	10.9	10.9	13.0
Head of Service (6 hours/day)	-	11.7	19.6	11.1	11.9	13.4
Head of Section (6 hours/day)	-	10.0	19.4	9.8	17.0	14.3
Registrars (6 hours/day)	-	11.5	19.2	13.3	20.5	16.0
General increase in the cost of living	11.5	15.7	17.0	17.7	24.6	19.8

Source: Own elaboration from INE and INP data (today 'INSALUD').

For indicative purposes, we should distinguish the share of the three main components of the physicians' wage remuneration. Taking figures from 1980, these were their basic proportions:

Staff categories	'baseline wage'	'position complement'	'extraordinary payments'	Total
	%	%	%	%
Chief of Department (6 hours)	26.3	55.5	18.2	100
Chief of Service (6 hours)	27.4	53.2	19.4	100
Chief of Section (6 hours)	35.0	49.8	15.2	100
Registrar (6 hours)	45.8	26.7	28.5	100

'Registrars' is here used as the equivalent to the Spanish category of 'Adjunto'.

APPENDIX 2. 1.

As it can be seen in the following tables, although an effort was made to group the initial observations belonging to the overall sample of "Health Residences" into subsamples with higher "within group" homogeneous characteristics and a lower "between group" degree of homogeneity (as described in Chapter V), the values of the coefficient of variation for some basic indicators remain relatively high, far above those which could be expected from some statutorily regulated condition of provision.

TABLE 1

Year 1979	Overall Sample	Subsample Selected	Only teaching 'Health Residencies' (HR)	Only non-teaching HR	HR with out-patient services	HR without outpatient services
Number of beds available per physician						
- Coef of Variatian	-	27.6	21.6	30.6	24.0	45.5
- Mean	3.48	3.55	3.90	3.27	3.58	3.41
- Standard Deviation	-	0.98	0.844	1.00	0.86	1.55
Number of nurses per physician						
- Coef of Variation	-	27.4	19.3	31.3	24.9	38.9
- Mean	3.46*	3.725	4.175	3.359	3.687	3.930
- Standard Deviation	-	1.021	0.805	1.046	0.918	1.529
Average cost per case (pts)						
- Coef of Variation	-	23.2	19.8	25.9	20.8	23.3
- Mean	75747	71034	73493	69013	68142	86577
- Standard Deviation	-	16468	14540	17904	14157	20176
Average cost per patient day (pesetas)						
- Coef of Variation	-	20.3	16.3	23.3	17.2	18.6
- Mean	11200	8010	7979	8036	7654	9922
- Standard Deviation	-	1623	1301	1870	1320	1848

Source: 'Informe Económico - Funcional de los Instituciones Sanitarias'. 1979

(*) - This figure has had to be referred to 1976, given existing data available.

We explored afterwards a greater range on indicators for the final subsample selected:

TABLE 2

	Mean	SD	Coefficient of variation
Average length of stay	8.94	1.53	17.11
Bed Occupancy rate	77.68%	14.45%	18.60
Total personnel per bed	1.81	0.93	51.4
Number of cases per personnel			
- per physician	114.79	32.54	28.3
- per nurse	31.90	8.84	27.7
- per ancillary worker	62.55	24.15	38.6
- total	17.42	4.72	27.1
Number of beneficiaries and affiliated per total personnel	468.16	197.9	42.3
Average cost per bed available (thousand pesetas)	2317.3	741.8	32.0
Average cost per case on (pesetas)			
- pharmaceutical inputs	4447	2799	63.0
- surgical inputs	1931	1840	95.3
- Lab. inputs	1092	639	58.5
- radiology inputs	567	308	54.3

Source: 'Informe Económico Funcional de las Instituciones Sanitarias'. 1979

Finally, we analysed some performance indicators for the different categories of staff in the subsample finally selected.

TABLE 3

ANALYSIS of some PERFORMANCE INDICATORS per HOSPITAL STAFF

Staff Categories	Number of beds available per each staff Category — Sample values (1979): Mean	SD	Maximun Value	Minimum Value	Coef. of Variation	Total number of: Mean	SD
RANK A (Physicians chiefs of Medical Services and Section, and Pharmacists) . . .	8.81	2.36	14.8	3.33	26.8	9.2	2.41
RANK B (Registrar physicians and Assist. Pharmacists).	6.52	1.73	11.8	2.71	26.5	6.8	1.75
RANK C (Junior physicians)	50.	50.	—	237	—	54.	50.
A + B + C	3.5	0.92	5.83	1.38	26.3	3.65	0.91
Rank a (Midwifes and physioterapists) . . .	26.4	8.0	54.5	10.	30.3	27.6	8.14
Rank b (Nurses, ATS) . . .	1.87	0.41	2.95	0.57	21.9	1.94	0.37
Rank c (Clinica Auxiliar) .	2.32	0.64	3.93	1.01	27.6	2.44	0.71
a + b + c	0.97	0.20	1.43	0.40	20.6	0.97	0.19
Rank x Ancillary workers (technical and special services) .	8.6	2.9	18.3	3.11	33.7	9.0	2.8
Rank y Professional and clerks. . . .	16.2	18.0	81.7	3.88	—	16.8	17.9
Rank z Ancillary workers (auxiliar and 'for services') . .	3.2	0.95	5.84	1.23	29.7	3.19	0.95
x + Y + z.	1.91	0.60	3.63	0.90	31.4	1.91	0.6
TOTAL PERSONNEL	0.53	0.10	0.77	0.23	18.9	0.55	0.09

beds per each staff category			Number of patients treated per each staff category					Number of potential beneficiaries and affiliated per each staff category				
Sample values (1979)			Sample values (1979)					Sample values (1979)				
Max Value	Min. Value	Coef. of Variation	Mean	SD	Max. Value	Min. Value	Coef. of Variation	Mean	SD	Max. Value	Min. Value	Coef. of Variation
15.8	4.7	26.2	288	77.6	443	128	26.9	7562	3030	14280	1936	40.1
11.8	3.54	25.7	213	54.2	347	91.7	25.5	5608	2313	10984	2369	41.2
—	208.	—	1666	2000	—	6516	—	38226	47121	—	16581	—
5.83	1.92	24.9	114	30	187	49	26.3	3016	1242	6208	1005	41.2
54.5	15.4	29.5	872	306	1852	395	35.1	23924	13765	70432	6915	57.5
2.95	1.27	19.1	61.8	16.8	102	26.5	27.2	1637	815	511	785	49.8
4.6	1.02	29.1	76.2	22.4	127	35	29.4	2027	976	5469	531	48.1
1.43	0.40	19.6	32.1	8.16	53.4	167	25.4	849	374	2108	306	44.0
18.3	3.1	31.1	287.9	113.1	591	85.3	39.3	7567	3711	16977	2298	49.0
31.7	5.1	—	572	696	28.38	84.16	—	15218	19878	105648	2732	—
5.84	1.24	29.8	106	38.6	148	54	36.4	2783	1363	7546	931	49.0
3.63	0.90	31.4	64	26	127	25	40.6	1683	880	4695	692	52.3
0.77	0.36	16.4	17.6	4.4	28.8	83	25.0	464	198	983	184	42.7

CHAPTER I I I A BUDGET POLICY AIMED AT PROMOTING EFFICIENCY IN THE HOSPITAL SECTOR.

CHAPTER III

A BUDGET POLICY AIMED AT PROMOTING EFFICIENCY IN THE HOSPITAL SECTOR.

3.1. INTRODUCTION.

We have seen in the previous chapters that there exists a need to give INSALUD the scope to provide hospitals with incentives to be efficient, and that the existing budgetary mechanisms cannot be used as a means of promoting efficiency. The question to be answered in this chapter is how could we do this.

The Outline of the Present Chapter is as follows:

We will examine first which factors need to be taken into account in designing a goal-orientated budget policy and, particularly, (i) which is the appropriate setting for the policy; that is, how should management delegation and the budget relationship be regulated; (ii) whether we can specify within the resulting framework, a policy able to raise incentive compatibility on both sides of the delegation.

For this, we need to recognize that the present relationship, while treated as a hierarchical type of relation, is to all intents and purposes, an Agency-type of relationship (section 3.2.1.).

This recognition will imply, as we will see, the need to set up the terms of an optimal reward structure, capable in principle of producing incentive compatability (section 3.2.2). According to the Agency theory, some form of output-sharing can be expected to be appropriate for this type of situation. Thus, we will search for the optimal financial arrangement which, together with the insertion of incentives for efficiency motivation, could contribute to changes in the work-effort over and above the minimum levels (section 3.3). Given that efficient output-sharing can both create incentive compatibility and lead to a work-effort in excess of some minima, the next step is to study what type of output sharing is 'optimal' for our purposes. That is: how we could design a form of output-sharing capable of modifying behaviour over time (removing X-inefficiency) and accounting for the phenomenon of variable productivity (section 3.4.1).

We will take Malcomson's result on internal incentives and wage hierarchies, applied to our modified model for a reimbursement policy under differential levels of performance, to prove that this policy for internal incentives could in fact be constructed and could be optimal under a specific set of objective rules. (section 3.4.2).

Several implicit assumptions are made in deriving this result (e.g. that a suitable unit of "work" can be observed and measured with sufficient accuracy). We will examine then whether all these conditions will hold in the hospital sector (3.4.2) and, if this is the case, how, according to the theory, the use of a budget-based contract as a form of output-sharing could modify hospital behaviour over time. The properties of this budget-based contract will be derived mathematically and represented in diagramatic terms for illustrative purposes.
Section 3.4.3 explores the way in which such a budget-based output-sharing system could be operationalised in the Spanish hospital sector. Particular attention is paid to issues of monitoring and measurement and to some general considerations about the behaviour of the agents, with respect to the characteristics of the output-sharing system, to ensure that their reaction will work in accordance with the system goals.

Assuming that we can operationalise the system, the question that then remains in our actual context is "What the likely effects would be, (or would have been) if the policy were (or had been) implemented?".

A simulation and other tests will be presented for these purposes (section 3.5) prior to the specific design of a Prospective Revenue Allocation policy which addresses all the former points. This is the objective of this chapter.

Summarizing this third chapter aims to provide an insight into the economics of a Prospective Reimbursement Policy which, being linked to some observed levels of performance, is orientated finally towards the correction of some cost deviations, as reflected by the notion of variable productivity.

3.2. FACTORS TO BE TAKEN INTO ACCOUNT IN DESIGNING A BUDGET POLICY.

Under the existing arrangements, the budgetary mechanisms in the Spanish hospital sector cannot be effective in the promotion of hospital efficiency. This would seem to be due to three factors:

(i) INSALUD being organized as a 'pyramidal' structure, based on a hierarchical communication network, cannot monitor the performance of its hospitals closely; even if it were capable of doing so, the inforcement of its directives is likely to be a rather complex and costly task.

(ii) Hospital physicians would have no incentives to honour the directives laid down by INSALUD, given the scope for discretional decisions which remain in their hands, and the technological knowledge available to them for the exercise of the profession with respect to the utilization of inputs 'scientifically' required.

(iii) The payment to physicians of a fixed wage remuneration and the cost-based retrospective nature of the hospital reimbursement system, (which constitute the financial environment surrounding the hospital decision-making process) cannot provide the incentives required for closing the gap that exists between technical knowledge and motivation for improving the hospital performance.

Therefore, it seems obvious that there is a need for a change in the basis of the relationship between hospitals and INSALUD if an appropriate budget policy, aimed at promoting efficiency by changing agents' behaviour, is to be implemented successfully.

Two main factors need to be taken into account in the design of a goal-orientated budget policy. First, what is the appropriate basis for framing the relationship analysed. That is, which is the setting in which the actual network or budgetary inter-relations occur, regarding particularly the scope of the decisions and powers of both sides of the delegation. Second, whether any problem of incentive compativility is present and, if this is the case, whether it can be tackled through an efficient reward structure, grounded in a budget policy.

Both points are examined in the next subsections.

3.2.1. The appropriate setting for the policy.

If there is to be a successful implementation of a budgetary system aimed at improving hospital efficiency, then clearly there needs to be an appropriate regulation of the relationship between INSALUD and Hospital Boards. As described in previous chapters, the present inter-relation is regulated statutorily within a hierarchical system, with unity of command and spanned control of management. This is done via central directives, enhanced by the INSALUD territorial delegates.

In this context, if INSALUD is going to have some impact on hospital behaviour, it needs to have the power to enforce its directives on policy targets set up by the Health Authorities. This has proved virtually impossible to achieve through the existing hierachical network, which fails to recognize the nature of the actual relationship between hospitals and INSALUD (more similar - as we will see - to an Agency type of relationship than to a hierarchical type of relation). This is mainly due to the following factors:

(i) Less than full information is available to INSALUD Authorities about hospital agents' performance. The nature of the output provided and the technological information barrier which are used frequently as justifications of the actual input utilization, are two factors which make performance particularly difficult to check.

(ii) In addition, problems of behavioural dependence can accompany those problems of informational dependence, breaking the postulates of the efficient communication network required for the operation of hierarchical structures under conditions of no-clear observability of the outcomes and no direct monitoring of the inputs utilized by the agents.

(iii) As a result, hospital agents, performing under less than completely specified contracts, come to dominate the behaviour of the institutions. In these contexts, only the identity of interests between parts can guarantee the achievement of the organisation's targets. This is, however, unlikely whenever rational agents take advantage of the behavioural and informational assymmetries support the power with which they are vested.

3.2.2. Incentive compatibility and the efficient reward structure.

However, the current regulation of the institutional relationship has failed to recognize the existence of the efficiency motivation problem. Indeed, hierarchical structures, by nature, do not treat specifically the issue of the potential shortcomings caused by the lack of the identity of interests between the individuals and the organisations. Economic institutions are monitered to perform in these contexts as a sort of neoclassical "black-box" in which individuals behave in a blind mechanistic fashion or, alternatively, operate within a context in which rewards are expected to stimulate the individuals productive response, preventing any form of economic "shirking".

This view of economic reality exhibits ovbious deficiencies, either because of failing to recognise the factors (i) to (iii) mentioned in section 3.2.1 or due to the fact that, in practice - and in particular in the hierarchical context - wage rates do not relate to marginal productivities and, therefore, market rewards cannot provide 'per se' incentive motivation to the agents (Lazear, 1979). In these cases it remains to be explained why any goal is ultimately subscribed to by the individual agents, and then, the problem of incentive compatibility may need to be explicitly treated and taken into account for the design of an appropriate budget policy for promoting efficiency amongst revenue recipients.

In fact, the problem of incentive promotion, and the efficient reward structure to achieve it, can be considered a sequential one. When the incentive problem appears - e.g. when management has only imperfect information about the abilities and willingness of individuals to work, and the outcome of the Agency (but not eh Agents' work effort) can be observed - the reimbursement policy needs to address, in some way, the question of efficient funding. Otherwise, postulates of rationality may involve a constant perverse incentive for reducing the agents' work-effort instead of moving it to some targeted outcomes. As Alchian and Demsetz have pointed out (1), in these circumstances each individual may be able to adjust rather freely his work-effort with respect to the realized reward, so as to equalize the marginal rate of substitution in consumption.

In other words, under less than complete contracts and a lack of performance monitoring, each agent will feel induced to take more leisure, because the effects of relaxing his rate of substitution between output and leisure will be less than the effect on the true rate of substitution. Thus, any contract that would pay a pure wage in these circumstances would involve no incentive to honour the agreement, since the worker could take his wage as given and select an effort level which tends to zero.

Given then the entire set of factors surrounding the institutional relationship under consideration, the Agency type of relationship is thought to be the most appropriate setting for analysing the nature of the delegation between the INSALUD and the Hospital Boards. This approach does obviously need some changes in the current statutory and financial arrangements which accompany the regulation for the allocation of overall value responsibility to the decentralized agencies (i.e. holding the funding at the central level), with the purpose of providing the basis for the promotion of efficiency improvements in performance.(2) In addition two questions need to be dealt with:

a) which postulates can support in this relationship a model of incentive compatibility for the achievement of some common goals?.

b) What is the optimal reward structure on which to base the required budgetary arrangements?.

The first question is analysed in the next pages within the framework provided by the theory of the Agency relationship, with central funding and attribution of overall responsibility to the decentralized agencies in the sector. This is considered both generally and in the specific case of hospital units within the health care sector. The second problem will be solved in the context of the intra-sectorial financial agreements, verifiable by those parties involved in the relationship.

They will be presented thereafter as the basis of a wider theory of resource allocation and, in particular, of the design of a Prospective Reiumbursement policy.

In this context, we will consider the potential effects of a change in the funding methods of hospital expenses aimed at breaking the causation of system inefficiency. The new funding methods will introduce incentives for efficiency improvements and will seek to control, in addition, the factors which are thought to cause variable productivity.

In sum, the formulation of an incentive policy for a cost-conscious management of hospital resources can be designed from a budget policy which focusses on the Agents' compensation for a given work-effort and, more widely, on the Agencies' reimbursement according to performance.

These points will be examined in the next sections.

3.3. CENTRAL FUNDING AND HOSPITAL PERFORMANCE UNDER THE AGENCY RELATIONSHIP

In light of the discussion above, the relationship between hospitals and INSALUD Central Board can be better framed from the view of an Agency type relationship, where the first (designated as the Agents) act on behalf of the second (or the Principal) in a particular domain in the hospital decision-making process, and mainly in the case of recurring decisions and sequential transactions involved in hospital activity.

In this framework, the principal is the one statutorily nominated with the responsibility for the provision of the service and for the overall management of resources. Whether this is designated as the Central INSALUD Board, the Regional Health Authority or any other more decentralized level of management, it does not change the nature of the principal. However, it has consequences for the type of monitoring which can be exercised, given the amount of information available in each of those organisational steps.

The agent in our model is, in principle, the Hospital Board. In reality, given the nature of the internal management and organisation of Spanish hospitals, the hospital physician can be considered the ultimate unit of analysis. Thus, the Agent-physician is the unit supplying work-effort in terms of medical services to produce 'output'.

Several reasons exist for recognizing the Agency relationship to be most appropriate. First, the enviroment surrounding the relationship, as we have seen, creates problems of information access and 'adverse selection', and hence involves behavioural dependence from the agent activity (or the moral hazard issue). This is due fundamentally to the fact that the extent of the individuals' choice of the time-activity-pace cannot be fully observed.

Secondly, the Agency type of relationship is more appropriate, having seen the particular relation of interests and powers involved: That is, whereas the principal is vested with reimbursement powers and concerned with the cost-conscious management of resources, the Agents' objective can be defined mostly in terms of the remuneration achieved for a given work-effort. The lack of identity of interests in this context involves pure wage contracts being inefficient in motivating Agents, in a context in which a complete monitoring of the Agent activity is extremely complex.

But in defining in practice the nature of the particular type of Agency relation relevant in our context, some important questions need to be considered. They are:

(i) What are the predominant reasons giving rise to the Agency delegation (i.e. problems of information access or behavioural dependence) and which part of the relationship is most affected by the implementation of the control policy designed? (For instance, is it the improvement in the information available for purposes of greater control for further central monitoring -i.e. providing Agents with incentives to reveal information but keeping overall Agency autonomy - or the attempt to achieve direct changes in the work-effort supplied by modifying the agents' behaviour)?

(ii) The intrinsic character of the relationship with regard to the potential limits for cooperative behaviour (either with reference to an initial point or through the sequential search for an outcome equilibrium).

(iii) The nature of the delegation from the principal to the agent, and whether the long run operational point of view dominates for this purpose any short-term influence (given, for example, the time horizon of a public agreement with privately owned hospitals).

(iv) Finally, whether there exists a random element or risk-bearing factor for any of the parties involved in the relationship with respect to agent's productive knowledge and the outcome of his/her activity.

In consequence, according to the terms and the nature of the relationship defined, the principal enjoys the outcome of the agents' activity (i.e. the statutory provision of the service and the achievement of certain targets related to policy planning). For this, he needs to remunerate the agent appropriately, in order to motivate his work-effort and compensate him for the disutility he suffers from it. The reward structure must then cope with both elements of compensation and motivation, given the dependence of the outcome on the Agents' behaviour, in order to orientate their effort towards some planned targets of performance.

3.3.1. The Optimal Reward Structure in a Single Agency Relationship.

We have seen above that in the absence of non-fully observable work-effort, a pure wage contract which would pay directly a fixed wage for a specific task involving a particular level of effort might be sufficient for job motivation. However, there are several reasons why, in reality, the construction of a contract of this nature is not optimal when information assymmetries and behavioural dependence problems exist between the parties concerned. This is due, inter alia, to the fact that the non-observability of the agent's effort for purposes of work monitoring would allow the agent to take the salary as given and reduce his effort, consequently, towards zero levels. In addition, if uncertainty were present, agents would have no incentive to move from too cautious or too venturesome positions for a comprehensive analysis and a complete evaluation of risks.

In these contexts, many authors (e.g.Hurwicz, 1973) have pointed out the convenience of the definition of an efficient reward structure in terms of an output-sharing formula, with a common functional form dependent on some variable observable by both parties involved in the relationship.

Defined in simple terms, an output-sharing contract, with no random states influencing the final outcome other than the agents' work-effort, can be designed in the following terms:

$$U^A\ (\ Z,\ M^A\ ),\ \text{where}\ U_1^A < 0 \qquad U_2^A > 0,\ \text{and}$$

$$U^P\ (\ \theta,\ M^P\ ),\ \text{where}\ U_1^P > 0 \qquad U_2^P > 0,$$

where U^A and U^P are the Agent's and Principal's utility function, $\underline{Z}$ refers to work-effort, $\underline{M}$ income and $\underline{\theta}$, outcome, with $\theta = a.Z$ and $M = \gamma\ (\theta)$.

The agent's objective function would be then to maximize his utility,(or expected utility where an uncertainty element influences one of the arguments of the function), this being a positive function of his share in the outcome (compensation) and a negative function of the level of work-effort. Similarly, the principal's goal involves the maximization of his utility function, which is a positive function of the difference between the total outcome achieved and its specific share with the agent. The introduction of some constraints in this basic relationship would not change its nature - e.g. by requiring the provision of a minimal amount of output, or as a result of fixing a certain financial compensation over some basic amounts.

In order to derive a determinate solution to this initially joint maximization exercise, the two following approximations can be suggested:

(i) Approaching the solution from a bargaining view with the tools of the Game Theory, in which each individual party of the Agency relationship pursues his own interest, but no one can dictate the outcome.

(ii) Searching for a solution where one actor maximizes his utility, subject to the constraint that the utility of the other part of the relationship is not lower than some pre-specified level. This position can be defined as that stationary point which satisfies the first order conditions for an utility maximization or a 'selective rational' position with respect to the choice of the action.

Thus by confronting the problem of output-sharing as a bargaining situation with the tools of Game Theory, the existence of transaction costs, information assymmetries and the characteristics of the output provided do not suggest a clear-cut solution valid for operational purposes. For example, in the context of incrementalistic budget strategies of an interactive process type as a form of output-sharing, which has been tentatively studied by Wildavsky et al., some alternative solutions can be suggested with respect to the pattern followed by the budget requests, budget approvals and the budget amounts finally settled. In this context, a solution of the Von Neumann-Morgernsten type of "two persons non-zero sum game" or a Nash type of solution[(3)] are obviously unlikely, given the communication network required for an efficient inputation of all the alternatives available. In similar terms we could reject the Zeuthern-Harsany result, although it might appear as more appropriate in the present context, regarding the postulate that 'the bargainer who can afford the least risk of disagreement is the one to make the concession'. But this approach does not account for the 'needs constraint' operating as a minimum for service provision. Moreover, the possibility that initial uncooperative strategies may become in a sequential approach "cooperative" ones, with respect to the final equilibrium achieved, should be taken into account (Radner, 1981) since, as Cross[(4)] has pointed out, 'actors can exhibit expectations and learning over time, which may allow finally for an implicit agreement in the common interest'.

In our view, a constrained joint maximization type of solution can provide a more realistic approach: that is, in searching for a reward structure in which one of the parties is at a stationary point with respect to the choice of the action, or in a 'selective rational' position,(5) and the other maximizes his course of action under this constrained framework. A cooperative outcome then can be derived which is more in accordance with the postulates of the bounded rational behaviour described in Chapter I I. However, this constrained maximization solution seems to require a previous agreement by both parties in the relationship on the general terms in which this framework is specified.

This framework can be set in our modified context of an explict separation between the demand division - or that generated by the medical staff - and the supply division of Central Administration. Under these terms, budget-based contracts have already been presented in other fields or research as an efficient response to the existing dichotomy between the actual management and the control and ownership of modern economic organizations. As Fama (1980) has noted, this solution requires only that:

(1) The activity can be carried out under a minimum degree of discipline among Agencies (in competing, for instance, for a fixed amount of output or resources), and

(ii) some "market" for the recognition of agents' capacities exists with regard to the Agencies performance (e.g. with principals willing to compensate according to performance).

In this last scheme, the agent's compensation as a form of output sharing consists of two parts, one defined over the subset of favourable outcomes and another defined over the subset of unfavourable ones. The partition is specified in the contract with reference, for example, to a standard benchmark. This divides the set of possible outcomes (into favourable and unfavourable ones - although there is not a binary limitation for this purpose) and according to them, the agent's remuneration.

As a result, the outcome of the partition comes to depend, partly, on the relationship between actual and the pre-stated standard performance, with 'performance' expressed in terms of some observable attributes of the outcome derived from the Agents' activities.

This general setting for incentive rewards has been proposed in the specific field of public sector management by some authors - see Bergson (1978), Ben-Zion and Spiegel (1979) - with the proposal of working a system of compensation ofpublic officials, mostly based on a 'career function' competing for the allocation of jobs and levelled down by potential dismissals. They offer a 'bonus function' as a form of output-sharing or compensation according to the Agency's performance.

This incentive structure could be optimal for our purpose, under some more specific conditions which we will explore in the following section. In particular:

(i) The prior commitment to share a pre-stated amount (the total budgetary funds available to the sector).

(ii) The possibility of monitoring 'minimum' levels of effort as resulting, for instance, from average benchmarks).

Summarizing, we will study the potential effects of the re-stated relationship between INSALUD and the individual Hospital Boards in the context of the Agency relationship, selecting a budget-based contract which accounts for the second of the solutions proposed - this is, to be at a point of the interior solution for a "satisfying" position for one of the parties - and it is addressed to the problem of providing incentives for a behavioural change by allowing to the other party the maximization of utility according to a prospectively advanced budget contract. This provides a differential system of reimbursement with regard to some observed levels of performance, which is able ultimately to motivate a change in the work-effort of the Agent over and above some minima.

3.4. AN INCENTIVE REIMBURSEMENT POLICY UNDER DIFFERENTIAL LEVELS OF PERFORMANCE.

We saw earlier that any reimbursement policy aiming to provide a global and objective financial restraint under postulates of rationality needs to recognize the inital levels of differential performance of the institutions concerned.

In this section,we will examine whether an ouput-sharing formula which incorporates some proxies for the notion of variable productivity can modify behaviour over time, leading to an optimal budget arrangement. This will be proposed from a budget contract which provides two types of reimbursement according to the observed levels of performance under conditions specified prospectively. The properties of this model for changing the agents' behaviour are mathematically derived from the maximization of the agent expected utility, having specified first an output-sharing arrangement which satisfies the principal's selective rational position (Leibenstein, 1975). The diagramatic representation of the properties of the formula, and some other general considerations for its implementation, are presented at the end of the section. The outline of the model is introduced first.

3.4.1. Outline of the Model

The main hospital reimburser is defined here as the 'principal', funding the i hospitals ($i = 1...n$) over t periods ($t = 1...m$). With no uncertainty surrounding the Agent's decision-making process,(6) his actions or decisions D depend only on the technical knowledge and the work-effort Z . Assuming that the agents' skills are known (medical affiliation, etc.),this defines an outcome (e.g. produced under assumptions of constant returns to scale) such as:

$$D = D\ (\ Z,\ M\ ) \text{ and } \sigma = \sigma\,(D).$$

The utility function of the principal is defined in principle over the difference between the level of utility derived from enjoying the total outcome and the loss of utility from having to reward agents:

That is,

$$U^p = U(\theta) - U(M).$$

However, as we saw before, the reward compensation $\underline{M}$ is itself a function of the outcome, i.e. $M = \gamma(\theta)$. (This proposition is in the core of any budget-based contract).

To avoid indeterminacy from this joint-maximization interaction, we will focus on the constrained Agent's maximization of his utility function, from which the principal's equilibrium is defined residually. In doing this, we will assume that:

(i) Hospital agents of different hospitals behave in similar terms. Thus being each hospital dominated by the medical agents, they can be treated as the unit for analytical purposes: They may share a common utility function, at least within the subsample studied, with a similar type of basic argument influencing the relevant physician and Agency behaviour.

(ii) Consequently, the hospital's performance can be approached as the 'Agent' performance in the basic Agency relationship and the 'individual income' as the total reimbursement at the hospital level. Individual Agents (and hospitals) seek in this modified context to maximize their utilities in terms of the income earned against the work effort required.

(iii) Moreover the reimburser is able to make some type of general(and objective) assessment about the hospitals' performance using the available statistics - although possibly subjected to measurement errors - in order to partition the set of relevant outcomes with respect to the total amount available for reimbursement purposes.

Assumptions (i) and (ii) are natural consequences of the model of hospital behaviour presented in the second Chapter. Hence, either from an organisational approach or from a similar set of individual agents which influence the Agencies' behaviour, these assumptions can be plausibly sustained. (For instance, with respect to a dominant group of physicians sharing similar motivations, such as doctors with a given status in the staff hierarchy, exercising their profession in the private and public practice, etc.).

The assumption contained in (iii) will require some adequate estimates of output, but, as we will discuss later, this measure can be plausibly achieved for those comparative objectives.

Therefore, we can define a utility function for the hospital agent defined over the work-effort Z and the income M, such as:

$$U_i = U\,(Z_1 \ldots Z_t,\ M_1 \ldots M_t\ ),\ \text{with}\ U_Z' < 0\ \ U_M' > 0$$

Assume, for analytical convenience, a short term prospective revenue planning with t=1,2. This time factor will not change the essence of the model (although a period of four or more years will seem more appropriate for our purposes).

The reimbursing authority then advances a 'two period' programme with a prospectively specified reimbursement level M_1 for period 1 and a chance of adjustment for period 2, with M_2 or M_2' depending on the outcome achieved by the Agency in the first year. This entails two potential levels of reimbursement for the second period: M_2' for an individual hospital which is 'relegated' and, say, M_2 for one which is not. For example, M_1 and M_2 would imply the funding of all the initial requests within the general guidelines,(7) whereas M_2' would be a weighted average of actual and standard expenditure. For instance, if EA is the actual expenditure and EP the standard average or predicted one from the sample behaviour, the total revenue allocated (ER) can be stated as:

$ER = \beta^t.\ EA + (1-\beta^t).\ EP$, where $0 < \beta < 1$ and t refers to time. Other different combinations are however certainly possible.(8)

As we prove in the next section, by penalising those hospitals with a poor performance through relatively lower reimbursement levels, incentives may be created to increase the work-effort. Moreover, by allocating the defined budget revenue already committed in this way, expenditure will remain under control and on more rigorous bases of rationality for efficiency motivation targets.

A theory of resource allocation to the decentralized institutions can be based on a 'budget contract' under the former premises (Demski and Feltham, 1978). This can be written in verifiable terms by the parties, in the sense that the total amount to be distributed is initially comitted and the problem is reduced to that of its distribution regarding ultimately some notion about the Agents' (cost) effective performance.

The outline of this approach was first put forward by Malcomson (1983) in the field of the search for efficiency in a context of wage hierarchies. The elements involved in Malcomson's model are similar, in our view, to those involved in any process of budget revenue allocation: the core of the decisions required in this area also involves an initial commitment of the resources to be distributed and its specific distribution amongst the 'individuals' encharged of its provision. The nature of both decisions in the budget revenue allocation process can then be analysed in very similar terms to those described in the context of Malcomson's research.

3.4.2. The model at work.

In line with MALCOMSON's model of internal incentives for an efficient performance,[9] the effective work-effort $\underline{Z}$ is defined here by some proxy measures on certain observed values $\underline{X}^*$ depending on $\underline{Z}$, but with some random errors potentially associated with its definition. This will be done by fixing a proportion $\underline{P}$ of hospitals with observed values of $\underline{X}$ over the average values of the 'peer groups' considered. These hospitals will find their budgets initially reviewed and, in cases, submitted to the regulation policy in question.

Different indicators can be used for the definition of $\underline{X}$, mainly with regard to: (i) the average sample behaviour of some homogeneous groups of hospitals, with respect to certain relevant aspects included in the specification of the entire hospital production function - for instance, from a cross-section analysis in which we assume a potentially identical productive knowledge but a different realization in practice of their production possibilities; and (ii) the resulting values from a more accurate prediction of hospital performance or some other combined indication with regard to some disaggregated hospital functions, adjusted for certain external and internal factors affected to the activity.

The probability of relegation P will be defined as $P = 1 - F (X^* - Z)$, where $F (X - Z)$ is the cumulative distribution function for the observed effort, approached empirically.[(10)] Since P is defined in terms of number of standard deviations from the average values of X, assuming the presence of a normal distribution, once P is selected X^* is determined.

Hospital units are then faced with the necessity of improving performance (Z), moving towards X^*, or having to adjust themselves to the advanced policy of expenditure restraint.

A minimum management supervision inside the hospital to avoid utility levels falling below minima and a positive value of P ($0<P<1$) are required to sustain the incentives designed.

The system, by dealing with the terms of the distribution verifiable by all the parties involved in the relationship, makes it more costly for each element of the relationship to 'cheat' on the rules or give false signals about performance. This is because average rather than individual signals are taken into account by the system and the distribution of the overall expenditure already commited is based ultimately on these factors.

Let us present an example. Suppose the hospital is to join a programme for expenditure rationalization over two periods. P is fixed (e.g. at levels of 20%). This defines the benchmark X^* with respect to which the performance of each hospital is going to be compared within the peer group we are considering. The crucial point in this approach is that for an individual hospital, X^* can be taken as given, since, although for any fixed value P it will in fact depend on the individual's work-effort, collectively its own contribution in this is neglegible.

Mathematical derivation of the properties of the model.

Values over $\underline{X^*}$ will make for a M_2 type of reimbursement in period 2 (M_2', otherwise). In both cases we will assume that it gets M_1 in period 1 but it is advised that the new policy comes into operation in period 2.[11]

The hospital agents' expected utility is defined over the following arguments:

$$V (Z_1, M_1, M_2, M_2', X^*) = F (X^*-Z_1) . U (Z_1, M_1, \bar{Z}, M_2) +$$

$$(1-F(X^*-Z_1)) . \quad U (Z_1, M_1, \bar{Z}, M_2')$$

i.e. the hospital expected utility is a product of the probability of being out of the zone implied by P, the utility derived from U given Z_1, Z, M_1 and M_2, plus the utility derived if M_1 and M_2' are the respective reimbursement levels, weighted by their corresponding probabilities.

The hospital will choose Z_1 to maximize $\underline{V}$ subject to $Z_1 \geq \bar{Z}$. For an interior solution, the first order condition for this maximization is:

$$V_1 = f (X^*-Z_1) . U (Z_1,M_1,\bar{Z},M_2) + F (X^*-Z_1). U_1 (Z_1,M_1,\bar{Z},M_2)$$

$$- f (X^*-Z_1). U (Z_1, M_1, \bar{Z},M_2') + (1-F(X^*-Z_1)). U_1 (Z_1,M_1,\bar{Z},M_2') = 0$$

Rearranging terms,

$$V_1 = f (X^*-Z_1). \left[U (Z_1,M_1,\bar{Z},M_2) - U (Z_1,M_1,\bar{Z},M_2')\right] +$$

$$F (X^*-Z_1). U_1 (Z_1,M_1,\bar{Z},M_2) + (1 - F (X^*-Z_1)) . U_1(Z_1,M_1,\bar{Z},M_2') = 0 \qquad (1)$$

Since F (X^*-Z_1) is the cumulative distribution function of the observed effort, f(X^*-Z_1)-its derivative- will be the probability density function.

This result, together with the minimum effort level $\bar{Z}$, defines an effort supply function $Z_1 = g (M_1, M_2, M_2' , X^*)$. Note that since $U_1(.)<0$ and $U_4(.)>0$, there can be an interior solution for Z_1 if and only if $M_2>M_2'$ and if $f(X^*-Z_1)$ is strictly positive. In other words, that there is always the probability of an external intervention for the so called 'relegation'.

Since X^* is defined once $\underline{P}$ is selected, for purposes of the effort-supply function of the hospital, $\underline{P}$ and not $\underline{X}$ appears as primarily relevant.

The effort supply function of the hospital is then defined by:

$Z_1 = h (M_1, M_2, M_2', P)$. (2) where $P = 1-F (X-Z_1)$. And it follows that $\partial X/\partial Z_1 = 1$ and $\partial X/\partial P = -1/f(X-Z_1)$.

From the first order condition of the agent's expected utility maximization, choosing Z_1 subject to $Z_1 \geq \bar{Z}$ and using standard comparative static methods, we can derive the effects of changes in variable Z_1 by total differentiation of the expresion in (1) :

$$(V_{11} + V_{15})\, dZ_1 + V_{12}\, d M_1 + V_{13}\, d M_2 + V_{14}\, d M_2' + V_{15}\, d X^* = 0$$

Since $P = 1-F (X-Z_1)$. $dP = \frac{\partial P}{\partial X} . dX + \frac{\partial P}{\partial Z_1} dZ_1$,and $dX = -\frac{dP}{f(X-Z_1)} + dZ_1$,

therefore,

$$(V_{11} + V_{15})\, dZ_1 + V_{12}\, d M_1 + V_{13}\, d M_2 + V_{14}\, d M_2' - V_{15}\, dP/f(X^*-Z_1) = 0 \qquad (3)$$

Assuming conventionally that $U_{11} < 0$ and, for simplicity, that the utility function is additively separable between periods, one gets the following results for an interior solution of Z_1 by solving $\underline{h}$, using the theorem of implicit functions:

$$\frac{\partial Z_1}{\partial M_1} = h_1\ (.) = -\frac{d\ M_1}{d\ Z_1} = -\frac{U_{12}}{U_{11}} \gtreqless 0, \qquad \text{as } U_{12} \gtreqless 0 \qquad [1a]$$

$$\frac{\partial Z_1}{\partial M_2} = h_2\ (.) = -f\ (X^*-Z_1)\frac{U_4}{U_{11}} \geq 0, \qquad [2a]$$

$$\frac{\partial Z_1}{\partial M_2'} = h_3\ (.) = f\ (X^*-Z_1)\frac{U_4}{U_{11}} \leq 0, \qquad [3a]$$

$$\frac{\partial Z_1}{\partial P} = h_4\ (.) \doteq f'\ (X^*-Z_1)\Big(U(Z_1,M_1,\bar{Z},M_2) - U(Z_1,M_1,\bar{Z},M_2')\Big) \Big/ \Big(f(X^*-Z_1).U_{11}\Big) \gtreqless 0,$$

$$\text{as } f'\ (X^*-Z_1) \lesseqgtr 0. \qquad [4a]$$

In [2a] and [3a], object of our attention here, we need to evaluate U_4 at M_2 and M_2' respectively. As we postulated in our prospective reimbursement policy, based on an internal type of incentive as reflected in [3a], its negative sign proves the positive influence on the work-effort $\underline{Z}$ of a marginal decrease in the advanced level or reimbursement M_2' - within the boundaries stated by the policy - whereas an increase would lower it.

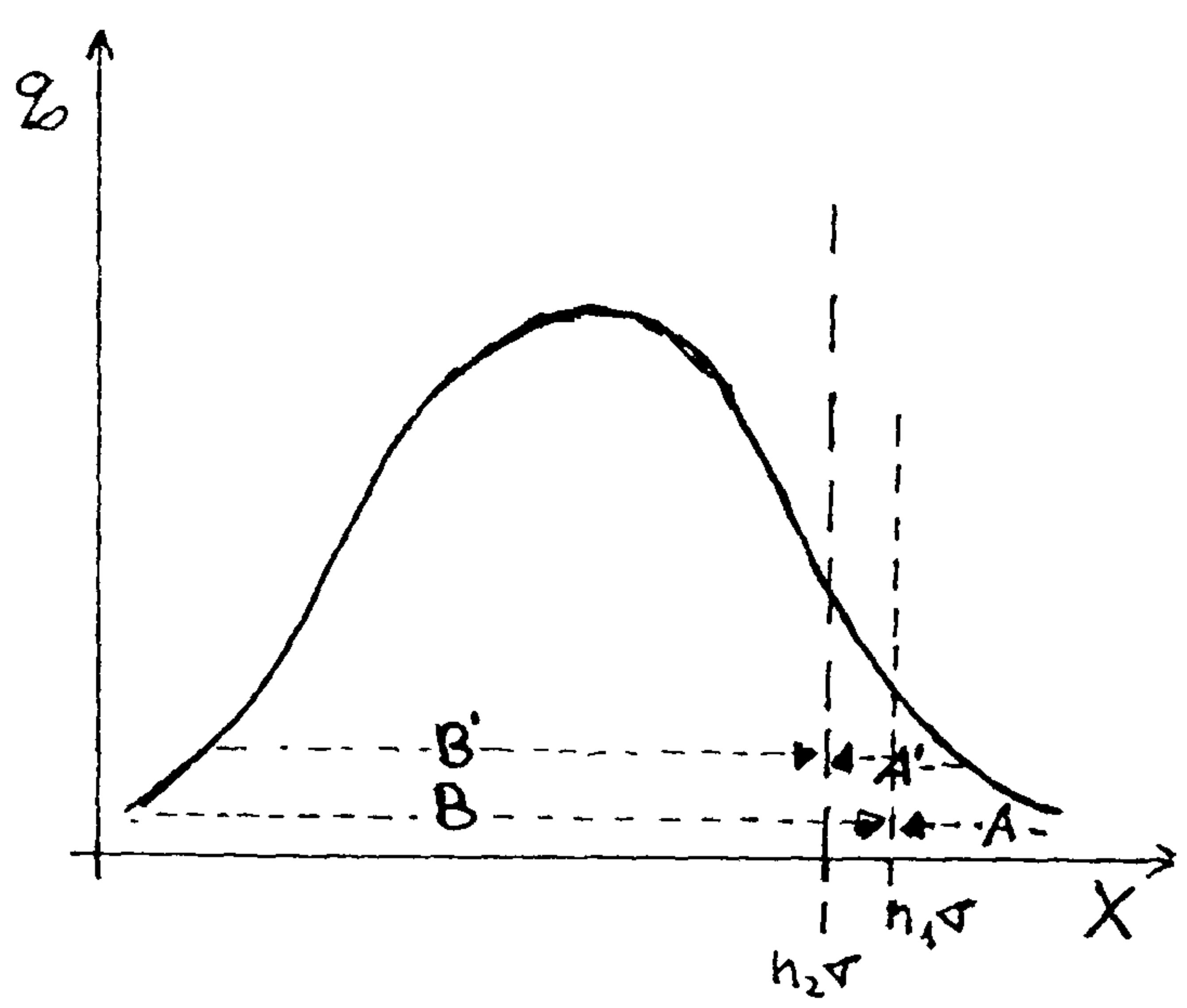
Therefore, the difference in the levels of reimbursement stated from a budget-based contract, advanced through some sort of prospective formulation of the reimbursement policy, well related to the Agents' performance levels and ultimately to their work-effort, can be considered a major element in improving efficiency.

Diagramatical representation of the properties of the formula

Diagramatically, the Prospective Revenue Allocation policy put forward earlier consists, in short of:

(i) The definition of the variable $\underline{X}$ as a proxy for hospital performance and its value of reference, say $\underline{X}^*$, to be utilized as the benchmark for its evaluation.

(ii) The determination of the number of units (percentage), to be submitted each year to the incentive reimbursement policy, with regard to the sample distribution of X. Assuming normality, this implies the possibility of fixing $\underline{P}$ from the product of the value of the standard deviation of $\underline{X}^*$ ($\underline{\bar{V}}$) and the number required for catching the stated percentage of the probability distribution function. (See Figure 1).

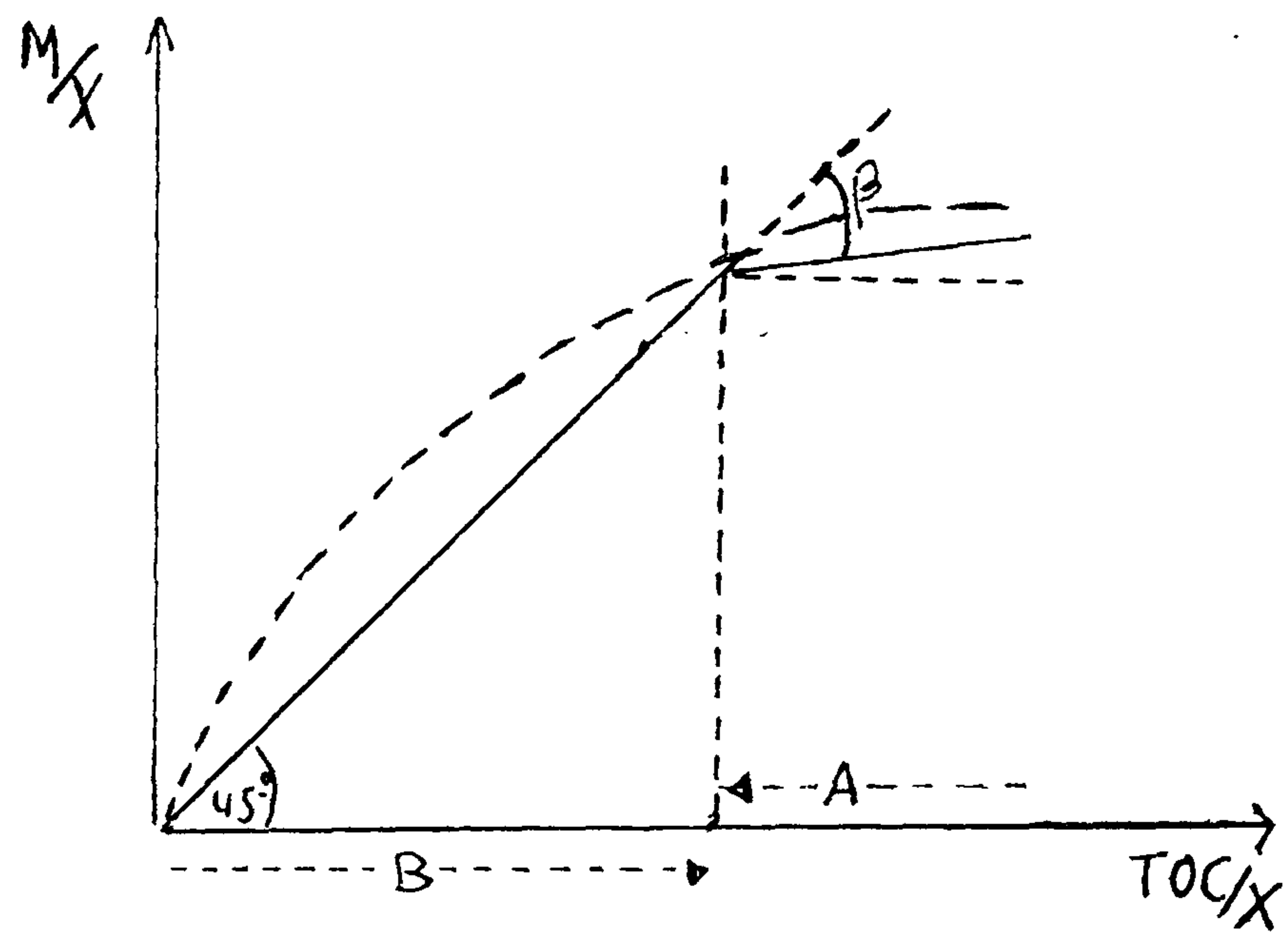


if A=20%, n_1 =1.78 X_1^*=1.78 $\bar{V}$

if A=30%, n_2 =0.68 X_2^*=0.68 $\bar{V}$

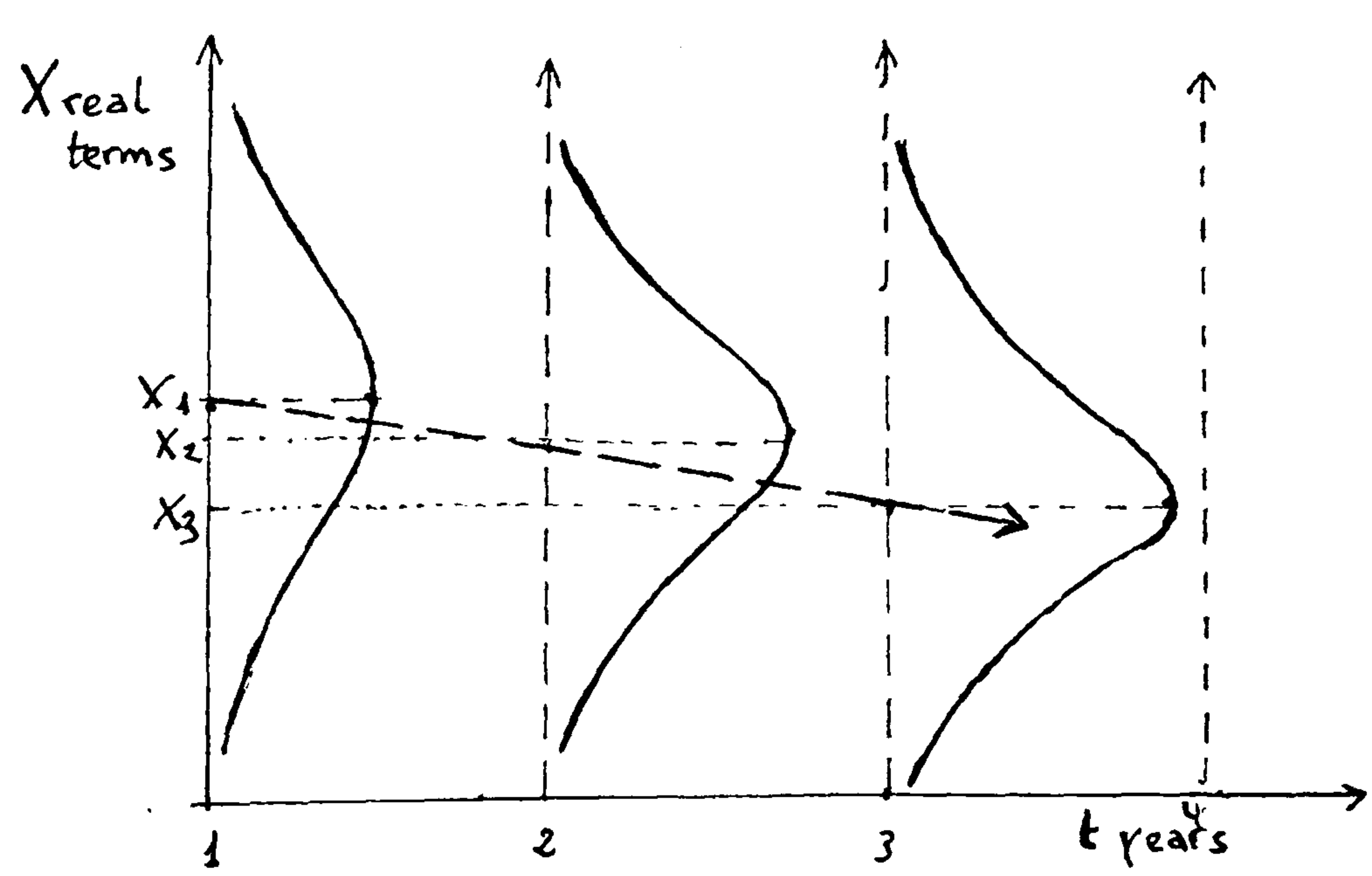
?

(iii) For those units included in the shaded area $\underline{A}$, the prospective control policy is applied in order to achieve an improvement in their performance. The operational mechanics of the system involve in this context some sort of functional relationship between the total reimbursement and the cost per unit, with the strength on efficiency promotion depending on the incentive parameter β. (See Figure 2).



(The continuous line involves the binary set of results which determines the levels of reimbursement. The discontinuous line implies an increasing revenue allocation strategy which follows a less than proportional rise in the level of reimbursement per unit with respect to costs as we move inside of the area A starting in the inflexion point).

(iv) The result of the efficiency improvement involves, therefore, a sequential approximation in real terms to the X* average values of the overal sample. (See Figure 3).



3.4.3. Some other general considerations.

Some other factors have to be pointed out, however, at this stage:
a) So far, financial incentives potentially raised from a reimbursement policy point of view have been only considered. They are by no means, the only set of incentives which can be proposed. For example, regulatory enforcement or education for behavioural change may be used with an appropriate reimbursement policy and complementary training to physicians given since, as the ultimate managers of health resources, they will have to confront those issues related to costs and bear thereafter the responsibilities derived from their decisions.
b) For our purposes, the possibility of 'relegation' rather than the possibility of 'promotion' has been selected as the incentive scheme. Some evidence is available from those countries in which funding for overall responsibility has been implemented[12] that 'negative' incentives in the financial field have produced more direct results. In fact, there seems to be evidence that hospital directors, and doctors in general, have not showed too much interest in saving money if the reimbursement system appears to be working under incrementalistic mechanisms for expansion of recurring expenditure. The reason is that one-off savings today may represent a threat against future expenditure and they are likely to result in permanent reductions of the budget bases.

In this sense negative incentives are not merely the reciprocal of positive incentives. In fact, hospitals do not appear to become more thrifty in order to earn and keep profits, but they may become less inefficient in order to avoid losses. Negative incentives seem to have the virtue of making over-spending risky, given the prospective settlement of the financial constraint with effects on the future revenue planning. Moreover, this view allows money rewards to be considered as 'awards' for excellence since they are not generated just from the hospitals' savings. Thus the question is not that of favouring underserving or pure cost-shifting, but to promote a cost-conscious management of resources.

However, we have to notice that for the process as a whole to operate, requires the validity of the assumption that observed measures of institutional performance can be taken by the parties as acceptable bases for the analysis of the 'budget fat' and managerial slackness. (This point needs further attention and will be explored later on).
c) If governmental incentives have to help achieve some plublic health targets, they must move doctors towards positions of cooperation. This may require as much political as technical elements of motivation. For

these purposes, a minimum level of utility should be guaranteed by the reimburser in terms, for instance, of a realistic wage settlement on which to base the reinforcement of the physicians' own responsibility in the hospital results. At the same time, some minimum levels or standard units of work-effort have to be monitored and strengthened as a condition of payment (e.g. the fulfilment of the minimum notional working time, the actual provision of the service within some statutory boundaries, etc.).

d) Finally, some sort of ranking, order or scaling of measures for performance should be capable of being utilized for basing the budget contract, although no cardinal unit is required for defining an absolute difference between the groups relegated and promoted.

3.5. <u>THE POTENTIAL EFFECT OF AN 'OUTPUT-SHARING' BUDGET SYSTEM ON THE AGENCIES' PERFORMANCE.</u>

The potential effects which can be induced over time from the implementation of an incentive policy of the nature here proposed will be seen in the following exercise.

We will insert for this purpose, our prospective reimbursement policy into a very simple model of expenditure determination, widely used for deriving behavioural inferences in the private versus public provision debate, such as that of Baumol (1967) concerning the existence of a 'productivity lag' in Public sector services with respect to the rest of the economy. This is a model with only one labour input explicitly included [13] - hence with the overall operating expenditure figure equal to the total wage bill - and with a postulated single constant input-output coefficient.

Hence, $\vartheta_{it} = Z \cdot L_{it}$, where $\underline{\vartheta}$ and $\underline{L}$ denote as before output and physical input - or amount of manpower - [14] $\underline{i}$ the Agent unit, and $\underline{Z}$ the true level of effort (or the proxy for productivity).

Crucial elements of the modified version of this model are:

<u>First,</u> the introduction of the hypothesis of variable productivity

(Z_{it}), according to the evidence available, is a more realistic approach than that derived from assuming a single homogeneous rate of growth and an uniform level of productivity for all the productive units in a sector of the economy.

Second, the fact of making explicit in advance the terms of the reimbursement policy. That is, for those units included in the **A** area seen before, we can define a budget-based contract where the budget allocation (ER) is

$$ER = \beta^t \; TOC_{it} + (1-\beta^t) \; \hat{TOC}_{it}$$

with $0<\beta<1$. $\hat{TOC}$ represents the predicted operating costs from the average performance of the peer group considered (once adjusted for exogenous variables through a multi-regression analysis), and t refers to time.

Third. the definition of Z in expenditure terms, and so X^*, by fixing the percentage P (i.e. 20%) - picking up the appropriate number of standard deviations from the predicted mean cost - in order to set the desired level of efficiency.

Fourth, the separation of the total wage bill into its several components in a way in which we can identify each of the specific elements to be considered for specific control.

$$\text{Hence, } TOC_{it} = WB_{it} = W_{it} \cdot L_{it} \equiv \bar{W}_t \cdot \frac{L_t}{N_t} \cdot N_{it} + (W_{it} - \bar{W}_t) \frac{L_t}{N_t} \cdot N_{it}$$

$$+ \bar{W}_t \left(L_{it} - \frac{L_t}{N_t} \cdot N_{it}\right) + (W_{it} - \bar{W}_t) \cdot \left(L_{it} - \frac{L_t}{N_t} \cdot N_{it}\right); \quad [1']$$

where L is the amount of labour input as before, N the population served(15), W the effective cost per unit of labour input and $\bar{W}$ its average sample value.

The first element of the right-hand side of the identity approaches the budget allocation which could be provided on pure population bases (e.g. regarding the average work-load indicators); the second element corresponds to the intra-sample relative price adjustment due, for

instance, to the different external labour market conditions in the rural areas, and the third and fourth elements reflect the 'residual' comprising 'price' and 'quantity' deviations from the average sample values.
Both elements can be represented by

$$\left(L_{it} - \frac{L_t}{N_t} \cdot N_{it}\right) \cdot W_{it} \qquad [1''];$$

and fifth, having postulated some 'target' value for θ_{it}/N_{it}, which would tend over time, with the support of the implemented policy, towards $\theta_t/N_t = k$.

Under these conditions, the effort for a cost-conscious management of resources, and for any potential efficiency improvement, has been postulated in our context to be a function of the binding budgetary coverage, sustained on the rationality of the Prospective Reimbursement Policy formulated. This 'tightness' can be measured as the difference between what the Agency would expect to get under unmodified conditions of reimbursement and what they will actually get according to the control policy implemented. That is,

$$TOC_{it} - (\beta^t\, TOC_{it} + (1-\beta^t)\, \hat{TOC}_{it}) = (1-\beta^t) . (TOC_{it} - \hat{TOC}_{it}) \qquad [2'],$$

where $\hat{TOC}$ accounts for the adjustment of the first and second element of our initial identity [1'], but does not include the 'residual' term [1''].

In terms of the effective work-effort and allowing for a one-period lag in the operation of the financial constraint, according to the model formulated for internal incentives, we can write:

$$\Delta Z = \frac{Z_{it} - Z_{it-1}}{\bar{Z}_t - \bar{Z}_{t-1}} = f\,(1-\beta^{t-1}) \,.\, g\,(Z_{it-1} - \bar{Z}_{t-1}) \qquad [3'],$$

given the assumption that $(TOC_{it-1} - \hat{TOC}_{it-1})$ is in itself a function $\underline{g}$ of the difference between the individual and average work-effort, with $g'<0$ and $f'>0$.
(Note that $\Delta \underline{Z}$ must be strictly greater than one if an increase in the work-effort of the relegated subsample has to be achieved in real terms).

Since in our X-efficiency analysis movements in $\underline{Z}$ are bound to be reflected in $\underline{X^*}$, we can write:

$$TOC_{it-1} - \hat{TOC}_{it-1} = (L_{it-1} - \frac{L_{t-1}}{N_{t-1}} \cdot N_{it-1}) \cdot W_{it-1}$$

Therefore, [2'] can be re-writen as:

$$(1-\beta^{t-1}) \cdot (L_{it-1} - \frac{L_{t-1}}{N_{t-1}} \cdot N_{it-1}) \cdot W_{it-1} = (1-\beta^{t-1})\left(\frac{\theta_{it-1}}{Z_{it-1}} - \frac{\theta_{t-1}}{\bar{Z}_{t-1}} \cdot \frac{N_{it-1}}{N_{t-1}}\right) \cdot W_{it-}$$

or

$$(1-\beta^{t-1})\left(\frac{\theta_{it-1}}{N_{it-1}} \cdot \frac{N_{it-1}}{Z_{it-1}} - \frac{\theta_{t-1}}{N_{t-1}} \cdot \frac{N_{it-1}}{\bar{Z}_{t-1}}\right) \cdot W_{it-1} \qquad [2'']$$

As we postulated, this equality should tend over time towards

$$\left(\frac{1}{Z_{it-1}} - \frac{1}{\bar{Z}_{t-1}}\right) \cdot W_{it-1} \cdot N_{it-1} \cdot (1 - \beta^{t-1}) \cdot K$$

Approximating the solution with $\underline{k}$ evaluated at t-1 on the optimal path as θ_{it-1}/N_{it-1}, the former expression can be writen as:

$$\left(\frac{1}{Z_{it-1}} - \frac{1}{\bar{Z}_{t-1}}\right) \cdot \theta_{it-1} \cdot W_{it-1} (1-\beta^{t-1}). \qquad [4']$$

or, $\left(\frac{1}{Z_{it-1}} - \frac{1}{\bar{Z}_{t-1}}\right) \cdot Z_{it-1} \cdot L_{it-1} \cdot W_{it-1} (1-\beta^{t-1}) =$

$$= \left(\frac{1}{Z_{it-1}} - \frac{1}{\bar{Z}_{t-1}}\right) \cdot Z_{it-1} \cdot (1-\beta^{t-1}) \cdot WB_{it-1}$$

and therefore,

$$\Delta Z_{it} = f\left\{\left(1 - \frac{Z_{it-1}}{\bar{Z}_{it-1}}\right) \cdot (1-\beta^{t-1}) \cdot WB_{it-1}\right\} \qquad [5'].$$

This provides a similar result to that derived from the theoretical approach seen before(16), with the increase in the work-effort made dependent on:

(i) The lagged component of the Agency work-effort with respect to its average value, weighted by the initial level of budget reimbursement (or output amount in [4']).

(ii) The incentive parameter β stated in the funding policy previously advanced.

Therefore, in basing the reimbursement levels on the Agencies' performance, the productivity rate of growth becomes endogenous to the reward structure and hence able to be submitted to control and manipulation. Correspondingly, the link between productivity and financial rewards, when compared with what the neoclassical theory asserts for the incentive motivation of the economic agents, appears reversed.

3.5.1. A Simulation Approach to the Solution.

The 'moving average' nature of the result that appears from the reimbursement system designed - given the average or predicted sample benchmarks to which hospital performance is going to be sequentially compared - raises a non-linear problem in the difference equation derived.[(17)]

In approaching its solution we can either:

(i) replace the target observed values $\underline{X}^*$ by some other standard values, discretionally defined 'a priori' in each period of time, independently of the sample behaviour;

(ii) simulate the behaviour of the derived equation through the construction of a computer program, in order to suggest the pattern followed by the $\underline{Z}$ values under some previously stated plausible conditions (mainly with regard to the analysis of the deviations between actual and predicted or potential levels of expenditure).

We adopt the second approach, since the first alternative is inconsistent with the nature of the objective conditions required for our optimal reward scheme and which must be in the core of the budget contract advanced.

For these purposes, we need:

1. To define those variables included in the analysis in a way that they can be empirically estimated;

2. to give some plausible values to the parameters which appear in the derived equation. These values can be in fact, a matter of policy choice; and

3. to make some explicit assumptions about the functional form and other related parameters of the incentive scheme formulated.

We proceed as follows:

a) The proxy for the observed work-effort, and 'efficiency' measurement, is defined in expenditure terms. The variable selected is the 'average cost per case', redefined according to the results of increasing the sample homogeneity, first by using the clusters derived from the inclusion of the external factors thought to influence legitimately hospital costs, and second, from adjusting the average values resulting from each of those subsamples for the individual internal factors, which prove an impact on costs, as analysed through the standard multi-regression techniques. (Both approaches will be adopted in the following Chapter).

The starting value for the proxy of $\underline{X}$ will be defined for each unit of the sample as the ratio between the estimated and the observed average cost. For the remaining years, individual productivity levels will be taken from the addition to those starting figures of the values calculated from [5'] , in relation to the subset to which the formula was eventually applied.

The 'exogenous' average productivity increase (18) is calculated year by year from the ratio between the estimated and observed values, but refered only to the subsample of observations to which the formula does not apply. This figure will be used thereafter as the denominator of [3'] to calculate a relative value of $\underline{Z}$ for defining a $\Delta Z_{it} > 1$.

b) A four year incentive scheme is stated, during which hospitals failing above the levels of $\underline{X}^*$, sequentially defined by the iteration, will be, in our terms, 'relegated'. This budget contract provides a controlled reimbursement which depends on a weighted consideration of;

(i) the budget level resulting from holding unmodified conditions of funding;

(ii) the budget level resulting from the individual prediction of the hospital average expenditure, according to the performance of the sample in which the unit is included, as considered in [2'].

The efficiency parameter will receive an initial value of 0.95, which over the four years is bound to decrease as t increases (i.e. 0.940, 0.884, 0.781), given the nature of the proposed formula:

$$TOC_{t+1} = (\beta^{t} TOC_{t} + (1-\beta^{t})\, \hat{TOC}_{t}).(1+r_{t+1}) \text{ .}$$

Due to the complexity of the estimation exercise and problems of data availability, we will not be able to repeat the exercise for each single year - although this should be possible with a more complete information base and an appropriate system of data processing, as we point out in Chapter V.

Hence, we will simulate repeatedly the initial estimattion - adjusted annually for inflation - using alternatively the retail price index, the projection of the trend of the health care expenditure and the increase in the Gross Domestic Product at market prices (or factor r).

The 'global budget' for funding purposes will be thereafter calculated from multiplying the average cost per admission within each of the clusters derived - adjusted for the estimated parameters of the regression equation formulated - by the observed actual number of admissions in each hospital unit.

With respect to the proportion P of units considered each year for budget utilization review and control through the revenue planning system, the value is fixed alternatively at levels of 20% and 30% with regard to the cumulative distribution of the differences between the estimated and actual values of the total operating costs in the overall sample. The number of standard deviations chosen for this purpose are 1.78 and 0.68 respectively. (For simplicity, we assume normality in the distribution of the selected variable). The values of the standard deviations are calculated using the conventional procedures in relation to the variable defined as the difference between the actual and estimated values.

The maximum cost reduction to be reasonably expected for each single year (hence, residually, the minimum utility 'safe' from the incentive scheme) is established with values in between 1 per cent and 3 per cent.

We take these floors from the cost savings criteria expounded on the Resource Allocation Working Party for England, and from the U S experience in the operation of some prospective reimbursement experiments.(19)

c) The functional form selected for f and its weighted factor WF - which accounts for how quickly the incentive scheme should develop its effects in moving towards the cost reduction target - is dealt with in the following way:

With respect to the specific shape of the incentive formula aimed to promoting a productivity change of the nature shown in Figure 4, the first hypothesis implies that for any single year, the maximum productivity increase (DZMax), or 'cost-efficiency' reduction, increases positively with F. This variable F is assumed to approach the 'fat' of the budget, or the 'sacrifice possible', this being defined, as before, as the difference between 'what they know they are going to get in relation to what they would expect to get in the case of holding constant conditions of reimbursement' (i.e. resulting from an incrementalistic pattern of reimbursement). (See Figure 5).

Figure 4

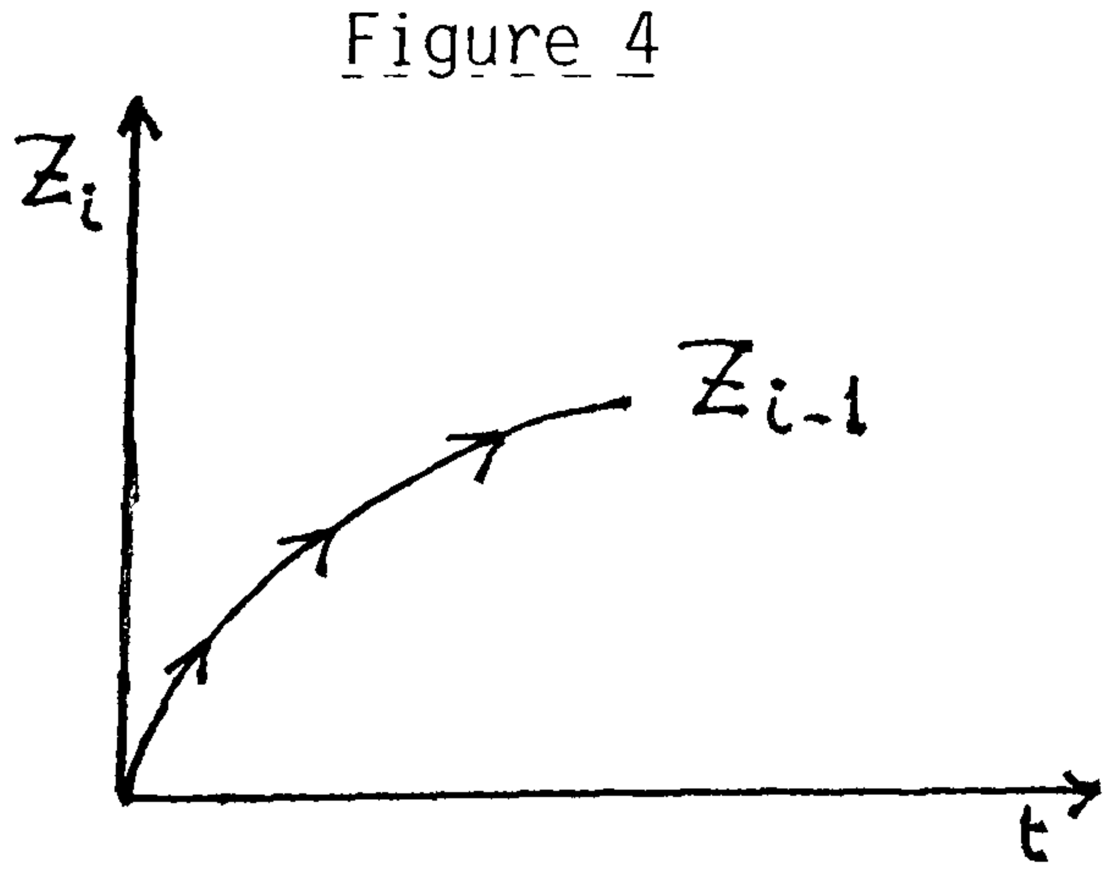
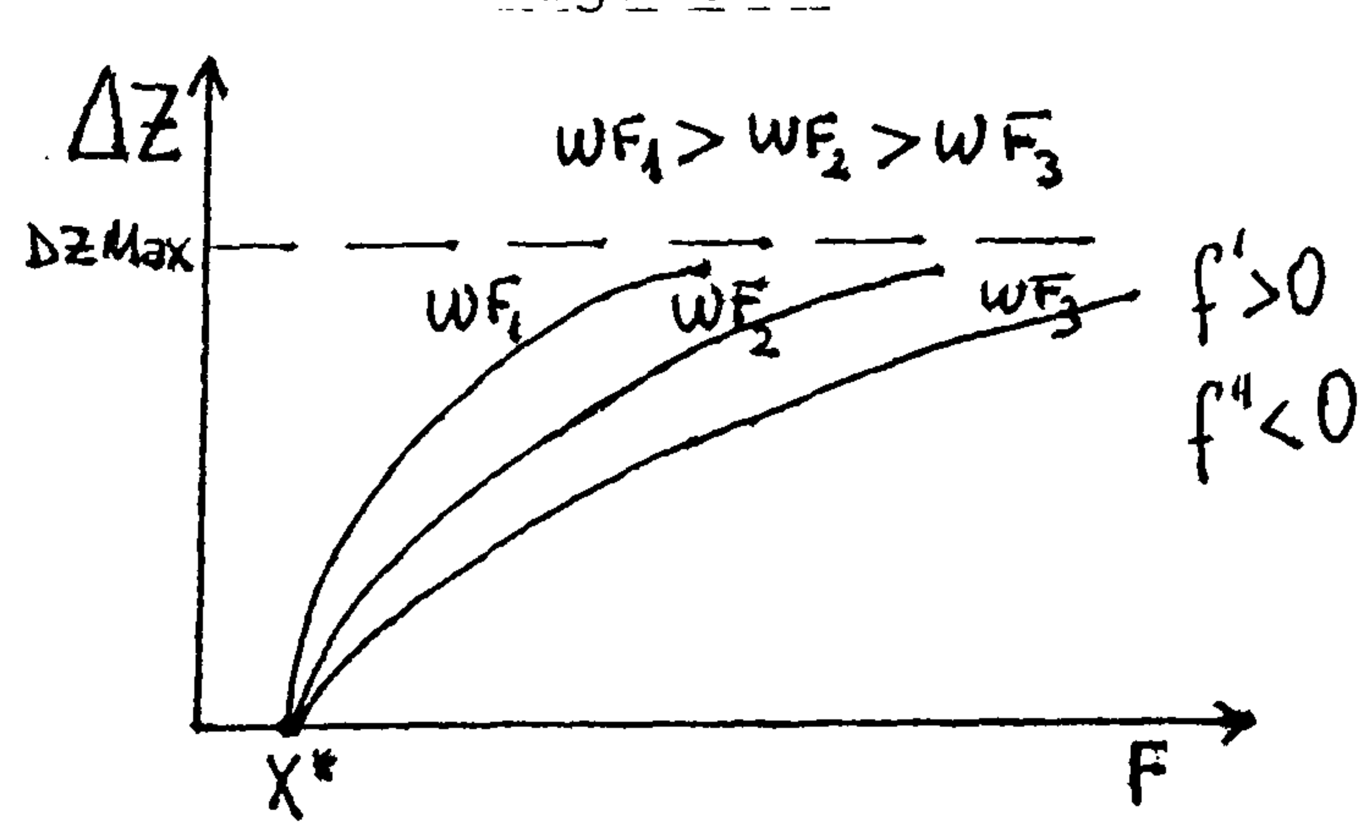


Figure 5



(DZMax defines in fact the 'possible set'. The factor WF receives alternative values for the prupose of accounting for how quickly the cost reduction targets are reached).

The assumption involved in f is that the incentive which stems from the announced penalty declines to the DZMax value once it overtakes a determined amount close to the cost-reduction target.

As a result, the specified form can be defined as:

$$\Delta Z_{it} = \text{DZMax.}\ (1-e^{-WF.F})$$

where F, as before, is equal to: $\left(1- \dfrac{X^*_{it-1}}{X^*_{t-1}}\right)(1-\beta^{t-1}).\ TOC_{it-1}$,

with $Z_{it} = 0$ when $F \leq X^*$, and $Z_{it} \longrightarrow$ DZMax when $F \longrightarrow \infty$.

Other forms considered give values to f' and f'' both > 0, with ΔZ_{it} = DZMax. (F^{WF}), and, from taking the functional form $y = b^X$, under the exact specification of ΔZ_{it} = DZMax. $(WF^F - 1)$, with WF > 1.[20]

The results of the exercise are presented and analysed in Chapter V.

3.5.2. Potential Hypothesis and Policy Testing.

At this stage, it would be attractive to set a formal test of some alternative hypotheses presented in this Chapter, mainly with regard to the positive effects of the reimbursement policy on a cost-conscious management of resources, from replacing a cost-based retrospective system by a prospective strategy of the characteristics described above.

This is not possible at this point, since there are not, obviously, two consistent situations to be compared. However, two simple tests can be designed to validate the operational insights of the system, once the projected policy is in operation:

a) With regard to the nature of the policy, its fundamental feature could be approached by estimating the following equation:

$$\Delta Z_t = \lambda_0 + \lambda_1 . \beta(t) + \lambda_2 . \Delta Z_{t-1} + \lambda_3 \Delta Z_{t-1} . \beta(t) + \varepsilon_t$$

where λ_3 tries to collect the joint influence of the efficiency parameter interacting with the lagged increase of productivities. The reason for the specification of this equation is that the value of the efficiency parameter is dependent on both the previous year's level of effort and on the present Agency performance with regard to their success in closing the gap between actual and predicted expenditure.

b) Looking at the operational way in which the policy as a whole works, the empirical test would consist of the study of the influence in the cost behaviour of the sample analysed once the prospective reimbursement policy is implemented. This second test can use a dummy variable for collecting the effect of the policy on the endogenous cost variable in a behavioural cost function of the Hospital units enrolled in the policy with respect to the cost function of those units not included.

Either in a cross-section or time-series analysis, conventional levels of confidence may be used in support of the statistical significance of the reimbursement policy, in approaching its differential impact once we control for 'everything else' in the regression equation.

However, in the absence of an incentive formula of the type here proposed, we will be able to test in our context only some alternative hypotheses about the effects that this absence may have had on hospital performance. They will be centered on the hypothesis of cumulative causation and of the absence of self-correcting mechanisms in the management of public hospitals. At 'sensu contario' this test will be suggested as a proof of the value of the policy here proposed for improving the allocation of resources in the hospital sector. In addition, by examining its results on a cross-section basis, some suggestions about the nature of spatial and institutional X-inefficiency will be further explored for policy purposes.

3.6. CONCLUSIONS AND SUMMARY.

In this Chapter we have examined the terms in which an alternative budget policy, aimed at promoting efficiency in the hospital sector can be designed.

We conclude that having recognized the appropriate institutional setting for the relationship between the decentralised institutions providers of services and the reimbursing Authority with control over the total amount of resources available, as described by an Agency type of relationship, a budget-based contract of the characteristics proposed in this chapter can raise incentive compatibility. In addition, a budget-based output-sharing system, which accounts for the differential performance of the institutions and allocate revenues correspondingly, can help to promote efficiency and to insert guides for the allocation of resources within the Spanish hospital sector.

The specific design of a Prospective Revenue Allocation Policy of the type proposed in this Chapter will be the object of attention of the next Chapter.

FOOTNOTES.

(1) in "Production, Information Costs and Economic Organization". American Economic Review, vol 61. 1972 (page 780).

(2) On one hand, the relevance of an explict separation between the allocation of resources and the financial funding limits for the analysis of the economic behaviour of the organisations has been recently expounded by Hoenack (1983). On the other hand, the potentialities raised from the insertion of economic incentives in systems of budget control have been studied in the context of incomplete markets by Demski and Feltham in The Accounting Review vol LIII nº 2, April 1978.

(3) The Von Neuman-Morgensten "two persons non-zero sum game" solution consists of accounting for all those inputations in which each player gets individually that amount which he can secure for himself, while the two together get precisely the maximum amount which they could secure together. A Nash type of solution would stem from the bargain in which it is maximized the product of the utility increment from the initial threat point. In both cases, less than complete information for the calculation of the required inputations and the lack of an efficient communication network seem to make these solutions less meaningful in our context.

(4) as referred in CODDINGTON A. "Theories of the bargaining process". Allen and Unwin, 1968. London.

(5) This concept of 'selective rationality', as defined in Leibenstein (1975), refers mainly to the differences in constraint concern, and it is reflected in a sense of 'obligation to commitments', 'degree of calculatedness', 'degree of satisfaction' and some other aspects ultimately related to the concept of managerial discretion.

(6) When uncertainty is present, the optimal contract can be approached through a rental agreement, contracting for a fixed fee and enabling the Agent to claim for any residual left. The optimal output-sharing has then a mixed nature, in which the consideration of the agents' risk behaviour is decisive to establish the correct balance between the rental fee and the risk-share in the outcome.

(7) Since the total amount to be distributed has to be previously committed, precaution has to be taken in order, (i) not to overrun the total amount available: a general maxicap might be required for this purpose on the non-relegated subset, which optimally should take into account the expected changes towards the cost-saving targets of those units included in the relegation set; and (ii) to commit any amount left unused over the total figure available, for example to some specific types of service developments or some particular quality improvements.

(8) Other formulas for incentive reimbursement could be however also studied. For instance, referring DI as the deficit incurred, in general terms:

$ER_{it} = f\left(1/n \sum_{\ell=1}^{m} DI_{it-\ell}\right)$, or $ER_{it} = f\left(\alpha \sum_{\ell=1}^{m} (1-\alpha)^{t} . DI_{it-\ell}\right)$, with $0 < \alpha < 1$.

(9) in "Trade Unions and Economic Efficiency". The Economic Journal, Papers and Proceedings, 1983.

(10) This form makes the implicit assumption that the distribution depends only on the error and not on the level of true effort. However, it would not make a difference if output is made to depend on $\underline{X}$ (observed effort) rather than on the true effort $\underline{Z}$, except for the minor complication of having to treat output, and hence remuneration, as a random variable (see Malcomson, op. cit.).

(11) Probably, some boundaries should be set up in a way that M_2' is just sufficient to avoid complete 'bankrupcy' (e.g. according to a budget utilization review).

(12) This is the case of Canada. See for this purpose, W.A. GLASER "Paying the Hospital: Foreign Lessons for the United States". Health Care Financing Review, Summer 1983, vol 4 No. 4, and particularly, the monography "Paying the Hospital in Canada", November 1980.

(13) However, some explicit assumptions about the role of the capital input in the production function of services can be made. For instance, in the sense that the capital factor is at any rate implicit in the sector, but with different ex-post relevance, given its character of limitational input (this being necessary but not sufficient for the growth of the output), or given the nature of the innovation process and its reflection on the technical substitution possibilities.

(14) Whether the number of "whole-time equivalent" physician hours is a more appropriate measure of labour input is, in our case, not a matter of choice but of data availability, and of the specific purpose of the analysis (e.g. the intensity factor relevant for output production can try to be captured entirely by the work-effort of L in terms, for instance, of the medical services offered).

(15) Other measures of output could also be selected instead of the pure population data, such as the number of beds per population available, for instance, in the case of being applied the formula to the inpatient care sector.

(16) [4'] and [5'] offer a first approach to the potential policies for cost-containment purposes. For example, in adjusting expenditure in real terms, by creating a gap between the adjustment rate allowed and the actual rate ex-post required from holding figures in real terms; by controlling manpower expansion, in fixing the level of inputs per capita, or any other output or throughput measure selected; or via the design of an incentive scheme and manipulatiing the efficiency parameter over the time horizon of the revenue planning policy.

(17) In fact, being the real increase in work-effort a function of lagged individual performance with respect to the average, and operating the mechanism designed in a selective way in which each period, different subsets of units will be formed for each year, depending broadly, on $Z_{it-1} - \bar{Z}_{t-1} < 0$ for the subset $I \in i$ and, otherwise, $Z_{it-1} - \bar{Z}_{t-1} \geq 0$ for the subset $II \in i$.

(18) The term 'exogenous' refers here to the fact that its increase is considered to be independent of the operation of the incentive formula.

(19) These figures have been taken from the England Resources Allocation Working Party (1976)(pages 32 and 33) and from the United States experience in Hospital Regulation (see Sloan et al. (1982), page 105).

(20) All of them respond however to the assumption, in line with Leibenstein's proposition, that the degree of X-inefficiency (and the potentialities for efficiency improvements) is a function of the size of the "shelter" or 'budget fat', given its direct influence on the agent's work-effort.
(H. LEIBENSTEIN, "Notes on X-efficiency and bureaucracy", published by R.C.O. MATTHEWS and G. B. STAFFORD (joint eds.) in "The Grants Economy and Collective Consumption", International Economic Association, London, 1982.

CHAPTER IV ISSUES IN THE DESIGN OF A PROSPECTIVE REVENUE ALLOCATION POLICY

ISSUES IN THE DESIGN OF A PROSPECTIVE REVENUE ALLOCATION POLICY

4.1. INTRODUCTION

We examined in the previous chapter the characteristics of a budget-based contract which is optimal for our purposes, in the sense that it can create the efficiency motivation required for improving performance in the hospital sector over time.

In this Chapter we will study the translation of the abstract setting of the former budget-contract into the design of a reimbursement formula, basing a Prospective Revenue Allocation Policy - 'PRAP' - in the field of the Spanish Hospital Sector. The construction of this formula will enable us to test in Chapter V some hypotheses on system inefficiency and variahle productivity that were advanced earlier on.

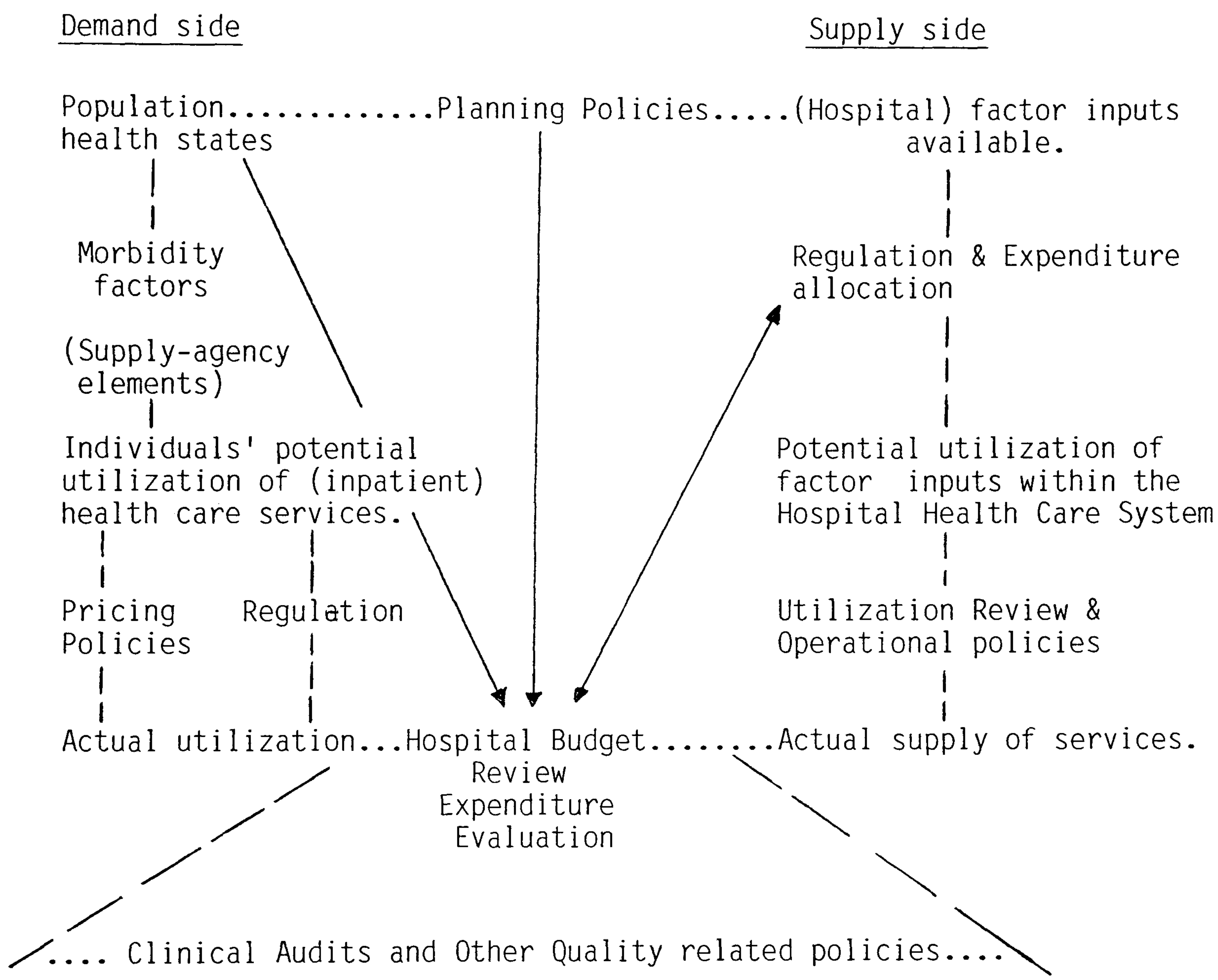
In this introductory section, we will frame the budget approach to a PRAP in the context of the determinants of health care expenditure, and the corresponding cost-containment policies in order to provide a better understanding of its workings and limitations.

4.1.1. Determinants of Health Care expenditure, Cost-Containment Strategies and Budget-Based Policies.

The general context in which a Budget-based policy applies can be seen in Table 1.

TABLE IV.1

The operational context of a Hospital Budget Review and Expenditure Evaluation Policy



Regarding regulation and expenditure allocation procedures in the context of the supply side, we must distinguish those policies that exert control over some partial aspects of the hospital operation from those which provide a global restraint on activity, but which allow hospitals to decide the details of production.

The first approach mainly refers to the control of inputs or levels of performance on the bases of 'engineering', economic, financial or epidemological evaluation. This can be done either;

(i) from a planning view: e.g. limiting manpower increases, fixing some input costs or controlling for the supply of capital - hospital beds, major equipment for service expansion, etc. - or

(ii) from an operational point of view, by exerting for example, controls over the length of stay of the patients admitted, the medical utilization of some specific inputs which appear to be related to the quality of some services, or by imposing normative criteria in order to restrict the utilization of certain services.

The second approach mainly refers to those budget policies which aim at a cost-conscious management of resources through the imposition of a global resource constraint, leaving to the hospitals the decisions concerning provision. If the Hospital Board enjoys an appropriate degree of autonomy, then it is considered likely that this type of policy will lead naturally to the application of techniques designed to improve the efficiency with which hospital budgets are spent. Otherwise, in the absence of a binding overall restraint on hospital resources, no reason would exist, in our view, for expecting the use of techniques of evaluation analysis in the internal allocation of resources.

How a budget policy of this nature can be designed and how it could operate in the context of the Spanish Hospital Health Care Sector will be described in the next subsection.

4.1.2. Considerations to be taken into account in designing a Budget-Based Prospective Revenue Allocation Policy.

In designing a PRAP of the characteristics formerly seen in the context of the Hospital Health Care sector, the following considerations need to be taken into account:

1. Which specific causal factors are the policy makers attempting to tackle through the Hospital cost-containment policy, and through which mechanisms those factors are postulated to operate.
This requires:

(i) the definition of the ultimate and intermediate targets of the policy (e.g. efficiency promotion through the achievement of the maximum feasible output, a given optimal output mix, or the utilization of an appropriate input combination);

(ii) the recognition of the elements by which the former outcomes could be achieved (e.g. once we have analysed the workings of the policy in the context of a plausible Agents' reaction function consistent with the behavioural model specified);

(iii) the identification of "where" the policy can be best implemented (i.e. once we have considered both direct and indirect effects which the policy is likely to provoke).

2. How those causal factors can be controlled by the introduction of the policy, and which mechanisms are thought to be appropriated for its correction. The design of this instrumental policy mainly involves:

(i) the definition of a particular functional form, according to the allocation of overall or specific value responsibility (looking, for instance, at the effects on the degree of concern over the constraints faced by agents, with regard to the targets defined in terms of work-effort), and

(ii) the specification, in consequence, of the financial unit to be considered for reimbursement purposes (i.e. a fixed or variable basis), accounting for its ultimate effects on the characteristics of the hospital output provided and the degree of utilization of the existing facilities.

The description of the actual type of PRAP proposed here, and its workings in the context of the Hospital Cost-Containment Policies, are presented in the next section.

4.2. DESCRIPTION OF A 'PRAP' AND ITS WORKINGS

Taking into account the considerations set up in the former section, the budget approach to a Prospective Revenue Allocation Policy in our context will be based on:

a) The identification of system-inefficiency (as defined in Chapter I) as the causal factor accounting for an important part of hospital cost increases;
b) the aim of putting forward a target for incentive motivation, able to promote, overtime, efficiency improvements (in the sense defined in the second chapter) out of the incrementalistic guides for current hospital expenses;
c) the recognition of the phenomenon of variable productivity to which the allocation of hospital funds needs to be mostly linked.

This approach views, moreover, the role of the physician as the dominant agent in the hospital decision-making process, and it proposes a policy on the overall allocation of revenue, accompanied by the allocation of specific responsibilities as the basic mechanisms of its control.

4.2.1. Basic Idea and Potential Utilization of a 'PRAP'.

The basic idea of a budget approach to a PRAP is that each hospital receives a revenue allocation which is deemed to be an appropriate operating cost figure. The rationale of such a policy consists of hospitals having to keep their expenditure within the limits imposed by the budgets objectively stated (see Evans 1970).

What distinguishes moreover a Prospective Revenue Allocation Policy from other Cost-Containment Policies of financial restraint, independently of the method of calculation, is that the rates reimbursed are set in advance for a future period, leaving hospitals to plan expenditure accordingly. As a result, hospitals come ultimately to share the risks from spending over and above the prior specified amounts.

In order to determine these prospective values, we can state in principle some normative criteria, either on input utilization or on overall hospital performance, or we can derive empirically some average or standard benchmarks. This second view is mainly pursued through assembling data into sets of "comparable" hospitals. Deviations from the average will distinguish in principle, the "efficient" from "inefficient" hospitals in the allocation of hospital expenditure.

This type of prospective reimbursement policy proposed here can serve two main goals:

a) As a Budget Review System. This system for evaluation of budget requests requires the revision of the prospectively set budget forecasts by the reimbursing Authority in the light of the former base-lines. As a result of the comparison, expenses which appear as "unnecessary" or "excessive" in view of the stated benchmarks are requested to be reduced by a certain time horizon.

Hospitals are generally given the opportunity to negotiate and/or to appeal against such decisions. However, once the final budget has been approved, a payment rate or a fixed global sum is established, covering what is understood to be the hospital "allowable" expenditure.

b) As a Formula-Based Revenue Allocation Policy. A prospective rate or total amount is determined here for purposes of a direct reimbursement through a revenue allocation formula. The formula is periodically recomputed and adjusted for external increases in order to set up each year the prospective expenditure figure.

This method of distribution of funds involves relatively little direct interaction between the Hospital Boards and the rate setting Authority, and it is, therefore, less costly to administer. But this method is more likely to be controverted given that no bargaining is provided for the hospitals. Therefore, it can be expected to be less preferred by hospitals than the former method.

In both cases, the revenue allocation base-line is grounded in factors which usually refer to certain hospital characteristics, actual performance and some other relevant factors in the hospitals activity. Based on this information, an expenditure or standard cost-line is fitted, which is subsequently used to evaluate hospital performance and to allocate the budget funds. Hospitals above the fitted line, and hence considered relatively inefficient, are requested to reduce costs so as to approach standard values.(1)

The residuals between the actual and estimated figures distinguish, in principle, the efficient from the inefficient hospitals. The degree of 'relative efficiency' is defined, therefore, in terms of the discrepancy between actual and predicted costs, as derived from the estimation of the hospital cost function.(2)

4.2.2. Approach Adopted in Designing a 'PRAP' in our Context.

The specific approach proposed here consists of:

a) The determination of a revenue allocation formula and the allocation of the responsibility for the disposal of the resources. This is a power vested entirely in the hospital decision-makers.

This determination will require: 1.-The selection of some homogeneous subsamples for which an average cost figure could be plausibly derived; and

2.-the estimation and prediction of the individual levels of hospital expenditure, according to the 'average' performance of the subsample.

b) The implementation of a budget evaluation strategy, supported by the revenue allocation formula and planned over a time horizon. For this purpose, the criterion for revenue allocation is defined as a weighted sum of actual and predicted costs, with the weighting factors changing over time (see page 103 Chapter III).

Summarizing, the policy aims to reinforce the role of the budget constraint in the hospital decision-making process. This result is expected from the incentive/penalities motivation of sharing between the Hospital Boards and Reimburser the difference between the prospective amount set and actual expenditure. The process of correcting discrepancies between actual and targeted values are phased over several periods through a revenue planning system. Such an approach may indeed be needed, given the evidence that costs at any point in time depend on both current and past behaviour, in such a way that past decisions cannot always be reversed immediately, even if current effort is very great.

In the rest of this Chapter, we will work out the mechanisms of the 'PRAP', by showing its construction in the applied field of the Spanish hospital sector and trying to establish the potential it has for reducing variable productivity in this sector. √

4.3. COST FUNCTIONS: CONCEPTUAL ISSUES

The construction of a base-line for the allocation and budget evaluation of hospital expenditure requires first the consideration of the following steps with regard to the nature of the cost-relation formulated: 1. The definition of the basic relationship to be analysed (and of those variables included therein).

2. Its mathematical specification and the measurement of the relevant variables.

3. The study of the alternative approaches to its (empirical) estimation, methodological issues and some other related econometric problems. √

In this section we will examine the first of the three points mentioned above. Points (2) and (3) will be explored in sections 4 and 6 of this chapter, respectively.

4.3.1. The Definition of the Basic Relationship Analysed.

The definition of a total cost function requires, from a theoretical point of view, at least three sets of variables (see, for instance, Henderson and Quandt, 1980, page 84).

a) The quantity (Y) and quality (Q) of the output produced.
b) The vector of factor input prices (Ip) and their quantities (Iy) utilised to produce this output (at a given quality level).
c) The service conditions (Sc) which affect the level of inputs required, together to the current state of the technology available.

This is:

$$TC = f\ (\ Y,Q,Ip,Sc,St\ )$$

Determining relevant hospital expenditure by establishing relative efficiency across hospitals has one main difficulty. This arises from the definition of the underlying production function supporting the associated (dual) cost function. In fact, as we will see, the elements included in the production function of hospital care will not always allow for a clear-cut identification of the cost equation.

4.3.2. The Definition of the Components of the Basic Relationship.

The first issue to be addressed in defining the components of that basic relationship is the definition of what constitutes hospital output.

It should be recognized that the final output of the hospital activity is the improvement in the health states of the patients treated or perhaps, ultimately, the increase in the levels of satisfaction that this improvement represents in the individuals' utility function.

However, as we will see in the next section, operational reasons usually prevent the utilization of this approach. This is mostly due to measurement problems, particularly complex in the context of health care, and to the identification of the output of the health care industry, in which the production-supply aspects of hospital activity can dominate the existing hospital information systems.

A further problem in defining hospital output is that of accounting for its multidimensionality. Several different aspects of care and cure are usually present in hospital activity. In addition, training and medical research may accompany the former hospital services. They can be certainly analysed as increases in the human capital stock or as inputs for future outputs, given their potential impact on patient health states. But this opens problems of trading-off one dimension against other and even within each of them.

In fact, in order to deal with this hospital multiproduct approach, we need to assume that each basic relation can be identified; that is, that the production function for each of the separate products is identifiable. We must then have a method to select and classify each individual relationship between inputs and each of the resulting outputs. For this, the assumption of the separability of the underlying technology requires that the use of an input in the production of a treatment, say of one type, does not affect the (efficient) method of producing the treatment of other types (i.e. assuming that "treatment" is regarded as the primary output unit from the hospital activity). This is unlikely, however, to be a valid simiplification of reality.

A second important component in the definition of the components of our basic relationship is the vector of the input factor prices and quantities and other external conditions which affect the provision of the service.

The adjustment for factor price usually arises from the need to control for differences in those exogenous factors which affect differential (territorial) purchasing abilities. A 'geographic relative price effect' of this nature can be due, in the present context, to the following factors:

a) The direct influence of differences in the territorial index of prices (and cost of living) on the price of the factor inputs.

b) Local scarcities in the amount of certain inputs available. These can stem as a cause or a consequence of the former differences in territorial prices.

c) The influence of the geographic variations in the rate of unemployment, and in the price of the labour input in particular, if a competitive labour market can be assumed to exist.

A third component is that related to the differential service conditions which may influence territorial costs. In fact, regional variations in social and economic characteristics, external to hospital activity, can make for differences in costs due to circumstances beyond the control of the hospitals (e.g. 'sunk' costs, enviromental externalities, social deprivation, etc.. See in this line those factors described in footnote 8 of Chapter II).

Finally, the state of technology is a relevant variable in the light of its impact on costs. This is deemed to be, however, a minor factor in the present context, since, in principle, the level of technology available to any hospital is likely to be highly similar in any given year.

It should also be noted that questions of output quality will not be separately examined here. This is because the definitional line adopted above assumes that outputs with different quality aspects constitute in reality different outputs. Moreover, this approach avoids the discussion of the conceptual issues involved in the definition of the 'quality' of hospital output, a simple presentation of which can be seen in Cullis and West, 1979 (page 120).

Summarizing, important conceptual problems exist in the definition of the basic cost-relationship analysed and its relevant components, once translated into the hospital care sector. These difficulties and, in particular, the question of the hospital output and the need to identify a vector of input prices and quantities in correspondence to its multidimensional aspects, have orientated the attention of the literature on hospital cost functions towards a reduced set of behavioural factors, which, without being strictly speaking definitionally related to the former variables, can actually explain hospital costs.

The operationalisation of this approach will be presented next.

4.4. OPERATIONALISING THE COST FUNCTION: BEHAVIOURAL ANALYSIS AND HOSPITAL OUTPUT.

In this section we will explore the rationality of the behavioural analysis of hospital cost functions and the measurement problems involved in the specification of the hospital output. In this context, the search for a sample selection of greater hospital output homogenity and for an explicit hospital output-mix adjustment, will be particularly considered.

4.4.1. Behavioural cost functions and hospital output.

Given the difficulties found for an appropriate specification of the fundamental components which should be included in a conventional production function and its associated cost function, the existing literature has turned its attention towards the behavioural cost approach.

The basic relation is then defined from the exogenous factors which prove to be of "ex-post" relevance in explaining the cost behaviour of hospitals. In this approach the efficiency line can be drawn from the analysis of actual sample-average costs, associated with changes in the unit costs of the services, holding the character and complexity of the output constant. The fundamental logic behind this is that if all the other explanatory variables remain constant, differences in the efficiency in which input prices are used will necessarily be inversely related to costs. That is, other things equal, the cost per unit will fall as the efficiency with which resources are utilized increases, and vice-versa. However, the connecting link that supports this relationship remains the production function. It should be noted too, that this approach makes most sense in a static world in which the major problem is that similar hospitals produce the same services with different levels of efficiency.

Under the former behavioural approach, the issue of measuring hospital output has received a preferential attention.

Problems of measuring improvements in health states, due to the lack of data (i.e. no indicators are available of health states when the patients enter and leave hospitals), have brought the attention of the researchers back to the use of some proxies of hospital output.

A first proposal has been to reduce the several aspects of the hospital output to a single global value unit - incorporating, for instance, inpatient services, outpatient activities and teaching and medical research (see Ruchlin, 1973). These efforts have failed in our opinion to provide a fully validated measure of hospital output, given the strong assumptions required for adding up its multidimensional aspects to a single unit.(3)

Therefore, most of the literature has sought, instead, to employ various adjustments on certain proxies,in building up an appropriate output -measure to be used in the context of a cost-behavioural analysis.

Regarding this objective, two variables have been traditionally proposed:

a) The number of patient days. This measure would seem to be better grounded in the 'hotel' and therapeutic aspects of the hospital activity than alternative ones. However, as different authors have argued (see, for instance, Williams, 1981) the use of this indicator requires the rather dubious assumption that the longer the patient stays in a hospital, the greater the output of the health industry. In addition, this measure comes to assume implicitly a fixed input-output technology, with output measured through the measurement of those inputs related to hospital stay.

However, as a whole, this measure is less sensitive to differences in the output-mix than the existing alternatives, although this may, in itself, be a matter of statistical testing.

b) The number of patients treated. This measure captures at least one dimension of the outcome of hospital activity, although it is still a misleading indicator of changes in patients' health states (among other reasons, because it does not take into account how many of the discharged cases are due to death). In this sense, the number of cases treated is generally agreed to be a more consistent measure of the hospital output than patient days, although its use requires a greater effort on the part of the researcher to perform an accurate adjustment for the hospital output-mix characteristics. As we will see, this necessitates the development of some acceptable criteria for grouping, and weighting by case, the resulting case-mix vector.

4.4.2. Increasing Hospital Homogeneity.

In practice, researchers have responded to the need to adjust for the different nature of the hospitals' output, once one of the former basic "units", in the following terms, have been accounted for:

(i) Looking at some hospital characteristics which it can be postulated may increase hospital homogeneity within each subsample with respect to the initial group and against the similarity of the rest of the subsamples.

Thereafter, the assumption to be established for each of the subsamples is that the output of the hospitals is similar enough for the application of a single accountancy measure (e.g. number of cases).

In practice, this approach to searching for a greater homogeneity of output comes to consider external factors to the hospital activity, in order to group 'ex-ante' those hospitals with similar characteristics thought to reflect similar output-mixes. (A study of the different techniques used for classifying hospitals in this way can be seen in Klastorin, Watts and Trivedi (1978)).

Amongst the techniques utilized for this purpose, Cluster Analysis seems to provide a convenient classification procedure, as we will analyse below.

(ii) The introduction of some output-mix proxies (e.g. a case-mix vector) to be used for a direct adjustment within the overall sample of hospitals. (With respect to the use of this type of adjustment, see Hornbrook's survey, 1982).

(iii) A combination of both methods is also possible. This will be the approach finally followed in our estimation exercise (Chapter V).

<u>The use of cluster analysis to increase subsample homogeneity.</u>

Cluster analysis is a type of multivariate technique which searches for the partition of the overall sample of cases under consideration into groups, such that the degree of association is high between members of the same group and low between members of different groups.

The purpose of cluster analysis in our context is, therefore, to establish whether there exists any systematic difference between groups of hospitals in the character of their outputs. This difference then justifies the use of subsamples of hospitals according to the classification criterion for greater similarity.

This technique of sample selection for increasing output homogeneity must be regarded as an informal check on some prior views and expectations about the potential benefits derived from the grouping of subsamples, in terms of the minimization of the variance within each subgroup and maximizing the variance between subgroups. These results can then be taken into account in the model-building process, for example, by splitting the data into blocks, the composition of which reflects some institutional factors (see Cooper and Weekes, 1983).

Cluster analysis is not, therefore, a technique free from value judgements. As we will see in the exercise presented in Chapter V, judgement is necessary at several stages. Nevertheless, this technique has some advantages over other multivariate techniques: Mainly, it does not require the assumptions of multivariate normality, linearity or additivity for the determination of the groups, neither of an explicit dependent variable nor a predetermined functional form for classification purposes. On the other hand, it avoids the need to set any fixed discrete interval for the utilization of the classificatory variable. Moreover, clustering techniques are computationally efficient compared with other multicross classification techniques.

A set of variables which try to catch differences in the nature of the hospital output provided, are commonly used for these classification purposes. Examples of such characteristics include variables related to the allocation and extent of the hospital coverage; the territorial provision of care and the level of medical services offered-weighted, in case, according to some other quality aspects; the hospital manpower mix - within and between different staff categories; variety and clinical hierarchy of the hospitals services provided, etc.

In consequence, if it is found that, according to some 'a priori' views, such groupings seem to exist, that is, if it is found that hospitals with some particular characteristics appear consistently to have an output which differs from that of hospitals which do not posses this characteristic, then it makes sense to estimate different models for the two sub-sets of hospitals.

Thus, in the context of the hospital cost analysis, such a finding would suggest that the instrumental variable in the analysis differs in character across the groups, and hence the recorded costs per unit of output refer to different outputs, and perhaps different models of hospital behaviour may then need be estimated. This point seems to have been neglected in the literature on hospital cost functions. Researchers have primarily focussed on the use of some dummy variables for catching the influence of those different hospital output charactersistics, on which a single cost per unit of output has been regressed. (See for example Culyer et al., 1978).

The specific consequence of using this second approach will be analysed in examining the econometric exercise presented in the next Chapter.

The use of a direct output-mix adjustment.

This method for increasing hospital output comparability utilizes some internal factors to the hospital activity which are postulated to incorporate output-mix differences. Examples include:

(i) indirect factors related to the effects of different output-mixes on some utilisation variables, such as the use of a weighted measure of the average length of stay (e.g. for specific diagnosis categories); those types of proxy measures resulting from the application of Information Theory to data on hospital activity (see Evans and Walker, 1972); or the indicators resulting from the comparison of the mean of some hospital utilization variables with its predicted values according to the sample behaviour - that is, had the individual hospital performed as the 'average sample hospital' (see Johnson, 1977); and

(ii) direct factors, such as an output-mix vector, based on the classification of hospital cases, per diagnosis or group of diagnosis, per specialties (see Feldstein 1967), clinical departments, etc..

The utilization of a case-mix vector requires, in practice:

a) The definition and classification of the components of the vector in terms of some criterion of similarity (e.g. an iso-resource utilization criterion, the identity or similarity of the primary diagnosis, of the management path from the admission to the patients' discharge, or, ideally, in terms of changes in the patients' health states achieved).
b) Some type of weighting for the different importance of each of its elements, either within each group or between several groups. For instance, in weighting with reference to an iso-resource criterion - weights could be derived on the basis of actual (average or predicted) costs or potential costs (such as defined from certain benchmarks). The first approach to the determination of weights has the obvious shortcoming of including actual episodes of "unnecessary" care, embodied thereby as elements of "unnecessary" output. The second approach involves, on the other hand, a high degree of arbitrariness, since it accepts the value assigned by a panel of experts, say, or that derived from the experience on costs as registered by some 'pilot hospitals'

for this purpose selected.

Given the objective of constructing a cost-guide-line, for the analysis of relative efficiency, two aspects of the hospital output-mix , in particular, need to be captured with regard to their impact on costs:

(i) From the resource utilization point of view, the 'intensity' of the services administered (or their quantity per day of stay) and the 'duration'(length of stay). The 'intensity' component can be mostly related to the complexity of the case in itself, whereas 'duration' is affected by the degree of severity or criticalness of the case, which depends on the acuteness and prognosis of its disease stage.

(ii) From a resource generation point of view, the actual condition of the patient and the treatment process employed are the relevant variables. The first aspect depends on the demographic characteristics of the patients and some other related issues (clinical history, secondary problems, etc.). The second one depends on the nature of the treatment itself (therapy, secondary diagnosis, surgery, etc.).

On the other hand, the implications of both aspects (amount and cost of the treatment) in hospital expenditure have to be considered in the context of the characteristics of the hospital, that is, with reference to its specialization - in terms of the relative concentration of hospital care in certain type of treatments (hence greater sophisitication of the equipment available, medical competence, etc.) - and of the complexity of the cases treated (e.g. hospitals coping with more of the less common specialties, etc.). Both factors in the output-mix definition may have opposite effects on costs, and both need then to be analysed for an appropriate definition of hospital "relative efficiency".

In sum, clustering strategies and direct output-mix adjustments are the two prominent methods utilized in the literature for capturing hospital output differences.

In the next pages, we will introduce the rest of behavioural explanatory variables which have commonly accompanied the former output-proxies in the literature to date.

4.4.3. Operationalisation of the Cost Relationship: Behavioural Explanatory Variables used in the Literature to Date.

A review of the existing literature from the seminal studies of MS Feldstein (1967) and JK Mann and De Yett (1970) on Hospital cost-Functions, can serve to summarize the following components of our basic cost relationship or behavioural explanatory variables (see Shuman et al. 1973):

a) Locational characteristics of the hospital. These have been mostly incorporated through dummy variables, which take values of zero or one depending on whether the hospital is situated in a metropolitan, urban/non-urban, outlying area, etc.

b) Size class, usually approached by the number of beds or some other variable related to the physical component or structural side of the cost relationship.

c) Outpatient activity. The indicators commonly considered for this are the percentage of outpatient visits on total hospital activity, the proportion of staff input time employed in outpatient services or the ratio of expenses directly apportioned to outpatient clinics over the total hospital expenditure.

d) Medical Staff (or hospital-based physicians; e.g. the full-time versus the part-time component; the staff composition per hierarchies, etc.) and non medical manpower (e.g. nurses per bed, unqualified manpower per square meter, etc.). This variable has proved to be particularly correlated to the case-mix factor and to the teaching status of the hospital.

e) The range of services provided, or the number of routine and non-routine services, sophistication of the equipment, quality and technical level.[4]

f) The research and teaching status. As before, its significance may be collected by a dummy variable, or some other continuous variable, such as the number of trained students weighted by the different educational categories.

g) The output-mix variable. As we pointed out, this factor seeks to capture directly the "output" heterogeneity across hospitals, by measuring its differences, mostly in terms of the hospital complexity and specialization.

For these purposes, the following indicators have been used in the literature (see for instance, Barer, 1982).

(i) The average length of stay, in simple terms or weighted by some other measure such as the relative bed occupancy rate. This is obviously the crudest of the indicators proposed.

(ii) The share of the number of cases whose length of stay exceeds the average length of stay by a predetermined amount, or the length of stay which could be predicted from the average work-load figures.
Variations in the incidence of chronic diseases between hospitals can obviously distort this measure.

(iii) The ratio of single to multiple diagnosis per case and of surgical procedures over the total number of cases, given their postulated relation to the hospital output complexity.

(iv) The service-mix component, accounting for the character of the services available and the supply of other related facilities.

(v) An index of the actual case-load of the hospital, based on well identified diagnosis catergories (e.g. ICDA groupings). The diagnosis-mix vector can be also combined with an index of complexity (e.g. according to the radiological and surgical procedures). The weights required can be also assigned by a panel of experts, looking at some more complex treatments according to their experience.

(vi) The construction of a single "resource index", derived for each particular hospital from summing the specific proportion of cases in each group and weighting them by a set of, say, charge-based factors (Comission on Professional and Hospital activities, 1978). As we saw, these weights can be derived from the average cost or charges within a sample group of similar hospitals.

(vii) A similar measure to the former one comes from the application of Information Theory to hospital data, in order to build an Index of the relative case complexity and hospital specialization (Evans and Walker, 1972). In brief, this method uses the information gain resulting from the observation that a type of case $\underline{j}$ goes to be treated in hospital $\underline{i}$ rather than in hospital $\underline{k}$. If we assume that complex cases tend to be handled in a few hospitals with more intensive facilities and more specialized staff, we could measure "complexity" via the degree of concentration of a case type across different hospitals. Hospitals containing more of the less common specialities would be given consequently a higher 'complexity' rating. Otherwise, hospitals with a limited range of specialities would receive a lower specialization rating.

(viii) The use of Diagnosis Related Groups (Mill et al, 1976). This approach groups the patient-cases that require a similar package of hospital services in terms of the utilization of resources, as measured basically by the variable length of stay.

For this purpose, an algorithm is constructed to partition observations into groups. They are derived from the minimization of the total within group sum of squares with respect to those attributes considered to have significant impact in the hospital consumption of resources.(8)

The following interaction is examined for this purpose:

"Output" = f (length of stay and intensity of the hospital
(DRG) services required)

(---A---)

A = g (condition of the patient and treatment process)

(---B---)

B = h (age, sex, primary-secondary diagnosis & primary-secondary surgery)

(---C---)

The possibility of observing the 'C' type of attributes for each patient, allows a classification of the output-case mix on this basis. Although problems of generalization and of potential extrapolation of these results over time are still present, (6) the DRG classification represents an obvious improvement with respect to the former crude case-mix groupings.

(ix) Other more recent developments include the "Disease staging" technique, the split of "Patient management categories", the creation of "VA multilevel care groups", and the introduction of some other generic algorithms as the "AS-Scores", the "Index of severity of the illness", "MD-DADO" (see Bentley and Butler, 1980).

h) Finally, some authors have pointed out the 'output-quality' component as a separate aspect of the hospital output, different from the output-mix variable in itself. In approaching this factor, proxy-measures in relation to the character of the hotel type of service facilities, some indicators on the activity of the hospital staff – trying to capture the human factor in the medical treatment – and certain quality enhancing type of services have been proposed.

In brief, the former variables have been mostly used in the literature for the operationalisation of the conceptual issues identified earlier on in the basic cost-relationship, as defined in section 4.3. On the other hand, the rationality of the behavioural approach followed and the specification of the components usually included in the estimated cost functions have been also examined in light of the existing problems in measuring hospital output. The analysis has highlighted the need to increase hospital output homogeneity prior to compare hospital performance and how this has been pursued in the literature. Clustering strategies and direct output-mix adjustments have proved their major relevance for these purposes.

In the following section we present the selection and measurement of the variables included in our basic cost relationship. They will be used in the empirical exercise of Chapter V. For this purpose, we will discuss first the implications of the utilisation of alternative dependent variables for reimbursement purposes in view of the efficiency motivation targets to which they must serve as formulated in the previous Chapters.

4.5. THE SELECTION AND MEASUREMENT OF THE VARIABLES INCLUDED IN THE COST RELATIONSHIP

A cost function which is linear in its parameters will be specified for our purposes,[7] as consisting of:

$$\text{Cost} = \beta_0 + \beta_1 \begin{pmatrix}\text{vector of institutional}\\ \text{variables}\end{pmatrix} + \beta_2 \begin{pmatrix}\text{vector of behavioural}\\ \text{characteristic variables}\end{pmatrix}$$

$$+ \beta_3 \quad (\text{vector of the output-mix}) + \varepsilon_i$$

This section analyses the selection and measurement of the variables included in the former cost relationship, discussing its implications in the design of a PRAP.

We proceed from the selection of the average cost dependent variable (section 4.5.1) to the remaining institutional, characteristic and utilization variables (section 4.5.2).

4.5.1. The Selection of the Average Dependent Variable.

For the sake of econometric convenience (as is shown in 4.6.1), an average cost function (instead of a total cost function) is estimated.

In performing this transformation we need to assume that the 'case treated' constitutes the homogeneous unit of output in the sense specified before, and that variations in the hospital-mix can be explained by a systematic component in the cost equation.

The choice of this output variable, and hence of the average cost dependent variable has, in addition to econometric problems, some implications for its utilisation as the prospective rate or reimbursing unit. It was precisely, in the light of these implications that we selected the number of cases, once the following alternative units were examined:

a) A per capita value. This is, at first glance, the unit which seems to relate better to the concept of output defined as the individuals' health states in the community, once we take into account some other influencial factors such as the morbidity differences in a natural catchment area. As a result of this definition, a payment to the agents through a fixed capitation value could provide a maximum incentive for efficiency, in so far as it discourages both inefficiency in production and excess of utilization, since reimbursement does not vary with increasing levels. However, it involves in practice a great deal of difficulty in its implementation: This depends on whether it is possible to group the population subscribed to this policy in panels attached to specific hospitals. In addition, it may tend to discourage any cost-increasing innovation in treatment methods to a greater extent than other alternatives, due to its lesser flexibility with regard to the adjustment of the reimbursement figures to such changes.

b) A per patient-day figure. This is probably the simplest measure to apply and is to some extent appealing, given its minor dependence on differences in the output-mix vector. Its utilization may have, however, a perverse impact on the utilization factor (e.g. occupancy rate and average length of stay), and to the extent that hospitals can influence the utilization of hospital inputs, its use may reduce the chances of attaining a cost-conscious performance. Furthermore, since the first days of a patient's stay are commonly more resource-intensive, and thus more costly, than the final days, the use of average 'per diem' rates creates incentives to keep patients longer, given that these additional days will tend to be "profitable". Moreover, if rates are paid identically for each diagnosis, hospitals will also have an incentive to admit more patients that are less seriously ill, hence less expensive to treat, whenever the 'needs' constraint does not fully bind hospital admissions. (See Hornbrook and Rafferty, 1982).

c) Having rejected the former reimbursing units, the number of patients or cases treated was considered as the more appropriate measure,(8) although it must be recognized that the hospital revenue allocation should not be based on this factor alone.

The choice of this output unit is based on the fact that, in principle, a reimbursement per case can encourage efficiency in the supply of care, for instance, through reductions in the average length of stay.

It must be recognized, however, that it generally provides zero or perverse incentives concerning unncecessary admissions . Its utilization is then particularly acceptable if it is followed by quality-monitoring policies since they too, take into account the patient case as the unit of analysis.

At any rate, the use of this measure will need to be accompanied by an adjustment for case-mix differences, since, otherwise, it could encourage discrimination in favour of the admission of short-stay cases and, hence, it may dilute 'output-quality'.

On the other hand, the utilization of data on patient bases makes possible thereafter the construction of alternative values for hospital analysis at different levels of aggregation. In fact, research is being developed in the United Kingdom (see Magee et al. 1981) on clinical, speciality and departmental costing and budgeting, and in the United States on patient and diagnosis costing and reimbursement (see Thompson et al., 1982, on Diagnosis Related Groups).

Although the use of values based on cost centres at the specialty or departmental level remains relatively attractive, in the sense that it makes decision-makers directly accountable for their financial actions, statistical problems oblige one to be cautious about its direct utilization here. That is, gaps in the information at present available rarely allow for such accurate attributions of costs required for the corresponding apportionment of responsibilities. (See for this purpose the recommendations of the Körner Working Group in England, 1983, in relation to changing in this direction the Hospital information system).

These problems can be found, in particular, in using the patients' "diagnosis", even after taking into account the special characteristics of the hospital and the nature of the treatment process administered.

Although this basic unit involves in itself the consideration of case-mix differences - given that its measurement can also be refined by patient related aspects, such as the inpatient demographic characteristics, the character of the diagnosis and therapeutic aspects of the treatment, etc. - its utilization raises, at some stage, the question of grouping and weighting. Since groups and weights are derived on the basis of existing information on actual utilization, they are unlikely to be valid for their utilization in contexts different of those from which they have been initially derived.

In short, there are both theoretical and empirical reasons on why an average cost per case value may be preferred as the unit of analysis (once one has accounted for hospital case-mix differences) and as the reimbursing unit for purposes of a prospective revenue allocation policy of the type described in this thesis. For these reasons, the total number of cases treated in the hospital will be taken as the output unit and used in the denominator of the average cost variable.

In the remaining pages of this subsection, we will examine the measurement of the total hospital costs, which constitutes the numerator of the ratio defined as the dependent variable.

Total hospital costs.

Measurement problems stem here from the way in which capital costs and expenditure on outpatient services are dealt with within total operating cost figures. In fact, there has been in the literature to date no discussion about the need to include in the total operating costs the maintenance and normal replacement costs of the existing equipment. In general, the problem is whether hospitals apply in practice homogeneous criteria for the determination of depreciation rates on building costs and purchases of major equipment. If such criteria do not exist, it would seem advisable to deal just with the remaining operational costs for purposes of hospital cost comparisons. However, since total costs per unit of output reflect the influence of the rate of utilization of hospital facilities on additional capital expenditure the omission of capital depreciation might, in principle, bias over time hospital expenditure figures.

The most significant result of omitting capital depreciation seems to be to underestimate the extent to which total cost per output rises with scale. This is particularly the case when the number of patients, rather than the number of bed-days, are used. In fact, to the extent that larger hospitals have lower cases per bed, they may consequently have higher capital costs per case, unless they achieve very substantial economies of scale from larger capacity production, etc.. This problem then also requires the use of an appropriate measurement of the institutional scale variable.

But in our context, we can use total operating cost figures given that depreciation costs are at present homogeneously regulated by the central reimbursing Authority. This fixes the depreciation rates and determines across the public hospitals the values on which rates have to be applied.[9]

Therefore, in ending this subsection, it should be pointed out that the total operating costs (in the way defined above) per number of cases treated in the hospital constitutes the average cost variable selected in our basic cost relationship.

The remaining explanatory variables will be presented below.

4.5.2. The Explanatory variables: Institutional, characteristic and Utilization Variables.

Among the independent or explanatory variables included in a behavioural cost function, 'institutional', 'characteristic' and 'utilization' variables have to be considered.

Institutional variables:

The "size" variable is the most widely used of those explanatory variables. In principle, it can be approximated by the 'output-productive' capacity, the hospital staff, or the 'physical' size of the hospital (e.g. beds, capacity, or square meters of floor). In practice, the interest in using some type of 'size' factor arises from the objective of tesing the existence of scale economies in the hospital production of care (see in this sense Appendix One to the next chapter).

With regard to the proxy measures for hospital scale, some issues need to be noted. In this sense, Friedman[10] has pointed out the potential presence of a 'regression fallacy' in the analysis of hospital costs. This is the case when output is used as the proxy for scale, and it would be due to the fact that depreciation costs from conventional accounting procedures will not be fully measured. In fact, when output rises, plant depreciation does not usually increase in the same proportion. Thus the unit cost of production will be, "ceteris paribus", negatively correlated to the output-capacity ratio.

On the other hand, if hospital size as indicated by the number of beds is used as the institutional indicator for scale, one must take into account the problems arising from the fact that this variable is typically correlated with several other explanatory variables in the behavioural cost function. In addition, there may exist a functional-relation problem of this variable with the dependent variable: as hospital size increases, the cost of acquiring and maintaining any increase in the types of care provided with specialized inputs will increase too. If, as is usually the case, this increase is not accompanied by a proportional rise in the number of cases treated within each specific category, those inputs will no longer be increasingly utilized. Therefore, the costs of acquisition and maintenance will not be spread over larger volumes of output. When this is the case, an increase in hospital size will lead to an upward shift in the average cost curve rather than a movement along the same curve (Feldstein, 1967).

In addition, the physical size variable may imply two different cost effects with regard to the hospital utilization factors:

a) If larger hospitals assume responsibilities for referrals within wider geographic areas, or provide a more sophisticated type of care, the cost-distribution aspects of care provision may rise with size. This may be due,for instance, to the sparsity of the area served,from which higher costs result from greater hospitalization in avoidance of emergency risks on those territorially based services. This factor may then lead to the provision of inpatient care to persons who could otherwise have been treated as outpatients.

b) Lower unit costs may be recorded, on the other hand, in so far as scale economies are significant, due to the advantages of larger plant capacity. They are likely to stem from existing output indivisibilities.

Moreover, a wide range of indirect effects on other independent variables can appear too, For example, some collinearity problems may appear between (i) hospital size and the flow of cases per bed available and with respect to the quality aspects of the output(although, in this last case, with a less unabiguous relationship),and (ii) hospital size and the case-mix variables, training programs and other utilization variables (average length of stay and bed occupancy rate), mostly with a positive association.

Finally, a direct association between scale and input substitution may affect costs if:

a) larger hospitals can afford greater opportunities for manpower substitution, given some rigidities among the factors of production, or

b) larger hospitals are less efficient in actually combining inputs, in adjusting themselves to different utilization rates with more difficulty than smaller hospitals, or if the incentives for efficiency steadily diminish with size.[11] In this sense, technical and economic efficiency may be pulled in opposite directions.

Therefore, we can conclude that the entire effect of "size" on total average cost figures is, in general, rather ambiguous from an 'a priori' point of view.

Among the institutional proxy variables, other external and enviromental factors can also be introduced, the most common of which are hospital structure, 'age', geographic location, inter-hospital cross-boundary flows and certain external factors affecting labour market conditions (i.e. cost-amount of the factor inputs). In addition, the teaching status of the hospital - which, as we saw, may present some indirect effects on hospital size and medical staff[11] - and some other general population attributes of the geographic area served can also be included.

For our purposes, most of the institutional aspects mentioned above will be used as the basis for constructing subsamples of homogeneous hospitals,following a cluster strategy. Thereafter we will select the hospital size variable 'number of beds' as the explanatory variable in the regression analysis for explaining actual variations in hospital costs.

The utilization and characteristic variables.

The vector of utilization and characteristic variables are intended to reflect the individual performance of the hospital, once we have accounted for the institutional factors which surround its management. This vector can include:

a) The flow variables related to the actual utilization of the existing facilities and/or some direct estimators of the labour and capital inputs available. They try to reflect factors related to the under-utilization of the existing capacity, the over-provision of some services, inefficient levels in production, etc. They are of crucial relevance to the definition of an accurate base-line for the analysis of efficiency for comparing hospital costs.

For these purposes, variables such as the turnover interval, the bed occupancy rate and the average length of stay are widely used as indicators of the actual performance of the hospitals.

b) The case-mix load. The introduction of this factor is required given the use of an average cost per case as the reimbursing value.(13) It accounts for actual differences in the hospital case-load.

The influence of these variables cannot be generally postulated in a single unambiguous manner. Existing collinearities between the variables included in the case-mix vector and other institutional, characteristic(14) and utilization variables prevent that identification. Moreover, the utilization variables included in (a) are linked to one another by definition through a single relation. Thus flow of cases per bed equal the bed occupancy rate (x365) divided by the average length of stay. This relation requires then a careful selection of the variable included and the one which is omitted - and which is consequently allowed to vary in the regression - if a careful interpretation of the influence of each single variable has to be made.

This last point implies important observations for the efficiency analysis in the utilization of the hospital facilities. In fact, from the relationship described, it is clear that an extra case can be treated by shortening the average length of stay or by increasing hospital capacity. Both can be legitimately termed, in principle, "changes in utilization", the first referring to changes in the utilization of the existing beds and the second to changes in total capacity. As we will see later, the way of achieving this apparently similar result has different cost consequences, depending on whether these changes relate to costs per case, per patient-day or per bed (see Appendix V. 1).

These differences will need to be taken into account in the analysis of the hospital relative efficiency.

In addition, the following points can be made with regard to the utilization variables mentioned above:

In relation to the average length of stay variable (LOS), we should note that cost per case is just the cost per patient-day adjusted by the length of stay. Therefore, the former utilization variable can easily be traded-off with average costs through changes in the length of stay. As a result, the average length of stay has the obvious drawback that it can easily be manipulated by hospitals for reimbursement pruposes, particularly if cost per patient-day is considered as the reimbursing unit.

If the bed occupancy rate (BOR) were the variable chosen, we would need to point out the sources of variation in the ratio of utilization of beds with respect to the total number of beds available, before making any statements about relative efficiency in the utilization of hospital facilities. Thus, BOR may increase due to higher LOS, with a constant flow of cases per bed, having no positive implications for efficiency improvements if we assume that those extra-days in hospital make no differences to the patients' health. On the contrary, it may give an indication of "inefficiency" given the unnecessary utilization of resources. Otherwise, BOR may also increase because of higher values in the flow of cases per bed available, holding LOS constant. This change can be considered, in principle, as an efficiency contribution by improving the utilization of hospital facilities. This is due to the fact that it allows for the treatment of additional inpatient cases through the increase in the utilization of the already available hospital capacity.

In conclusion, the Case-Flow variable (FL), or number of cases per bed available over the year, is, in our view, the utilization variable with less "perverse" effects for reimbursement purposes, once the institutional factors and the hospital case-mix vector are taken into account. In this way, it will be incorporated in our estimation exercise as the second explanatory variable.

The second "characteristic" variable, as advanced in (b), is the case-mix load. As said, its inclusion is particularly relevant given the decision to use an average cost figure as the dependent variable in the estimated regression. This vector will try then to recover part of the information lost by the use of an average cost per case value as the dependent variable, by being included as a separated explanatory variable.

Some authors have omitted this adjustment by assuming that the "output-mix" of the hospital is already captured by the actual stock and flow variables (Carr and Feldstein, 1967).(15) It is not at all evident, in our view, that the case-mix complexity and hospital output specialization can be systematically captured by the hospital size in the way approached by those authors,(16) nor that all their effects on the intensity utilization of resources are collected by the hospital actual utilization rates.(17) Regarding the facts, the balance of the argument seems to support the explicit inclusion of the vector of the actual case-mix load, according to the nature of the endogenous variable previously selected for estimation purposes.

As we will see later on, econometric problems will advise the use of a percentage version of this variable instead of absolute figures.

In sum, the size 1 variable - proxied by the number of hospital beds available - the flow of cases 2 per bed and the specialty 3 case-mix vector in percentage terms, will be used as the set of explanatory variables selected for the estimation of the hospital behavioural cost function.

Before going to the actual estimation exercise, we examine in the next section some methodological and econometric problems which arise from instrumenting this particular approach to the hospital relevant expenditure and from the use of multiple regression analysis techniques for its estimation.

4.6. ESTIMATING THE COST FUNCTION: SOME RELEVANT ISSUES.

In this section, we examine certain problems which are not always made explicit in the literature. They are commonly present in the estimation of the type of regression equations used for the calculation of a base-line for hospital cost analysis in the way proposed above.

In this sense, we will refer to those issues derived from the utilization of estimation techniques to predict hospital expenditure from the set of determinant factors assessed as relevants for reimbursement purposes.

As we will see in the next pages, this sort of exercise contains methodological issues and involves econometric problems which need to be taken into account when examining the strengths and limitations of such an approach.

4.5.1. General Methodological Problems and Responses.

It should be noted that the definition of a base-line of 'standard expenditure' from a predicition based on the results of the estimation of equation will refer, in principle, to the figure that an hospital unit would spend if it conformed to the regression formula; hence, if it spent the average of the hospitals which have the same characteristics, such as approached by the set of exogenous variables considered. Therefore, it does not provide in itself an absolute 'efficiency locus' for theoretical comparisons.

For these purposes, the methodology of analysis is to some extent tautological, since the determinant component of what has to be spent is derived from what is already spent. In other words, the system contains a "backward looking" character, since it looks at the past to predict the future. In order to fully validate this method we need to assume that,

(i) on average, hospitals are already providing services at a standard level - as defined for instance by existing planning policies- and at standard values - as monitored respectively by the operational policies inforce. Hence, their expenditures on a cross-section basis are equally worthy of consideration; and

(ii) the statutory obligation of the services which have to be provided is the same for all the hospitals.

In the second place, and as a result of the 'backward-looking' character, a cumulative causation effect may appear in the absence of specific measures of correction. In fact, those hospitals which have managed more resources in the past will obtain from the prediction relatively greater amounts for the future. Moreover, if the initial data system contain biases, these will be perpetuated, or even accentuated, by the regression analysis (Godley, 1976). As a result, in the case of underestimating, say, a cost difference through the regression formula - that is, if we fail to take into account that it costs more in certain areas than others to provide a given level of service with a given degree of efficiency, we will be introducing a persisting pressure to reduce the level of services in these areas.[18]

In the third place, econometric problems, which we explore further in the next section, can affect the conclusions derived from the estimation and interpretation of the results obtained. That is, in addition to the caveats that the selection of the 'relevant' variables may offer themselves, the techniques utilized for their estimation may not lead to an unambiguous interpretation of the results (for instance, and typically, in terms of the marginal impact on costs from unit changes in the value of each explanatory variable).

Viewed as a whole for reimbursement targets, these problems may not matter very much for an "average" hospital, since, by definition, overall expenditure will be largely accounted for by the coefficients of the estimated equation; that is, those predicition errors will be smaller for average hospitals as compared with other outlying hospitals. But for deviant hospitals, there could be no confidence that, say, a relatively low allocation because of one of the factors will be compensated by a relatively high allocation because of other factors.

These problems should be then explicitly appreciated in considering whether or not the equation as a whole has the properties of generating the 'appropriate' reimbursement.

Finally, the uncertainty which may surround the results of the revenue allocation process formulated in these grounds could be argued to create certain risk-aversion - for instance, in terms of hospital technological innovations - and a lack of planning against the running needs of hospital operation. This caveat may be also recognized within a bargaining framework for the allocation of hospital revenues.

In dealing with those points and providing operational responses, it should first be remembered what the objectives of the exercise are. They consist, in short, of providing an external outline for budget evaluation and revenue allocation, out of the pace imposed by past expenses and from existing political pressures and lobbies for its allocation.

In this sense, with respect to the first observation, it can be replied that "in general terms, what the decentralized average sample unit has found necessary to spend on the whole, is a sufficiently satisfactory measure of what it was needed to spend responding to a local view of the population needs". Moreover, this "backward-looking" character is common to all those methods based on data referring to actual values, which seem to be, after all, the only feasible alternative given data limitations.[19] What we do in addition here is to redefine those values in some more appropriate ways.

At any rate, the core of the limitation pointed out above for this type of exercise remains valid in general terms, but it has to be well understood. What one is going to obtain at best from the estimation is an average cost function instead of the locus of economically efficient input-output combinations. Nevertheless, if the objective is to use the estimated parameters to predict the amount that "on average" can result from the use of some given parameters, obviously, an average cost function can plausibly be used. If the objective were to obtain from the pure marginal costs that under real conditions derive from the increase in the use of one factor, then the average cost function would yield a valid answer only in a very rare case: This is, when at any rate of factor utilization, the slope of the actual cost function is the same as that of the hypothetical function along the slope of the line (e.g. as measured, precisely, by marginal costs). (20)

4.6.2. Econometric problems.

In addition to the problems discussed above, which are fundamentally associated with the nature of the regression approach to the method of assessing 'relevant' expenditure, econometric problems may stem from the way in which the regression equation is estimated.

In estimating the hospital cost function in a first step by Ordinary Least Squares (OLS) we need to deal first with those factors which may lead to violation of the basic assumptions underlying OLS. In a cross-section approach to the estimation of the overall sample, these factors are likely to be:

a) The violation of the "zero mean" expected value of the random component of the expression.

This violation may appear, in our context, when we incurr in misspecification errors in the systematic part (e.g. by omitting a variable which should be included in the basic cost relationship), in the functional form (e.g. indicating incorrectly the specification of an equation as linear in its parameters), or in the structural stability of the stated relationship (e.g. pooling observations which belong to different cost behavioural models, as may be the case from the inclusion of teaching and non-teaching hospitals in the same regressed sample).

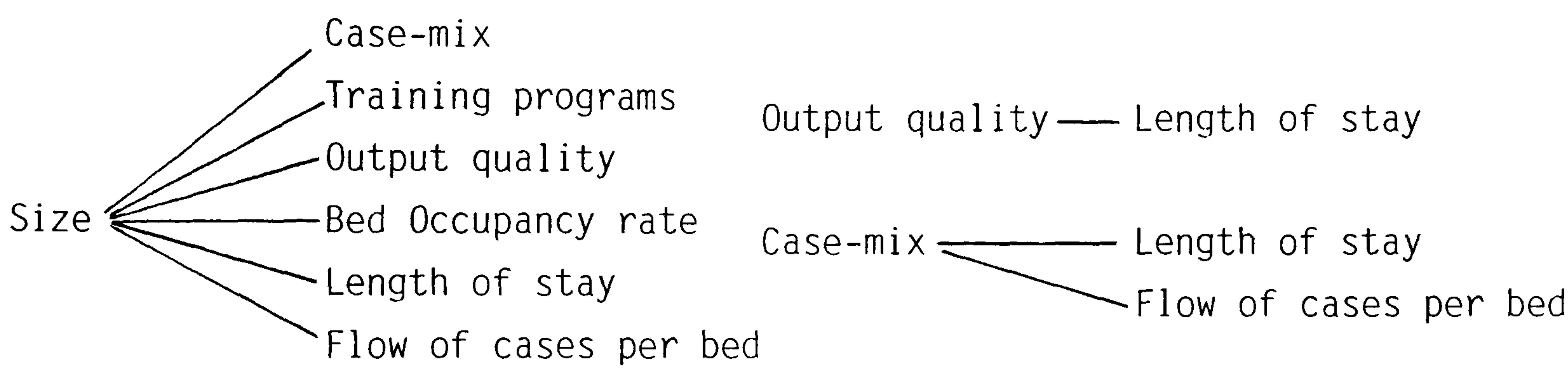
Potential bias, inconsistency and inefficiency of the OLS estimators are in this case the common consequences of this violation.

b) The violation of the hypothesis of randomness or non-stochastic property of the regressors.

This problem commonly appears when regressors are mutually correlated, as happens in the case of multicollinearity. As we have argued earlier this is the case in our context of some of the utilization and characteristic variables.

The violation of the non-stochastic property of the regressors may create large sample variances, and lead to imprecision in the estimated parameters. It has the effect that the partial regression coefficients cannot be interpreted any longer as a measure of the change in the dependent variable due to a change in the independent variable, other things being equal. (Obviously, "other things" are not equal due to the problem of collinearities).

In brief, the following set of collinearities have been discussed in the literature estimating hospital cost functions:



Therefore, if some linear relationship exists between those independent variables in the model set up earlier, given that most of the variation of the highly correlated regressors will be common to them, this will leave little variation unique to each variable.[(21)] This means that OLS procedures have little information to use, and any estimate based on little information cannot be held with much confidence.

As a result, an econometric trade-off has to be made on these cases at some stage of the estimation between the collinearity problems arising from the inclusion of some of those variables, and the misspecification erros resulting from their omission. This invovles a compromise of whether we can reduce the variance of the remaining estimates, in omitting those variables suspected of being collinear, by enough to compensate for the additional bias which may be otherwise introduced.

The presence of multicollinearity has been particularly argued, amongst the elements of the output (case)-mix vector. In dealing with this specific problem, Principle Components Analysis has received most of the attention.

The Principal Components technique apppears for this purpose as an appropriate procedure, because it seeks to achieve a sufficient reduction in the number of parameters (say, case categories) by "combining" instead of excluding variables, and thereby avoiding the resulting potential bias from that omission. Principal components are created as linear combinations of those variables which explain the maximum variance and keep the orthogonality property among them. In this sense, this technique eliminates the problem of multicollinearity and it offers an implicit solution to the problem of grouping, weighting and aggregating the different factors used in a case-mix measurement. However, the use of the regression coeficients, resulting from the inclusion of principal components as weights, can be made only with difficulty in our context. Indeed, they do not provide a clear-cut interpretation for building prospective rates of reimbursement.

It seems then advisable to search for the diminution of the multicollinearity problem amongst the case-mix variables in some other indirect ways: for instance, through the use of the share of each case-diagnosis category in the total, since evidence is available of a substantial reduction in this way in the value of their correlation coefficients.[22]

c) The violation of the assumption of homoskedastic disturbances. This violation occurs when the disturbances do not have constant variance. This has been currently argued to be the case when hospitals of different size are included in the same sample (Lave and Lave, 1972). The reason in this case for the violation of the assumption of homoskedastic disturbances is that there seems to be more variation in the services offered, and hence in the output-mix, in larger hospitals than in smaller ones and, therefore, that the variance of the disturbance term in the estimated equation changes with hospital size.[23]

In these cases, an observation associated with a relatively high disturbance varience is likely, again, to contain less information about the deterministic relation between the explained and explanatory variables than one with a minor disturbance variance. As we have seen, OLS will provide in these cases inefficient estimators in small and large samples, although they will retain the properties of unbiasedness and consistency.[24] This result is due to the fact that OLS procedures will place equal weight on the observations which have larger error variances than those with smaller ones. This missleading weighting occurs because the sum of squared residuals associated with large variance error terms is likely to be substantially greater than the sum of square residuals associated with low variance errors. It is then reasonable that the observations with large variance errors should be differentially weighted. This rationale is in the basis of the alternative Weighted Least Squares option selected in our estimation exercise, that offers a solution equivalent to the one resulting from the deflection of the whole equation by the value of the independent variable which causes heteroskedasticity in the disturbance term. A similar reason lies behind the estimation of the average cost value of the dependent variable, assuming that the level of output and hospital size move in similar directions.

A second general issue is that of the identification of the equation. So far as behavioural costs collect supply aspects closely related to demand factors, this may cause an identification problem.

The predominant view has come to postulate in this matter some sort of Say's Law: that is, that supply creates its own demand in the provision of inpatient health care and, in this sense, demand factors could be correctly ignored. It may not always be so absolutely clear that supply factors alone determine the quantity of care demanded, since, among other reasons, as Berki (1970) has pointed out, it may be difficult to distinguish between the aspects of creation of demand and those of elimination of the excess of demand.[25]

However, given the use in our context of a behavioural cost function instead of a conventional production function - which may raise additional problems of separability - we can plausibly assume that the influence of the resource constraint has an increasingly prominent role compared with that of the "needs" constraint. But if this is not the case, problems of error misspecification, and hence biased, inconsistent and inefficient estimates, will again stem from the results of a conventional OLS estimation.

A final point can be made here with respect to the possibility of the existence of sample selection bias.

In short, this problem may arise in our context from the construction of different subsamples. Although they exhibit in principle a higher degree of internal homogeneity, its grouping criteria may affect the functional form and structural stability of the overall regression equation. In these cases, the implicit variables which make for the interdependence among subsamples may need to be examined. For instance, when the "cluster" definition is based on the hospital size, variable factors such as the extent of the area actually covered by the hospital, the level of transfers of patients between hospitals belonging to different groups, and the influence of territorial indices of the cost of living on the voluntary allocation of labour inputs, and some other externalities in general, can cause that type of distorsion, given its association to 'size'. Although this problem does not seem to have been considered as extremely important in the literature, it may need to be taken into account in some cases.

In sum, from the analysis of the problem involved in the definition and measurement of the variables selected in our basic cost relationship and of the econometric problems surrounding the estimation of the regression equation specified above, in the next and final section we will conclude some considerations and recommendations which will be taken into account in our estimation exercise presented in the next Chapter.

4.7. CONCLUSION AND SUMMARY

The following considerations will be incorporated in the estimation exercise of Chapter V:

a) Given the nature of the factors currently causing the violation of the hypothesis of homoskedastic disturbances, required for an OLS estimation - and which are likely to be present in our context - we need to test for heteroskedasticity.

In fact, given the cross-section nature of the sample and the range of hospital size, a differential weighting for those observations with larger variance error terms may be required.

According to the analysis performed in this Chapter, the number of hospital beds is the main candidate for weighting, and so for deflating the initial equation, opting consequently for a Weighted Least Squares type of approach. In addition, the use of an average cost function, as far as hospital output is related to size, may also involve an implicit deflation which can reduce the problem of heteroskedasticity too.

b) In the final specification and estimation of the regression equation, a trade-off may be needed between the problems derived from multicollinearity and those resulting from the omission of the correlated variables - or specification errors; that is, between the possibility of reducing the variance of the remaining estimates by excluding those variables suspected of collinearity, by enough to compensate for the bias which otherwise will be introduced.

The use of principal components, prior to applying clustering procedures(26), and the utilization of a percentage version of the speciality case-mix vector as the explanatory variable included in the regression equation,may serve too for dealing with the multi-collinearity problem.

c) The problem of a potential sample selection bias will be considered by re-selecting the initial sample and searching for external institutional characteristics with a major effect on hospital singularity. The existence of a constant functional form and structural stability for the subsamples created will be consequently validated, testing for misspecification errors. (Thereby, the greater internal homogeneity achieved for the subsamples may have also indirect effects in reducing the problem of heteroskedasticity).

d) However there remains the problem of the potential simultaneous equation bias in the estimated regression. The lack of exogeneity of some independent variables will remain, in our view, as long as we are unable to define a more complete model of hospital behaviour, offering a better understanding of the full network of 'inter-relations'. Research is certainly required in this field.

In consequence, the following variables will be considered:

a) We will use an average cost per case as the value of the cost per unit of output dependent variable.

b) The specification of the case-mix variable is bound then to have a crucial importance, given that most of the explanatory power in the variation of the dependent variable seems to rest in this factor.
We should aim, in principle, to build a case-mix proxy for hospital output not tautologically defined only in terms of actual costs or performance levels, especially when grouping, weighting and aggregating the initial vector of values under postulates of functional homogeneity. The measure should be, in addition, sensitive enough to collect differences in the hospitals' output, and consistent in order to be used for reimbursement purposes. At present, the data available advise the use of specialty case-mix figures accounted in terms of percentages of the total, given the degree of multicollinearity otherwise involved among them and with respect to the rest of institutional and utilization variables.

c) The flow of cases per bed available is, in our view, the most satisfactory measure for the actual utilization of the existing hospital facilities. At any rate, this single variable cannot alone aim to capture fully the impact of the operational facture on hospital "efficiency", since its explanatory power depends on the specific assumptions made about the variation of the other related utilization variables.

The variables described above will be considered in the estimation exercise, finally presented in the next Chapter.

FOOTNOTES:-

(1) Under both systems, this required adjustment may be supported by an engineering analysis of activities at the micro-level or through an overall Functional Value approach, at a more aggregated one.
The first method consists of the study of changes in the use of facilities, the introduction of new operational methods, coordination of clinical services, manpower assignments and equipment supplies, in order to increase output per unit of inputs, or to reduce inputs per unit of output - seeking in this way to reduce expenditure.
Functional Value analysis is a technique to achieve restraint in overheads. For this purpose, expenses are broken down into cost centers under manageable budgets. Within these mini-budgets, percentage reductions or cost-savings are requested from each of the budget managers. These measures are further evaluated by the analyst in terms of 'risks and savings'. The results are ranked in this basis, asserting the most beneficial ones with respect to their cost-saving targets.

(2) This residual analysis has also been studied in other fields. ROSS and BURKHEAD (1974) have proposed this technique in a time-series context in searching for the specific contribution of changes in work-load, cost of factor inputs and productivity improvements from data on local public expenditure. They take the view that productivity is what is left once alternative explanations are eliminated (accounting for quality factors and measurement errors too). The rationale of comparing for productivity analysis a calculated or projected value to the actual expenditure figures has also been proposed by FARRELL. Farrell's functional approach consists of fitting a production (cost) function through all the points and use the deviations from fitted to the actual observed values as the criterion for the analysis of performance.

(3) RUCHLIN searches for a measure of "relative value" from an index of 'relative service intesity' and a 'resource absorption coefficient'. The first component is derived from the ratio of registered staff in the particular area considered in relation to the total number in the medical area. The 'resource absorption coefficient' is taken as a multplicative function of the expected length of stay and the expected cost per day. Both factors are defined in terms of the particular activity with respect to the total.
In a similar way, the College of American Pathologists have built a measure of this nature, the so called "Californian Relative-Value Scale".

(4) Some authors try to collect, in addition to the presence of some non-routine services, those other postulated 'enhancing quality' services by weighting psychometrically its expected relative effects on costs. However, it should be clear that this variable aims to collect the potential supply of a set of services rather than the actual hospital case-load.

(5) Major groups are split in a number of subgroups which are classified according to Primary and Secondary diagnosis, Primary and Secondary surgery, age and sex, in a recursive way through a step by step procedure. Groups are derived according to the percentage reduction in the explained variance of the dependent variable by minimizing the total within group sum of squares.

For each of the resulting groups, the "classify algorithm" produces, afterwards the mean length of stay,the size of the sample and the number of observations for each included independent variable. The procedure stops when the characteristics of the groups do not show satisfactory properties, such as a certain minimum size.

(6) Most of the problems involved in the application of a DRG classification stem from the need to revalidate its results in order to be extrapolated across the sample of different hospitals and over time, testing that similar patterns in the utilization of resources are held.

Since DRGs are based on the actual length of stay and the actual utilization of resources, hospital "inefficiencies" are incorporated entirely. The fact that DRGs employ actual performance patterns as characteristic variables, instead of those derived from a theoretical production function, leaves open the issue of the extent to which the hospital is accountable thereafter for its own efficiency.

(7) Although still linear in the parameters, a cubic adjustment could be, in principle, postulated, in a consistent way with the theory, of the type:

$$\text{Total cost} = \text{constant} + \beta_1\,(\text{Output}) + \beta_2\,(\text{Output})^2 + \beta_3\,(\text{Output})^3 + \text{error term}$$

It could expect that the value of the estimated signs of the parameters to be positive for β_1 and β_3 and negative for β_2. In addition, if a saddle shape has to appear - and hence the positivity of the marginal cost at the minimum - the condition will be given by

$$B_2^2 < 3B_1 \cdot B_3$$

(8) We will quote here for illustrative purposes some defenders of each approach: "The number of cases - says M. S. Feldstein - is the relevant output-proxy, perhaps by exclusion, once we look at the potential 'trade-offs' of the cost per patient day variable and length of stay". An opposite view is held by Ingbar and Taylor

(op.cit) They argue in favour of the use of the number of patient-days as the output measure, facing their critics by pointing out that for the same reason that 'beauty' does not need to be regarded as the output of a beauty saloon, so the hospital output does not need to be specified in terms of its effects on the patients". Finally, J.R. and J.B. Lave seem to adopt the view that "the best measure of output cannot be set a priori and it should be, in principle, the one which proves most homogeneous across the hospitals, given the cost function specified". In a first paper (1966), the number of patient-days was selected as the output-proxy. Later, in a second paper, they seem to reverse their initial choice in favour of the number of cases as the proxy variable for the hospital output.

(9) However, once implemented the Hospital Reform as today proposed, the former problem may raise, given that the new Bill aims to attribute an autonomous sphere of decisions to the hospital management in the field of these economic issues.

(10) Therefore, the regression of unit costs on output (scale) will be based downwards, towards constant marginal costs. M.S. FELDSTEIN points out, however, that the problem of higher utilization rates, which appear to decrease unit costs, may be due not so much to the problem of defining costs as to the question of defining an appropriate unit of output.

(11) In this sense, CARR and FELDSTEIN (1967) postulate that utilization and scale move together: small hospitals would operate more efficiently at lower than average relative occupancy in order to adjust to demand fluctuations more economically.

(12) See DUSANSKY R. M. and P. J. KALMAN (1974) in "Towards an economic model of the teaching hospital". Journal of Economic Theory, no. 7.

(13) For a good summary on this issue, see M. L. Barer "Case-mix adjustment in Hospital cost analysis". Journal of Health Economics, no. 1 - 1982.

(14) According to the empirical evidence available, these collinearities can be summarized in the following terms:

a) An inverse relationship for the flow of cases per bed (FL) is usually found with respect to size (number of beds) and the complexity of the output-mix, here mainly related to its characteristic variable "average length of stay".

b) with reference to the length of stay (LOS), a positive relationship is often derived with respect to size and the case-mix complexity vector(especially if a 'diagnosis related group' type of approach is followed).

c) The bed occupancy rate (BOR) variable usually shows its correlation to some other institutional variables, although here the evidence is less unambiguous.

(15) In this sense, see W. J. CARR and P. J. FELDSTEIN: "The relationship of costs to hospital size" Inquiry - 1967. In fact, they argue that the output-mix effect is already captured by the hospital size and the range and scope of the service facilities available.

(16) In the literature, some evidence has been pointed out for each of these suggestions. But it seems that if hospitals dealing with chronic diseases are excluded from the sample, "complexity" seems to increase with hospital size. This seems to agree with the data available about the cross-boundary flow of patients, moving from hospitals of different size for more complex treatments. The effects of the "specialization" factor does not appear to be, however, so unambiguous.

(17) Longer average length of stay can be, in fact, associated with those hospitals providing a more complex output-mix. But hospitals specialization may reverse the initial effects by reducing the turnover interval or by increasing the flow of cases per bed available. Moreover, "complexity" and "specialization" may not be contemporaneous for all the hospital inputs. But if this is not expected to be the case, the results of applying Information theory to the case-mix measurement can provide a more relevant approach.

(18) In the Spanish case, the study of the interterritorial differences in the provincial index of cost of living gave a ratio of 1.70 between the mini-max values. This proves the importance of the factor indicated.

(19) Some authors - see W. GOODLEY, 1980 - have argued for the need of a "client-group approach". This view requires the definition of some standard units for the services provided and their associated costs, independently of the actual levels provided and the amounts spent; hence, the definition of a minimum, average or standard value; the assessment of the objective in terms of inputs availability or outcome results, and the need to account for the influence of cost variations on a geographic basis.

(20) This condition may be approximately met if for any input factor combination, the average absolute differences between the ideal output corresponding to that combination, and the actual observed rate of output is the same for all the input combinations. Otherwise, if the unit deviates from the ideal output rate by an equal percentage for all the input combinations, the empirical cost function would tend to understate the hypothetical cost-efficient allocation of the inputs used.

(21) Some of the collinearities resulting from the estimation exercise presented in Chapter V can be here advanced for illustrative purposes. They show the following correlation values:
a) Hospital size - number of beds has a positive coefficient of 0.487 with respect to the average length of stay.
b) 'Size' has also a significant negative correlation coefficient with respect to the ratio of physicians to nurses.
c) The number of hospital beds has a positive correlation coefficient of 0.411 with respect to the flow of cases per number of beds available ratio (against what it appears to be asserted in the literature).
d) Finally, no significant correlation seems to exist between the bed occupancy rate and the number of beds available.

(22) See footnote 32 on page 230 - Chapter V

(23) We will examine this assertion and the acceptance of this assumption in the econometric exercise and Appendix of the next chapter.

(24) The OLS estimation will, in fact, overestimate the variance of the errors, increasing the precision of the estimators otherwise required (see J. JOHNSTON "Econometric Methods", 2nd. edition McGraw Hill (1972).

(25) In this way, S. E. BERKI ("Hospital Economics", 1972) has pointed out the need to use an interaction function for the aggregate demand of each output category and the number of hospitals prepared to provide each particular type of care in a given area in order to introduce demand factors into the final determination of output.

(26) Cluster techniques have been also used, aiming to achieve an aggregate measure of the case-mix vector, in the sense, for instance, of grouping those categories with similar impact on costs. However, once again, this clustering procedure cannot solve the problem of collinearities because, so far as the marginal costs derived are based on the original estimators, they retain the basic structure of the collinear variables. Hence, it reduces the number of parameters to estimate but it does not reduce its multicollinearity.

CHAPTER V THE CONSTRUCTION OF A COST BASE LINE FOR PURPOSES OF ANALYSIS OF RELATIVE EFFICIENCY IN HOSPITAL EXPENDITURE .

5.1. INTRODUCTION

- Outline of the chapter.

5.2. DESCRIPTION OF THE HOSPITAL DATA AVAILABLE

- The Bases for its Reform

5.3. THE SELECTION OF THE SAMPLE FOR THE ANALYSIS

5.3.1. Some initial groupings.

5.3.2. Resulting subgroups from cluster analysis.

- Selection of the variables.

- The use of cluster analysis to increase subsample homogeneity.

- The sequence of the exercise.

5.4. DESCRIPTION OF THE COST EQUATION AND RESULTS OF THE ESTIMATION.

5.4.1. Empirical results of the estimation of the cost function.

5.4.2. The results of the prediction.

5.5. ANALYSIS OF THE RESULTS OF THE PREDICTION.

5.5.1. The need for a geographical relative price adjustment.

5.5.2. The simulation of a PRAP 1979-1981.

5.6. TESTS.

- The institutional pattern of variable productivities.

- The hypothesis of cumulative causation vs self-correcting mechanisms.

5.7. SUMMARY.

FOOTNOTES; APPENDIX I,II; SUBAPPENDIX I, II

C H A P T E R V

THE CONSTRUCTION OF A COST BASE LINE FOR PURPOSES OF ANALYSIS OF RELATIVE EFFICIENCY IN HOSPITAL EXPENDITURE.

5.1. INTRODUCTION

We have presented the basic elements for the setting of a PRAP. It now remains to show the mechanisms through which this policy can be constructed in practice. This will be done in this Chapter.

This Chapter also aims to test some of the basic hypotheses about hospital behaviour put forward in the previous pages.

Outline of the Chapter.

The exercise will be undertaken with reference to a set of Spanish General Hospitals, amongst the Spanish 'Health Residences' selected for this purpose. Although they have already been chosen according to certain external characteristics (section 5.3.1.), cluster analysis will be used in order to increase internal subsample homogeneity with regard to those factors that are thought to refer to differential characteristics of the hospitals output (section 5.3.2.).

Once the clusters are derived, according to the resulting classificatory variables, subgroups will be incorporated into a multiple regression analysis in a first estimation of the predicted cost-base line, or 'relevant expenditure', for the analysis of relative efficiency in allocation of hospital expenses (section 5.4.).

The series predicted from the estimation of the regression equation for each different subsample will be analysed with a view to making an appropriate 'relative price' adjustment on geographical bases (5.5.1.).

Once this is made, we will explore potential patterns of variable productivity, testing the hypothesis of the absence of self-correcting mechanisms in relatively high spending hospitals. In the context of system-inefficiency in the allocation of resources (Section 5.6.).

We will describe in the next section, the hospital data available in Spain. After pointing out the characteristics of the present hospital information system, we will indicate some directions for reform.

5.2. DESCRIPTION OF AVAILABLE HOSPITAL DATA

There does not exist in Spain an integrated system of hospital information (HIS from now) designed specifically for purposes of monitoring and analysis of hospital activity. In general, it can be asserted that Spanish Health Authorities have followed an accountancy approach of a French type, which tentatively incorporates some analytical information, but mainly this is for administrative purposes. This has been collected mostly at the aggregate hospital level, and only partly on specialty-departmental bases.

In short, the Spanish HIS is characterized today by the following features:

(i) The fragmentary composition of the statistics available. In this sense, at least six different reports are required for attaining a comprehensive view of hospital activity.(1) For instance, with regard to the study of hospital expenditure, three different statistical sources need to be consulted; similarly for the anlysis of hospital indicators of activity, and for data on hospital manpower.

(ii) The system lacks a goal orientated definition for data collection, (i.e. from viewing sources rather than just usages). Up to the present, too many indicators of doubtful relevance are requested to hospitals from different Authorities (Department of Health, National Institute of Statistics, INSALUD), often without explicit feedbacks for hospital management. For example, up to 71 indicators are requested for the analysis of the hospital "profitability", most of which are also filled for other sources that do not seem well related to the former uses. In this sense, the problem is not to request new indicators, but to rationalise the utilization of the existing ones and reinforce the implementation in practice of some others (as in the case of the Bills dated 25th of April and 2nd of September of 1978 about budgetary and performance indicators, or more recently, those related to the General System of Accountancy or 'Plan General de Contabilidad' - which should have been implemented from 1st January of 1983).

(iii) Gaps in the coordination of the information due to the actual process of decentralization of the Health Care organisation can also be found. For example, the complete time series data on INSALUD hospitals, collected in the Economic and Functional Reports from 1977, stops at 1981, since data referred to hospitals now depending on the Catalan Health Institute are not now included in the series. This problem can become especially important with the generalization of the decentralization process.

Regarding the former picture, an indication on the _bases for the reform_ of the Spanish HIS can be presented next.

(i) In general, there exists the need to provide a goal orientated definition of the hospital statistics (e.g. from the view of hospital management, utilization review, medical audit, health care planning, special research projects, accreditation requirements etc.). In addition, they should supply information at the most disaggregate levels, capable of being used thereafter for several purposes, such as external financial accountancy (e.g. the determination of 'allowable' and 'non-allowable' expenses on a national or a more decentralized funding basis) or internal hospital management (e.g. for cost-effectiveness analysis, profiles of performance comparisons, etc.). They would enable different levels of attribution of costs for programs, resource area units or other cost-evaluation centres, etc.) and, correspondingly, of accountancy in the allocation of responsibilities.

(ii) The definition of the basic unit of the HIS in each of those broad areas considered for analysis, such as the department (e.g. at the administrative level) and the patient (for costing purposes), or the ward inventory per unit of nursing management and speciality costing for evaluation purposes, as in the English case according to recent recommendations.(2) Specialty levels should recognize moreover the need of greater functional homogeneity (e.g. according to the notion of group disease or some other clinical functions), in view, for example, of the high degree of variation within specialties in terms of patients' clinical characteristics or their associated patterns of resource consumption. Some authors (Thomson et al. 1982) recommend the definition of the statistics for costing at the patient level, since they keep afterwards greater possibilities for further aggregation at different levels (such as disease stages, treatments, consultants, wards, client-groups, age and patients' characteristics and medical specialties). This can be done on the basis, for instance, of the implementation of the ICDA (International Classification of Diseases), complemented, when possible, by some medical notes conveniently coded.

(iii) The need to group data and to apportion consequent expenses for the design of responsibility centers of decentralized management units. For this purpose, we need to improve the present classification of hospitals for increasing 'output comparability' and to use the accreditation system for updating over time the levels of the health care provided.

(iv) Finally, some more detailed recommendations refer to:
a) the need to develop a complete computerized network coordinated at the national level for a more regular provision of information (e.g. monthly), supplying hospital profiles of activity, sample comparisons of cost and performance indicators, etc.;
b) the convenience of complementing the existing information with data related to waiting lists, transfers of patients among consultants of different specialties and the levels of readmissions resulting from past discharges;
c) the inclusion of variables related to patients' social class and to the overall utilization of the health care system.

Since we are constrained at present by the existing information, we will use the available hospital statistics for the construction of the basic cost-line of our approach to hospital 'relevant expenditure'. For these purposes, the selection of the sample is presented below.

5.3. THE SELECTION OF THE SAMPLE FOR THE ANALYSIS.

The overall sample of general hospitals initially selected comprises 90 hospitals. Although all belong to the common category of "Health Residences" (and in this sense they can be distinguished from the "Health Cities" - or Regional Hospitals - and from the "National Centres" which provide a more specialized set of care), they differ from one another in several respects, such as:

(i) The size of the hospitals included, which ranges from 50 to 1050 beds;

(ii) the level of provision of care. That is, the sample includes some hospitals which do not provide maternity care, due to their proximity to maternity hospitals;

(iii) the location of hospitals. In fact, the sample contains general hospitals which are able to rely on the high technology of neighbouring hospitals, and also general hospitals which are not in such a position, and which, therefore, are forced to purchase at least some high technology equipment themselves;

(iv) the fact that the sample includes hospitals with more easily identifiable "catchment" populations than others; and finally,

(v) the geographic variation in the public-private hospital mix, this factor should be taken into account when attempts are made to estimate relative health care provisions.

The need to select a more homogeneous sample for purposes of studying the hospital cost functions became obvious, in view of the impossibility of fully adjusting for all the relevant exogenous factors given the set of hospital statistics available. This conclusion is supported, in addition, by three other considerations which prove the complexity of the adjustment required for an accurate comparison. First, relatively small hospitals may cover services with staff under part-time or some other specific agreements with other hospitals in the area; second,general hospitals which are located far from larger ones are less likely to refer complex cases to these hospitals, being forced instead to treat these cases themselves. As a result, such hospitals are likely to provide a wider range of services than general hospitals located close to regional hospitals. Third, there may be a downwards bias in the hospital average length of stay and turnover interval for those general hospitals which provide maternity services themselves, as compared with those which are close to other maternity hospitals and which provide this type of care on a separate basis.

5.3.1. Some initial groupings

An initial exploration of the overall sample was undertaken, looking at some territorial factors, such as:

a) The geographic location of the hospitals. According to this criterion, it became clear that 14 observations could be analysed on the basis of the correspondence between one general hospital-one province.[(3)] A second subsample was obtained as a result of grouping those observations with more than one general hospital in the province. Residually, another more heterogeneous subsample was determined, which mostly included those general hospitals close to regional and/or maternity hospitals; and
b) some geographic variations in the public-private hospital mix.[(4)]

The overall sample and the subsamples described above were used to study the existence of scale economies and the impact on average costs of changes in the utilization variables. This analysis was constrained, however, by the amount of information available and the number of degrees of freedom. Its results are presented in Appendix V. 2.

The search for a consistent cost-line on which to base a revenue allocation policy led us to the exploration of a more homogeneous sample, resulting of the application of the following criteria:

1. Exclusion of those hospitals with less than 300 beds and more than 650. Although these were, to some extent, arbitrary boundaries,

the aim was to capture a similar scale of management and range of organisational problems which could be related to hospital size.

2. To drop those "Health Residences" geographically close to regional hospitals (see the illustrative map presented in the Appendix to this chapter), given the potential distortions of geographic location on standard set of care as described earlier on.

3. To control for the criteria related to the function, extent and level of inpatient care, such as it is established in the Spanish Catalogue of Hospitals (in the way we will explain later).

A second and final sample of 50 observations was selected as a result of the application of these criteria. For these observations we managed to obtain the unpublished 'raw' data referring to each individual hospital from the National Survey of Hospitals (EESRI)(5) compiled by the Spanish Institute of Statistics (INE).

This information includes data about the structural characteristics of the hospitals (buildings, departments, utilization of materials, etc.), existing equipment, clinical specialties provided, staff and resources allocated (per specialties) and the range and category of the medical services available.

Data about expenditure were also taken from the EESRI Survey, although given its degree of departmental aggregation, they had to be complemented with some other sources. These were taken from the Budget requests sent each year to the Department of Health by each individual hospital, and on the expenditure finally settled, through the statistics published in the INSALUD Functional Economic Report.(6)

(A summary of the data used in other alternative hospital groupings - such as the 'Catálogo Nacional de Hospitales' - or for hospital accrediation and planning purposes, can be seen in sub-appendix - 2 to this chapter).

We will use then our second sample of fifty observations in order to estimate the behavioural cost function as the basis for the prospective revenue allocation policy designed earlier on. In order to achieve greater intra-sample homogeneity, separate cost analysis will be undertaken for each subsample derived from the results of cluster analysis.

5.3.2. Resulting subgroups from cluster analysis.

As advanced in the previous Chapter, the object of the cluster exercise here is that of grouping hospitals with similar 'output' characteristics from which similar performance could be expected 'a priori'. According to this objective, the variables selected for adjusting the former comparisons should aim to capture those aspects of hospital activity which represent legitimate differences in hospital output. In other words, they must reflect those factors related to planning policies in operation (e.g. in terms of the scope of services available, status of the hospital, extent of the provision and the characteristics of the population served),(7) and those factors which refer to enviromental or structural aspects which make for legitimate differences in the way in which services are actually provided.

Selection of the variables.

According to the former view, the factors below are thought to account for differences in the character of the 'hospital output' provided. The following variables were selected for this purpose:

(i) 'Age' of the hospital:
- :A dummy variable (DV) equal to one if the hospital had been reformed during the last ten years (i.e. since 1973), and equal to zero otherwise.
- :A dummy variable referring to the Administrator's subjective appreciation on whether the state of conservation of the building was 'good or very good' versus 'just regular or bad'.

(ii) 'Structure' of the hospital
- :Hospital floor space in square meters.
- :The difference, if any, between the number of licensed beds of the institution and those actually utilized.
- :The number of beds in rooms of one or two beds as a proportion of the total number of beds.(8)
- :Ditto for the proportion of rooms with four or more beds.

(iii) 'Status' of the hospital:
- :A dummy variable for teaching status (doctors) and zero otherwise.
- :Ditto for nurses;
- :dummy variable equal to one in case of provision of ambulatory services together with inpatient care in hospital, and zero otherwise;
- :dummy variable equal to one if the hospital is classified as a

District General Hospital according to the Spanish 'Catálogo Nacional de Hospitales' and zero otherwise.

(iv) Locational aspects of the hospital:
:DV to reflect its relative proximity to the most complex set of care provided by a Regional Hospital - or Health City - using as the criterion its location or not in the same province (if so DV=1 and 0 otherwise).
:DV regarding the urban concentration in the province (equal to one if more than one third of the population in the province lived in the place of location of the hospital, and zero otherwise).
:DV equal to one if the Hospital is located in the capital of the province, and zero otherwise.
:DV equal to one if the density of population in the province is greater than the national average, and zero otherwise.

(v) Population characteristics, or nature of the population served:
:Number of population covered per total beds available.
:Family mean income of the province.
:DV equal to one if the hospital is located in an area considered according to some economic indicators, as having a relatively high standard of living and zero, otherwise; (9)

(vi) epidemiological 'environmental' characteristics:
:A dummy variable for perinatal mortality in the province where hospital is located, with value equal to one if figures are above the national average values and zero otherwise.
:Ditto for the incidence of morbidity (as defined in the corresponding survey on this matter (10) on reported cases).
:The child-mortality rate in the province per one thousand inhabitants.
:The total mortality rate per one thousand inhabitants in the province.
(These last two variables were introduced once it became clear that 'morbidity' did not necessarily move in a similar direction to 'mortality');

(vii) 'geographical context' of provision of health care.(11)
:The total number of beds available in the province (including private and other local public sector beds) per thousand inhabitants.
:Number of habitants per total number of physicians in the province.
:Ditto, but in relation to specialist consultants;

(viii) 'nature of the hospital supply-mix' (12)
:Percentage of hospital childbirths over the total number of cases admitted.
:Percentage of patients with surgery over the total number of inpatient admissions.
:Percentage of patients discharged and sent to other hospitals over the total number of cases treated; (13)

(ix) 'scope' of services a) regarding the inputs and the degree of specialization of the hospital medical staff.
:The ratio of physicians over the total number of personnel.
:The ratio of nurses over the total number of personnel.
:The ratio of ancillary workers over the total number of personnel.
:The percentage on total hospital expenditure of payments referring to 'part time' components in the total wage bill.
:The ratio of physicians in higher hierarchical service categories ('Chiefs of Departments', 'Service' or 'Section') to the total number of physicians.
b) Regarding hospital facilites.
:Number of operating theatres.
:Total number of medical specialties.
:Total number of surgical specialties.
:Percentage of medical specialties with the category of 'services' over the total number of medical specialties.
:Ditto, with respect to surgical specialites.
:Number of medical administrative services.
c) Regarding hospital equipment related to 'quality enhancing' services. (15)
:Dummy variable (DV) for the existence of heart surgery.
:Ditto for neurosurgery.
:Ditto for intensive care.
:Ditto for preventive medicine.
:Ditto for rehabilitation services. (14)

The use of cluster analysis to increase subsample homogeneity.

Cluster analysis is performed on the basis of the former variables to partition the observed sample of hospitals into disjoint groups in such a way that the degree of association is low among the memebers belonging to different groups and high among the members of the same group. This analysis aims, finally, to construct an operational classification based on hospital output characteristics.

Methods, procedures, criteria and measures.

There are two basic kinds of clustering techniques: those based on fusive methods, and those based on divise methods. Both start from the calculation of measure of similarity or difference (in other words, a 'closeness' or distance measure) amongst the observations. The fusive (or aglomerative) methods start by thinking of each observation (or case) as a'cluster of one' in order to fuse similar observations into clusters of larger size. They then fuse, at a new stage, similar clusters obtained into yet larger clusters, and so on until only one cluster remains.

By way of contrast, the divise methods start from a cluster of all the observations, to partition succesively this cluster until there are n clusters of one.

Amongst the fusive methods, hierarchical procedures, as opposed to nested ones, are mostly utilized.[15] Differences among them arise from differences in the way of defining the distance or similarity measure either between one ovservation and a group containing several observations, or between the groups of observations as a whole. This hierarchical procedure fuses observations, or groups or observations, which are closer or more similar at each particular stage, according to criteria such as the nearest neighbour or minimum distance.

The sqaured euclidean distance is usually taken as measure of the dissimilarity type, and the Wards' computational method - based on the calculation of the error sum of squares - as the similarity measure, internally compatible with the basic hierarchical procedures. This second method measures the loss of information at any stage of the analysis which results from the grouping of individual units into clusters. The measure used is the total sum of squared deviations of every 'point' from the mean of the cluster to which it belongs. Thereafter, the union of every possible pair of clusters is considered,selecting for fusion those two clusters whose results imply the minimum increase in the error sum of squares.

Prior to the application of the clustering techniques, we standarise our data in order to remove its scale and the range of its absolute variation.[16] On the other hand, the technique of Principal Components allows us to cope with the problem of multicollinearity among the included variables, in order to eliminate any potentially unwanted weighting of some factors. Otherwise, we would need to assume that some of the factors - those affected by multicollinearity - are more important than others for the pruposes of the cluster exercise, since they would be unevenly weighted due to their major influence through the remaining collinear variables.[17]

Some other considerations about the use of cluster analysis.

Some observations need to be made in order to make an accurate use of this technique.

1. Cluster analysis is not a technique free from value judgements. On the contrary, judgement is necessary at several stages. This is mostly due to the fact that:

a) We need to choose the 'shape' of the cluster. The problem here is that one cannot be sure a priori which shape a cluster may have. In two dimensions we could plot the data (as a scatter plot) and decide by inspection. What we have to do here is something more complex since this search for "shapes" has to be done in a multidimensional space.

b) We need to define the measure of similarity. That is, whether the units are required to be similar as a whole on average for each group of variables or whether similarity has to be established for each variable. This choice is involved in the selection of the arithmetical criterion of similarity.

c) We need to establish some mini-max group boundaries to achieve a balance between conflicting demands for internal homogeneity and a manageable number of groups. For this purpose, we do not necessarily know how many groups of clusters there ought to be. One can only proceed experimentally, bringing all of one's prior knowledge to bear on the interpretation of the results.

2. In addition, clustering procedures are argued to be just a useful tool for the analyst looking at the experiences of several cases (i.e. hospitals) simultaneously. The method appears to be less useful for an analysis looking at the reality of each specific case. They provide a convenient method of handling large data sets and of comparing in a systematic way the operation of some decentralized units, whenever the objective of the comparison can be made highly precise.

3. This technique has some advantages over other multi-cross classification techniques, since it does not require the assumptions of multivariate normality, linearity or additivity for the determination of the groups, neither of an explicit dependent variable nor a pre-determined functional form, nor any fixed discrete interval for the classificatory variable.[18]

The sequence of the exercise.

Having recognized these limitations, we will proceed to summarize the stages of the exercise.

We begin with the initial selection of the 240 variables available (cost, structure and quality-related variables) for each of the 50 hospitals included in the sample,[19] according to the initial expectations that hospital size and status are likely to be the primary grouping criteria for increasing hospital output homogeneity.[20] This expectation was reinforced by a glance on the behaviour of the basic statistics and on the findings in other countries studied.

We selected from these 240 variables, those variables to be used for the cluster exercise on the basis of the following criteria:

(i) To proceed from the full set of variables (using the maximum information available) to a more restricted subset. This was selected in pursuing an easier interpretation thereafter of the results.

(ii) To take into account continuous and binary variables. However, since CLUSTAN (21) could not handle a mixed data set of this nature, we had to transform those continuous variables into binary ones. We did this by assigning for each variable values equal to one to all those observations equal or greater than the average, and zero otherwise. Attempts were made to use multidimensional scaling techniques to convert binary variables into pseudo-continuous variables, but they proved impossible. (22)

The use of binary variables is bound to suffer the problem of the impossibility of removing any implicit weighting unvoluntarily introduced. In fact, whereas the former problem could be faced optimally in the case of the continuous variables through the use of standard scores and the technique of Principal Components, this cannot be done for the binary variables.

Given the nature of the difficulties arising from the use of cluster analysis, and those related to its subjective nature in particular, the issue of the validation of the results derived will also need to be considered. The following systems of validation will be used:

a) By showing that the groupings which emerge from the exercise are stable (or approximately so) under different clustering strategies (i.e. using agglomerative and divide procedures), and

b) by appealing to a visual exploration of the derived plots of the units which are in each group and their respective 'minimum spanning tree'. This is, looking at the 'dendogram' and the plot of the two first principal components derived,(23) with regard to their jointly explained cumulative variance (see, for this, the Appendix One to this Chapter).

c) In addition, we will examine the value of the F and t statistics for those variables included. In fact, the F value gives for each variable the ratio of the variance of the subset of cases included in the derived cluster to the variance of the overall sample. Hence, small F ratios will indicate those variables having comparatively low variation within the cluster and being, therefore, good diagnostic indicators. Otherwise, if no clusters appear, the expected value of F will be close to one.

T values are defined as the ratio of the difference between the mean of each continuous variable in the subset of cases included in the derived cluster and the overall mean in the sample for that variable, to the standard deviation of this variable in the sample. Therefore, large t values will indicate variables having cluster means which are substantially different of the population sample mean. Note that the expected value of t with no significance for the variable in the cluster will be zero, since the two defined members of the nominator should be identical.

A good cluster diagnosis variable could then be considered as one that has a small F value and a high t value in one cluster and a large F value and a low t value in the other clusters. The case of variables with t

values of opposite signs in two different clusters will indicate the inclusion in one group of all those observations of the variable above the mean, with respect to that other grouping of all the other units below.

In sum, hierarchical procedures - (nearest neighbourhood and Ward's method) - and product moment correlation and squared euclidean distance as their respectively consistent coefficients, are used in the first approach for checking the validity of the results. Thus, in order to validate the robustness and stability of the clusters obtained with respect to the method used, the RELOCATE and DIVIDE procedures are additionally performed.(24) In both case, the F and t tests are also analysed.

The results of the exercise.

The first 59 continuous variables initially considered were a mix of structural cost-related and quality variables.

As it can be seen in diagram A-1, the first cluster appears to group the first 25 observations with hospital size between 100 and 320 beds. Only 4 violations are recorded with respect to this 'size' diagnosis variable.(25) The second cluster comprises the observation 25 up to 44 (320 - 500 beds) with only two violations.(26) Finally, a third cluster seems to be present, grouping the six largest hospitals (from 500 to 650 beds) and including one observation which refers to a hospital of special characteristics.(27)

In correspondence with these size brackets, the F and t statistics suggest that hospital size appears to be the best clustering diagnosis variable (i.e. number of cases, total existing manpower and total number of members of the medical profession in higher hierarchical positions). It can be observed, on the other hand, that variables such as the average cost per case, per beneficiary, per bed and per patient-day contained within the groups show values highly dispersed. The same is true for those variables related to the average length of stay and the bed occupancy rate.

In graphs A-2 and A-3, we plot the first against the second principal component factors and the minimum spanning tree.

In a second step, and more in line with our initial aim of not basing the cluster formation on the specific way in which hospitals face the provision of inpatient care, we dropped all the cost-related variables from the selected set. 29 continuous variables were then analysed. We proceeded as before, performing identical procedures and operational transformations (correlations, standarisation and principal component analysis). Two clusters seem to appear (see A-4, A-5 and A-6) as a result of this more specific approach, mainly defined by the two following variables: the ratio of inpatient to outpatient cases and that referring to the proportion of higher medical hierarchies in the medical profession over the total number of physicians. The results are reproduced in Table V. 1 below:

TABLE V. 1

	First cluster		Second Cluster	
	F ratio	t ratio	F ratio	t ratio
(i) Proportion of outpatient cases/inpatient cases:	0.304	-0.553	1.130	0.427
(ii) Number of surgical specialties over total	0.153	-0.384	1.369	0.221
(iii) Number of nurses over the total number of hospital staff:	0.365	0.653	0.942	-0.228
(iv) Number of ancillary workers over total hospital staff:	0.541	-0.654	0.990	0.250
(v) Proportion of higher medical hierarchies over total number of physicians:	0.599	-0.550	0.978	0.310

The robustness of the clusters derived were indirectly tested by dropping from the analysis diagnostic variables (i) and (v) and performing again the corresponding agglomerative procedures. When this was done, the former clusters disappeared and no other significant diagnostic variable took their place.[28]

The third stage in our research in hospital groupings was based on the selection of 7 variables considered to be quality proxies. (On page 241 , A-8 plots the two first principal components which account for a 52% of the cluster variance). The results seem to suggest that from the use of this information, a single overall cluster could be considered. This cluster is particularly homogeneous if we drop the observation 1 (which is a rather old hospital and was going to be reformed in 1980, number 9 (one of the few hospitals included in our sample with a local extent of care), number 34 (a hospital operating on the basis of 130 beds instead of 400 beds initially licensed) and number 47 (which is in fact a regional hospital).

It can also be suggested from the visualization of the dendogram, that observation 40 (the hospital in the Canary Islands) and the last five hospitals involve a more complex set of hospital 'quality' output, reducing then the range of hospital size to be considered for increased homogeneity in between 100 and 500 hospital beds.

Finally, in order to arrive at a more manageable number of variables, we selected 9 continuous variables which were thought to be most relevant for our initial purposes of defining homogeneous subsamples for hospital performance comparisons. The two first principal components already account for a 50% of the cluster variance and the first four of them for 75%. The following variables were selected once we checked that they satisfied the classification rules and were normally distributed:[29]

(i) -The percentage of beds in rooms of only one or two beds over the total number of beds available.

(ii) -The number of patient days per ancillary worker.
- " " " inpatient cases per total number of hospital staff.
-The ratio of physicians to nurses.

(iii) -The proportion of beds in intensive care over the total number of beds.
-The number of medical and surgical specialties over the total number of cases treated.
-The number of medical and surgical specialties with the higher medical rank ('services' in our case) over the total number of doctors.
-The proportion of outpatient cases over the number of inpatient cases treated.
-The proportion of Junior physicians over the total number of doctors.

Two clusters appeared with rather well defined diagnosis variables, with the following statistics (see also A-9):

TABLE V.2

	Cluster 1		Cluster 2	
	F ratio	t ratio	F ratio	t ratio
1-Percentage of junior physicians over the total number of doctors:	0.019	-0.852	0.997	0.331
2-Percentage of outpatient cases over the inpatient cases treated in the hospital:	0.376	-0.25	1.226	0.097
3-Ratio of the total number of physicians over the total number of nurses:	0.597	1.12	0.483	-0.435

Variables 1 and 2 can be considered proxy variables for the 'teaching' attribute of some hospitals and for the relative importance of the outpatient work-load on inpatient activity, respectively.

We explored then the possibility that both variables were appropriate diagnosis variables for our classification purposes. In doing this we searched for the robustness of the classifications obtained in order to establish whether the groups emerged from the analysis were stable under different cluster strategies. RELOCATE was first run for this purpose. In fact, RELOCATE is an iterative procedure with a hierarchic fusion method which optimises the error sum of sqaures as the dissimilarity criteria selected. From a random assignment, it starts with a classification of the population of n hospitals into k clusters. During one 'relocation' scan, each object is considered in turn, and its similarities with all k clusters are computed.[30] The population is repeatedly analysed until no objects are relocated during one full scan. The similarity coefficients between all pairs of clusters are computed in each step, the two more similar clusters being subsequently fused.

However, whether this classification provides a global or local optimum cannot be established precisely. (For this purpose, several different starting points should be alternatively attempted. If the resulting classifications are similar, then the probability that the global solution has been achieved is higher).

We confined the search for two final clusters to 10 iterations for each relocation cycle, using standard scores to replace the numeric data, and calculating the error sum of squares as the criterium of dissimilarity type.

The two final clusters derived were defined by:

TABLE V.3

	Cluster 1		Cluster 2	
	F ratio	t ratio	F ratio	t ratio
Teaching status	0.00	-0.852	0.967	0.401
Outpatient/inpatient work-load	0.437	-0.371	1.188	0.175

These results are highly similar to those derived from the use of the hierarchical procedure- Ward's method - after selecting as the similarity coefficient the euclidean distance. (Only two cases failed to hold this correspondence).

The use of binary variables

The relevance of the teaching attribute appeared again from the utilization of the set of binary variables selected. The following dummy variables tried to capture the different aspects of hospital output:

(i) Hospital 'status':
- DV=1 for the provision of outpatient services.
- DV=1 for the case of teaching doctors and training programs.
- DV=1 if teaching nurses.

(ii) 'Extent' of provision of care:
- DV=1 if the extent of provision of care is classified in the National Catalogue of Hospitals at the 'provincial' level.
- DV=1 if the hospital is located close to a regional hospital (providing so, in principle, a less complex set of care).

(iii) 'Quality' aspects:
- DV=1 if the last reform of the building was made in the last ten years.
- DV=1 if the condition of the building was judged to be 'satisfactory' by the hospital administration.

(iv) 'Sophistication' of care:
- DV=1 if microcomputer.
- DV=1 if CAT-scanner.

(v) A 'distinctive' set of supply:
- DV=1 if provision of intensive care.
- DV=1 if Neuro-surgery.
- DV=1 if Preventive medicine.
- DV=1 if Rehabilitation services.

The results (see A-12) seem to support the grouping of observations into two main clusters according to the presence of the Dummy variable for the 'teaching doctors' status. That is:

TABLE V.4

	Percentage of occurrence of the DV per teaching status	Ratio of occurrence of DV-teaching in the cluster over its occurrence in the overall sample.
Cluster 1	16.1	0.35
Cluster 2	94.7	2.06

In fact, all 19 observations grouped in the second cluster had the characteristic of being teaching hospitals. Regarding the 31 cases in the first cluster, only 3 observations appeared to violate this diagnosis variable.

A similar result was also found when other 20 binary variables were selected on the basis of the values of the continuous variables judged as relevant in the previous grouping exercises. Given that, as we explained, the mix of data was not possible in the CLUSTAN version used, those continuous variables were transformed into binary variables. The results are reproduced in Table V.5.

TABLE V.5

	Cluster 1		Cluster 2	
	% frequence of occurrence of the binary variable	% occurrence in the cluster/% occurrence in the overall sample	% frequence of occurrence of the binary variable	% occurrence in the cluster/% occurrence in overall sample
(i) DV per ratio of no, physicians in higher hierarchical positions over total.	57.6	1.15	35.3	0.71
(ii) DV per percentage of outpatient cases treated over no. inpatient cases	12.1	0.51	47.1	1.96
(iii) DV per percentage of junior physicians over no. doctors	27.3	0.62	76.5	1.74

DV =1: ratio > average; 0 otherwise.

In conclusion, the results of the cluster analysis seem to support the view that 'size' and 'teaching status' are the two most relevant 'classificatory' variables for grouping subsamples with a higher internal homogeneity, in terms of hospital output and in the way in which this is approached here by the set of variables selected. However, these results have to be interpreted with care in light of the caveats and observations pointed out before.

In short, the exercise shows that if a wide range of variables are selected - searching for a full information approach, that is, including structural, cost and quality factors - those variables related to hospital size appear as mostly relevant for classification purposes. The number of beds appears in the event to be the most relevant diagnosis variable. Subgroups are formed for 100-320 and 320-650 hospital beds. In addition, a third subsample with six observations (those between 500 and 650 beds) could be also considered.

If a more reduced set of variables is re-selected, (most related to the nature of hospital output, independently of the way in which this is actually provided - dropping consequently the cost-related variables), the 'characteristic' variables related to the hospital teaching status appears as the best diagnosis component of the subgroupings. The relevance of the teaching attribute appears then robust enough given the application of alternative cluster procedures.

The influence of the 'structure' variables was also indirectly tested, once these variables were dropped and only quality-related variables were used. No significant clusters appear in this case. A rather homogeneous sample for the fifty observations selected seemed to be acceptable from the point of view of those specified quality related variables.

Since the object of the exercise is to predict an average cost per case figure, which being appropriately adjusted could be used with sufficient reliability for building our Prospective Revenue Allocation

Policy, hospitals were grouped according to those last two diagnosis variables. For these subsamples of increased homogeneity, we undertook separate cost analysis.

The <u>second part</u> of the exercise consists of performing this separate cost analysis within each subsample considered using multiple regression analysis techniques. We will search in this way for a more accurate prediction of the average cost value to be used as the basis of the reimbursement policy, once we have controlled for those exogenous factors which give rise to legitimate differences in the utilization of the existing service facilities.

5.4. DESCRIPTION OF THE COST EQUATION AND RESULTS OF THE ESTIMATION.

The structure of the exercise earlier on presented, can be summarized in the following diagram:

Cluster analysis:	structural and environmental hospital characteristics and quality proxy variables for the facilities available.			
Multiple regression analysis:	total cost	efficiency in operation	flow of cases	bed occupancy rate
		number of cases	hospital size	average length of stay
		complexity of the cases		

In order to make the aim of the exercise clear, we should remember the following:

(i) We wish to estimate something which can be used in practice for asserting a cost-base line for purposes of revenue allocation and/or a budget review evaluation. We must avoid then any adjustment which incorporates variables which may be unilaterally manipulated by the hospitals. On the other hand, we do not need to go further into the analysis of the hospital production function, since we are interested here mainly in the performance of cost-related behavioural factors.

(ii) According to the previous discussion, the specification of the behavioural cost equation incorporates:

a) The average cost per case as the dependent variable. The output unit is in fact the number of cases. Since "cases" are the counting measure, we need to incorporate the specific features associated with the case-mix characteristics. This is, as a work-load component for a further adjustment, in addition to those external or 'structural' factors considered in the cluster exercise. This we will do by including a specific "specialty case-mix" vector. Total costs are defined as total operating costs (including capital depreciation costs).

b) The right hand side of the equation includes the following explanatory variables:- The size of the hospital. (Other institutional and external environmental variables have been previously taken into account for purposes of cluster analysis). Given, as we will see, the presence of heteroskedastic disturbances, the size variable will also be used as the weight factor for an alternative estimation using the Weighted Least Square option.[31]

-Characteristic variables and those related to the actual utilization of facilities: The specialty case-mix vector and the flow of cases per bed available are included for this purpose. The case-mix variable tries to recuperate some of the information lost by using an aggregate unit of output. It was decided to use the proportion of the cases in each specialty in view of the evidence available suggesting the presence of multicollinearity.[32] The flow of cases per bed - defined as the number of cases treated per number of beds available - will be taken as the most appropriate utilization variable (since it appears to be the less easily manipulable amongst the utilization variables studied).[33]

In short, our specification follows, in general, Feldstein's approach on hospital cost analysis, as presented in "Economic Analysis for Health Service Efficiency" (1967).

5.4.1. Empirical results of the estimation of the cost function.

The results of the exercise are summarized in Table V. 6. The appendix gives the key of the variables used. Figures are in thousand of pesetas.

The initially estimated equation corresponds to the overall sample of 50 hospitals. (1') includes the equation as initially formulated, whereas (2') accounts, in addition, for scale economies in hospital size and throughput variables, and it incorporates two dummy variables collecting the presence of teaching programs and the supply of outpatient services in the hospital.

Regarding the R^2 value, it can be seen that about two thirds of the variance of the dependent variable in (2') is explained by the variation of the explanatory variables. The value of the F statistic did allow for the rejection of the hypothesis of the null coefficient vector at conventional levels of significance. However, the statistic calculated according to the Goldfeld and Quandt's test, did not allow for the rejection of the null hypothesis of homoskedastic disturbances at the 99 per cent interval of confidence - although it did at the 95 per cent. The equation passed the Chow test for structural stability and the test for hypothetical alternative functional forms, as specified in the Ramsey modified RESET test. (See Subappendix one).

From the results in (2') it can be also noted that the estimated value for the constant, throughput and throughput square variables, teaching status and the variables included in the case-mix vector referring to the proportion of Obstetrics and Gynaecology cases over the total, have a *t* ratio statistically significant at the usual levels of significance.

T A B L E V.6

Results of the regression analysis. (Figures in thousand pesetas)

	const	BA	BA^2	FL	FL^2	DVMIR	MIR	DVCE	MIECST	CGECST	OBGIST	PECST	CUICST
(1')	215.40	-0.065	0.000068	-7.702	0.088	—	—	—	-21.35	3.214	85.56	-15.12	-7.210
AC 1 - 50	(4.391)	(-1.08)	(0.946)	(-3.275)	(2.552)	—	—	—	(-0.554)	(0.115)	(2.588)	(-0.387)	(-0.064)
	R^2 = 0.601		$\bar{R}^2$ = 0.511		F(9,40) = 6.681								
(2') AC	231.1	-0.039	0.0000395	-8.542	0.097	11.5	—	-7.69	-10.92	12.04	92.954	-4.652	-50.73
1 - 50	(4.16)	(-0.581)	(0.486)	(-3.256)	(2.512)	(2.297)	—	(-1.31)	(0.255)	(0.386)	(2.527)	(-0.101)	(-0.387)
	R^2 = 0.665		$\bar{R}^2$ = 0.569		F(11,38) = 6.872		Goldfeld and Quandt: F = 3.668 > F(0.05,9,9) = 3.18 Chow's test: F = 0.783 < F(0.01,9,9) = 5.35					SQFVDV = (0.522)	
(3') $AC^{wf=BA}$	182.70*	—	—	-7.558	0.0827	10.164	—	—	34.53	19.65	108.22	12.186	-61.71
1 - 50	(3.267)	—	—	(-2.75)	(1.986)	(2.454)	—	—	(0.799)	(0.625)	(3.188)	(0.285)	(-0.496)
	R^2 = 0.637		$\bar{R}^2$ = 0.567		F(8,41) = 9.007							SQFVDV = (0.116)	
(4') $AC^{wf=BA}$	191.96*	—	—	-7.713	0.0850	11.290	—	-5.541	27.32	21.24	104.76	10.802	-77.61
1 - 50	(3.349)	—	—	(-2.788)	(2.028)	(2.574)	—	(-0.81)	(0.617)	(3.050)	(3.050)	(0.252)	(-0.614)
	R^2 = 0.643		$\bar{R}^2$ = 0.563		F(9,40) = 8.012							SQFVDV = (0.429)	
(5') $AC^{wf=BA}$	72.89*	—	—	-2.123	—	—	-0.091	—	95.47	50.99	115.66	51.29	-71.07
26 TEACHING HOSPITALS	(1.534)	—	—	(-3.729)	—	—	(-0.707)	—	(1.238)	(1.083)	(2.363)	(0.801)	(-0.388)
	R^2 = 0.598		$\bar{R}^2$ = 0.442		F(7,18) = 3.858		> F (0.05;7,18) = 2.58 > F (0.01;9,18) = 3.84						

	const	BA	BA^2	FL	FL^2	DVMIR	MIR	DVCE	MIECST	CGECST	OBGIST	PECST	CUICST
(6') AC	115.37	0.0722	-0.000076	-5.126	0.046	—	—	—	100.25	37.095	117.37	36.175	-95.10
26	(0.945)	(0.435)	(-0.411)	(-0.92)	(0.544)	—	—	—	(1.21)	(0.653)	(2.101)	(0.518)	(-0.44)
TEACHING HOSPITALS	R^2 = 0.588		$\bar{R}^2$ = 432	F(9,16) = 3.710		>F (0.05;9,18) = <F (0.01;9,18) =		2.54 3.78					
(7') $AC^{wf=BA}$	136.1*	—	—	-1.596	—	—	—	-5.179	-81.27	-20.98	106.93	-128.1	136.92
24	(3.23)	—	—	(-4.58)	—	—	—	(-0.965)	(-1.519)	(-0.489)	(2.147)	(-2.278)	(0.823)
NON TEACHING HOSPITALS	R^2 = 0.770		$\bar{R}^2$ =0.670		F(7,16) = 7.645							SQFVDV = (3.747)	
(8') $AC^{wf=BA}$	258.72*	—	—	-8.03	0.094	—	—	-6.159	-90.51	-28.47	67.45	-118.83	56.88
24	(3.692)	—	—	(-2.594)	(2.090)	—	—	(-1.26)	(-1.85)	(-0.73)	(1.38)	(-2.316)	(0.365)
NON TEACHING HOSPITALS	R^2 = 0.822		$\bar{R}^2$ = 0,727		F(8,15) = 8.643							SQFVDV = (2.787)	
(9') AC	200.1	-0.173	0.00023	-4.272	0.034	—	—	-7.885	-71.832	-26.48	94.50	-95.1	246.5
24	(3.17)	(-1.685)	(1.358)	(1.443)	(0.83)	—	—	(-1.41)	(-1.643)	(-0.76)	(2.08)	(-1.91)	(1.69)
NON TEACHING HOSPITALS	R^2 = 0.861		$\bar{R}^2$ = 0.755		F(10.13) = 8.09							SQFVDV = (2.466)	
(11') AC	66.69	-0.051	—	-1.433	—	—	—	—	-17.49	36.06	176.55	5.147	-79.40
	(1.02)	(-0.893)	—	(-3.140)	—	—	—	—	(-0.275)	(0.714)	(1.947)	(0.074)	(-0.371)
1 - 24	R^2 = 0.622		$\bar{R}^2$ = 0.456		F(7,16) = 3.754		> F (0.05;7,16) = < F (0.01; 7,16) =		2.66 4.03				

	const	BA	BA^2	FL	FL^2	DVMIR	MIR	DVCE	MIECST	CGECST	OBGIST	PECST	CUICST
(12')	17.98	-0.045	—	-1.761	—	—	—	—	165.12	86.95	195.67	53.38	272.0
AC	(1.01)	(-1.32)	—	(-3.09)	—	—	—	—	(3.281)	(2.931)	(4.504)	(0.948)	(1.656)
25 - 50	R^2 =	0.735	R^2 =	0.631	F	(7,18) =	7.12						
(13')	26.74	-0.07	—	-2.012	—	—	—	—	168.88	84.91	221.34	91.37	39.9
AC	(1.09)	(-1.29)	—	(-2.83)	—	—	—	—	(2.81)	(2.46)	(3.47)	(0.568)	(1.63)
1 - 43	R^2 =	0.744	R^2 =	0.595	F	(7.12) =	4.89						
(14')	139.72	0.0011	—	-1.807	—	—	—	—	9.998	42.82	44.55	-46-61	-205.49
AC	(1.63)	(0.06)	—	(-4.59)	—	—	—	—	(0.108)	(-0.502)	(0.471)	(-0.497)	(-1.11)
(0/1)>(0/1)	R^2 =	0.531	R^2 =	0.399	F	(7.24) =	3.88						
(15')	127.95	-0.063	—	-2.116	—	—	—	—	-55.12	-5.21	157.74	-111.43	434.44
AC	(2.84)	(-2.82)	—	(-3.842)	—	—	—	—	(-0.82)	(0-0.13)	(2.545)	(-2.124)	(1.973)
(0/1)>(0/1)	R^2 =	0.743	R^2 =	0.593	F	(7.12) =	9.95						
	const **	IBA		IFL		DVMIR		DVCE	IMIECST	ICGECST	IOBGIST	IPECST	ICUICST
(10')	0.0312	-0.534	—	-0.343	—	—	—	0.00155	-0.00072	0.000086	0.00041	-0.00029	-0.13E-8
IAC	(5.44)	(-1.493)	—	(-4.146)	—	—	—	(1.717)	(-3.61)	(-0.304)	(1.147)	(-1.592)	(-0.892)
NON TEACHING HOSPITALS	R^2 =	0.785	R^2 =	0.647	F	(8.15) =	6.95						
(10' bis)	0.0242	-0.188	—	-0.339	—	-0.002	—	0.00195	-0.00038	0.000087	0.00069	-0.000013	-0.56E-9
IAC	(6.767)	(-0.77)	—	(-6.642)	—	(-2.669)	—	(1.724)	(-2.34)	(0.426)	(2.297)	(-0.426)	(0.635)
1 - 50	R^2 =	0.687	R^2 =	0.616	F	(9,40) =	9.732						

Goldfeld and Quandt's test: F = 1.1567

Chow's test: 3.07 > F(0.05;9,32) = 2.20; > F(0.01; 9,32) = 3.05

SQFVDV = Square fitted value of the dependent variable.

In addition, it has to be noticed that, although not statistically significant, the number of beds and beds-square appeared with the expected signs, according to what can be postulated from the hypothesis of economies of size. Similarly, the expected signs also appear for the utilization variable (flow of cases per bed), for the dummy variable for outpatient services - although in this second case was not statistically significant - and for the specialty case-mix vector, with the exception of the percentage of cases in intensive care units which appear with an unexpected negative sign).

Given the impossibility of rejecting the null hypothesis of homoskedastic disturbances at the 99 per cent interval of confidence, we took into account the possiblity of having estimated inefficient Ordinary Least Squares estimators - although still unbiased and consistent. For this purpose, we considered the Weighted Least Squares option, using the hospital size (number of beds) as the weighting factor according to some initial expectations, and dropping consequently this variable from our initial estimation.[34] The results of the estimation are presented in equation (3') and (4'), respectively, in Table V. 6.
According to them we can point out that:

(i) The relatively high value of the intercept, which comes to indicate the influence of the fixed-costs on the average cost per case treated, once we take into account the rest of explanatory variables.

(ii) The negative sign, as expected, for the "flow" variable. In fact, as a result of a marginal increase in the ratio of the number of cases per bed available, a cost-saving of 7558 pesetas (see (3'))could be made, once we account for all the other factors included in the regression. Moreover, the positive sign for the statistically significant square value of the throughput variable seems to indicate the presence of scale diseconomies on the flow of cases per bed, once reached a turning point at a value of 46 cases per bed/year. (This would affect only a small number of non-teaching hospitals which show values with a maximum of 49).

(iii) The teaching status, as conventionally approached by the introduction of a dummy variable, shows an overall positive impact on the average cost per case of 10164 ptas. This represents an extra 14% above the mean average cost per case of the overall sample of hospitals and a 10% higher than the figure which results from the comparison of the actual values of the two separate subsamples; and, finally,

(iv) the low $\underline{t}$ statistics usually derived for the varialbes included in the specialty case-mix vector, with the exception of the proportion of cases in the Obstetrics and Gynaecology Departments over the total number of cases, with a marginal impact on the hospital average cost of about one hundred thousand pesetas.
(Note that (4') is the same equation as (3') but reinserting the dummy variable for outpatient services).

Nevertheless, in using the Weighted Least Squares option it has to be remembered that the weighted regression coefficients become just the OLS coefficients for the regression with the weighted data, and that the constant for the unweighted data becomes a variable which now takes on different values for different observations. Therefore, the R^2 has to be interpreted with caution since, strictly speaking, it includes the explanation achieved by this 'new' variable.

On the basis of the equation (4'), we incorporated the results of the cluster exercise in the regression estimation.

The teaching status of the hospital was considered first. As it can be seen in (5'), we replaced the binary variable for teaching by the actual number of junior physicians in teaching hospitals. Since 'size' still ranged from 208 to 655 beds, we kept the Weighted Least Squares option and, given that no evidence was found of scale economies - see (6') - for this disaggregated subsample, the squared values of the size and utilization variables were dropped from the regression. We considered then the F test of the vector of null coefficients for the included explanatory variables and the Ramsey's modified RESET test for the specified functional form.

We proceeded similarly for the subsample of non-teaching hospitals. The result can be seen in (7'). In this second case, we could not reject the hypothesis of an alternative functional form as given by the square value of the fitted dependent variable, according to the Ramsey's modified RESET test. In fact, the t ratio for this variable appeared to be statistically significant at the usual intervals of confidence (with a value of 3.747).

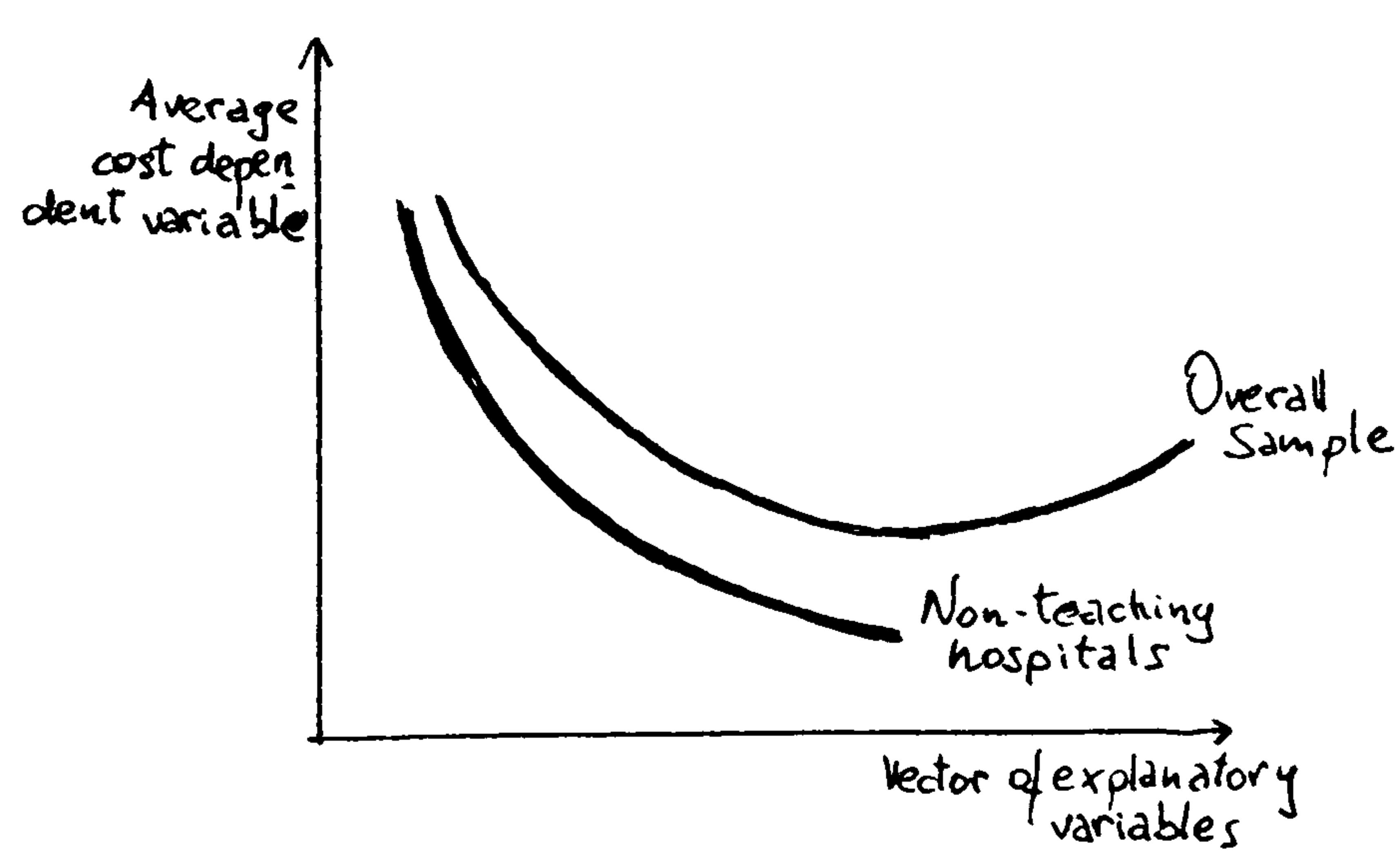
We explored consequently, the introduction first of the squared value of the throughput variable (FL^2) - see (8') - and, alternatively, given that non-teaching hospitals showed a reduced range of size (between 101 and 496 beds), we reinserted the size variable in the OLS option[35] - see (9'). In both cases, although the value of the former t ratio decreased, it still remained significant at the usual levels of significance.

We alternatively specified to these functional forms the simple and double logarithmic functions and an inverse function fitting a rectangular hyperbola to our data. Only this last functional form implied a substantive drop in the t ratio of the squared value of the fitted dependent variable, with values then not statistically significant at the usual levels of significance. This can be seen in (10').

Therefore, according to this result, it appears that, whereas a quadratic function on the utilization variable (adjusted for the specialty case-mix vector) linear in the parameters seems appropriate for the overall sample of hospitals, a rectangular hyperbola which does not meet the axes fits better the data for the specific sample of non-teaching hospitals.[36]

In graphic terms, this involves the following hospital functional forms:

FIGURE V.1



This result has the implication that to pool in a single model the different subsamples, using afterwards a dummy variable to collect the postulated influence of the relevant factor considered, can come to restrict unlegitimately the data to a single functional form and the whole set of parameters included in the model to remain constant for the two subsamples, when not affected by a 'slope' or 'shift' dummy variable.

In light of this result, we re-examined the other estimations, according to the alternatively derived functional form, in order to test for the presence of heteroskedasticity in the overall sample (and hence to check whether the WLS instead of the OLS option should be considered) and whether, in consequence, the Ramsey's modified RESET test before applied, for the specific subsample of non-teaching hospitals was invalid because of the minorated power of the test due to the presence of heteroskedastic disturbances.

From these results we can point out the following:

(i) The set of variables selected for our estimation exercise seems to explain a greater part of the variation of the dependent variable in the case of the non-teaching hospitals than of the teaching hospitals. It can be argued that some relevant factors for explaining the cost-behaviour of the sample of teaching hospitals might have been omitted. Thus, although restricted by the hospital data available, we tried to incorporate other additional variables which did not come to improve substantially the goodness of the fit. This was particularly the case when we used the absolute number of junior physicians,or its percentage in the total number of physicians of the hospital. It has to be noted, however, that, in both cases, the influence of the number of junior physicians on costs appeared with a negative sign. The reason for this could be that this manpower component represents ultimately a form of cheap labour which outweighs the influence of some other factors which can make for a more expensive output, once we have adjusted for their different specialty-case mixes.

(ii) The value of the intercept shows a value 15% greater in the case of teaching hospitals than of non-teaching hospitals. This result is very similar to that also found earlier on.

(iii) The value of the utilization variable shows a similar marginal impact on the average cost dependent variable in both subsamples analysed.

Finally, we took into account the influence of the size variable and of the proportion of the supply of outpatient services over the number of inpatient cases provided in the hospital.

With regard to the hospital size variable, the results of the estimation for the first 29 observations (up to 320 beds) are collected in (11'), whereas for the last 25 cases (up to 650 beds) the results are in (12'). Finally, the last six observations were dropped according to the plot examined in A-3; the results are summarized in (13').

Therefore, with respect to the initial subgroupings, the following results can be pointed out:

(i) The increase in the 'goodness of the fit' from the higher sample homogeneity achieved, as seen from most of the conventional statistics:

	Overall sample	First subsample (100/320 beds)	Second subsample (320/650 beds)	Third subsample (320/500 beds)
R^2 :	0.4969	0.621	0.735	0.744
$\bar{R}^2$:	0.4130	0.456	0.631	0.5951

(ii) The increase in the size of the economies which can be achieved from increasing the case flow variable as hospital size increases. This is particularly the case for the second and third groups.

(iii) The significance at the conventional intervals of confidence of most of the estimated coefficients for the specialty case-mix vector in the second subsample of larger hospitals. The relatively high values before found for OBGIST are repeated here.

In similar terms, our initial regression equation was estimated separately for the two subsamples derived from the previous cluster analysis according to the 'outpatient/inpatient care' attribute. The first subsample includes those observations with the proportion of outpatient cases over the total number of impatient cases greater than the sample average, and the second, those with a value of the proportion less than average.

The results of breaking down the overall sample according to the value of that attribute can be seen in (14') and (15'), respectively. From them, it can be pointed out that:

(i) As could be expected, the reduction in the average cost per case from a marginal increase in the 'flow' variable was greater in the second than in the first subsample with values for the ratio of the 'classificatory' variable greater than the average.

(ii) With regard to the second subsample, the expected negative sign of the coefficient for the proportion of paediatric cases over the total is statistically significant at the conventional intervals of confidence, as it is the large value for the coefficient of the proportion of cases in intensive care units over the total (as could also be expected).

(iii) The evolution of the R^2 and $\bar{R}^2$ adjusted for the degree of freedom was as follows:

	Overall sample	First subsample	Second subsample
R^2	0.4969	0.5307	0.743
$\bar{R}^2$	0.4130	0.394	0.593

On the basis of (4'), and the subgroups analysed in (5') and (10'), we construct the prediction series for further analysis and hypothesis testing.

This is done in the next sections.

5.4.2. The results of the prediction.

Although the significance of the estimated coefficients was not evenly acceptable at the usual intervals of confidence, we opted for the utilization of all the included regressors and not only of those which were statistically significant. [37]
On the base-line given by those estimated parameters we predict a value to which corresponds a reimbursement figure for each individual hospital, according to the characteristics taken into account in the regression. The exercise is repeated for the two main subsamples derived according to the two main classification criteria. The results can be seen in table V.7.

The series will be analysed first in the next section with a view to making appropriate relative price adjustments, before those figures are supplied as the input source for the final simulation exercise. We will test thereafter on this basis the pattern of variable productivity and the hypothesis put forward before of the absence of a self-correcting mechanism for relatively high-spending hospitals and, in case, of a potential cumulative causation effect in a context of generalysed system inefficiency in the allocation of hospital expenditure.

5.5. ANALYSIS OF THE RESULTS OF THE PREDICTION.

In analysing the results of the prediction series we consider first the argument for the need of a geographical relative price adjustment. The purpose of this adjustment is to achieve a similar impact in terms of the purchasing ability amongst the different reimbursed units, in order to provide the same amount of health resources through the financial allocation of any "peseta" no matter where this is spent.

TABLE V.7

ACTUAL AND PREDICTED REVENUE ALLOCATION (1979)

	Total Actual Revenue Allocation-1979	Total Predicted Revenue Allocation-1979 (4')	Total Revenue Predicted-1979 (5') and (10')
1	243065	253026	263176
2	309556	265832	272872
3	354395	342317	347057
4	381943	398239	395571
5	400574	388189	395334
6	354194	381339	387776
7	508114	430688	411583
8	504649	468420	479942
9	348366	423461	410158
10	491140	502861	477634
11	537108	542732	572679
12	606723	588835	618591
13	614578	716551	696549
14	447576	474407	447673
15	647219	637791	594364
16	918061	892566	897988
17	476231	670958	565150
18	553536	612223	567205
19	868760	837481	681309
20	532478	619110	572320
21	629878	588376	556224
22	728641	823968	820426
23	699504	537872	627076
24	523198	539126	545576
25	798617	861282	872159
26	836916	775555	901397
27	831700	795579	806009
28	605636	506067	569522
29	1013270	1047620	1019972
31	1004040	1003660	964322
32	1007470	1080460	1105095
33	1002700	1055510	1002700
34	571041	407825	581014
35	802277	913791	895811
36	961056	934976	868228
37	986430	1247600	1177931
38	1226010	1076800	1183885
39	1247830	1071800	1198423
40	1426940	1075300	1067880
41	1884840	1446250	1406762
42	1022760	1306400	1327957
43	1024460	1129490	1162970
44	1211960	1193380	1060050
45	1291240	1490860	1504741
46	1611390	1650020	1643697
47	1897130	1672840	1782639
48	1516070	1562810	1527676
49	1631060	1571060	1604565
50	1614090	1700800	1684782
51	1271750	1464070	1390515

5.5.1. <u>The need for a geographical relative price adjustment.</u>

For our purposes, the factors which make for the need of an inter-territorial adjustment have to be due, in principle, to labour market conditions exogenous to the internal policies in operation.

Following Bennett (op.cit.), geographical variations in the costs of provision of public services can stem from the following factors:

(i) Enviromental factors (of the type of pollution costs, stress, safety, influence of the climate...);

(ii) technical factors, such as "density" (costs of land assembly, water and sewage services, heating, etc.), scale of economies (e.g. from horizontal and vertical integration) and productivity differences (in the sense analysed in Chapter II);

(iii) 'locational' costs (e.g. cost of land and labour, transport costs, cost of living, etc.); and, finally,

(iv) 'sunk' costs, derived from decisions in the past with implications for resource allocation in the present,such as those derived from the design of plants with excess capacity - failing to appreciate accurately the existing levels of demand - or the effects of technological change on the investment in 'social capital'.

For our purposes, the factors which indicate the need for a geographical adjustment have to be due, in general, to external market conditions exogenous to the internal policies in operation. This excludes those technical factors related to variable productivity and perhaps some other "sunk" costs which are the result of the shortcomings of policies which could have been redressed in practice. Data availability does not allow, however, for a complete adjustment for all those factors (such as in the case of the environmental aspects surrounding the activity) and, for this reason, most of the studies focus only on the adjustment of the so called 'locational' costs.

Thus, the cost of living index is, in practice, the main factor in adjusting for territorial price differences. With respect to the cost of labour inputs (which account almost for two thirds of the hospital health care expenditure), differences in unit labour costs may arise due to the territorial level of unemployment (and, in our context, to the different extent of the possibility of exercising some forms of compatible medical practice) and to the lack of free movement of labour between geographic areas (e.g. given some scaricities and other locational preferences which require incentives to move). With respect to the cost of non-labour inputs, geographical cost differences may arise from locational aspects (e.g. transport costs), differences in the degree of competition and opportunities supplied by external markets for cost savings (e.g. for efficient contract services).

In approaching empirically the relevance of geographical price differences, we will examine first the extent of the variation in the cost of living index in a provincial basis in Spain. According to our own calculations (updating for 1979 the current statistics available), provincial differences in the cost of living ranged from 72.4 up to 123.2, taking the index of Madrid as the bases of 100. The distribution of these differences can be seen in Table V. 8 below:

TABLE V.8

Distribution of the relative index of cost of living (1979)
(Number of provinces)

— <80:	2	91 -95:	10	101 - 105:	10	—>110:	1
81- 85:	5	96 -100:	7	106 - 110:	4		

100 = Madrid;

Source: Own calculations, updating for 1979 the original source INE, 1973.

In sum, 33 of the 50 provinces had values under 100 - half of them less than 90 - wheras the other 14 provinces had a value between 100 and 110.

In addition, as could be expected, there existed a quite high correlation between provinces under value 'one hundred' and those provinces which could be considered rural areas, and viceversa for those provinces with values over one hundred and highly populated urban areas.

In testing, in particular, for the influence of the location factor on hospital costs, we used data made available by one of the larger private health insurance companies in Spain. These data are the only existing - although unpublished - statistics in which costs and utilization variables are referred to some disaggregated treatments. No other information related to patient and diagnostic characteristic variables is available.

The sample refers to 34 private hospitals of Catalunya, which can be clearly identified as located in urban (Barcelona) or non-urban areas, according to the National Catalogue of Spanish Hospitals. We selected then, according to the opinion of two expert physicians on the subject, those observations which could be referred to more similar patient population characteristics, amongst more than 300 treatments. We took for this initial selection, those cases with sufficient degrees of freedom for their analysis, through a cost-regression approach and for their utilization in a cross-section basis.

A priori, we could expect that hospitals in urban areas would be affected by higher costs of living (food, travelling and housing expenses), higher unit labour costs (and land prices) and higher distribution costs. On the other hand, they could be expected to have greater cost-saving opportunities (due, for instance, to the potential supply of more efficient services and to a greater degree of competition for their contracting, although the first aspects were expected to dominate over the second.

DVLH, the dummy variable collecting the location of the hospital, received a value equal to one for those hospitals located in urban areas and zero otherwise.

The results can be seen in Table V. 9.

TABLE V.9

The influence of the hospital location variable on average costs.

Average cost per...	constant term	DV location of the hospital	Length of stay	$\bar{R}^2$	F
Appendicectomy	17211	8380	4161	0.39	(2,26)=8.35
	(1.866)	(1.513)	(3.983)		
Caesarean operation	-23092	35107	9552	0.54	(2,13)=7.80
	(-0.663)	(3.758)	(2.452)		
Phimosis	832	6979	6629	0.35	(2,21)=5.65
	(0.17)	(2.36)	(3.32)		
Hysterectomy	18987	32427	4274	0.37	(2,12)=3.46
	(0.67)	(2.577)	(1.737)		
D & C (post abortion)	4928	4044	5777	0.63	(2,14)=12.01
	(1.84)	(2.177)	(4.836)		
Normal child delivery	9343	18403	3943	0.47	(2,24)=10.84
	(0.80)	(4.65)	(1.695)		
All above cases	8849	18104	8059	0.797	(3,158)=207.3
	(3.40)	(7.95)	(24.9)		

(*) = t-ratios.

The results of the OLS estimation prove, in fact, the significance of the dummy variable for all the cases analysed except for the first of the treatments seen (although the degree of existing multicollinearity of this variable with the average length of stay could be the reason of its minor influence on average costs).

Moreover, the correlation coefficients [18] indicate that the non-urban hospitals seem to be associated with higher LOS. (This can be due to the influence of a 'precaution element' in discharging patients).

The 'location' factor shows then, an inverse association with the average cost per patient day. This higher utilization aspect of the treatment do not seem, however, to be finally reflected on higher average costs per case.

Therefore, we can conclude the necessity of a relative price adjustment for territorial 'price' differences of the type proposed.

First, the analysis of the prediction and simulation exercise will be presented.

5.5.2. The prediction and simulation exercise (1979-1981).

From a detailed analysis of the results summarized before, the following features can be pointed out (see Table V. 10):

(i) The group of relatively high-spending hospitals includes 12 of the 50 observations analysed. Their values exceed, on average, by 17% from the figures predicted according to our adjusted regression estimation. These values range from 6.5% - Ciudad Real ('Nuestra Sra. de Alarcos') - to 34% - Las Palmas de Gran Canaria ('Nuestro Sra. del Pino'). 40% of these observations correspond to teaching hospitals. They show the largest deviations between the actual and predicted expenses, averaging at 20% of the actual expenditure figures. All of them range within the higher spending bracket.

(ii) The application of the relative price adjustment to the hospital sample suggests that some initial conclusions need to be modified. Thus, the observation corresponding to the regional hospital of Valladolid should be dropped from the relatively high spending group. On the contrary, another five hospitals should be incorporated in the group: The hospitals located in non-urban areas and their inclusion in the relatively high spending group depends crucially on the effect of deflating the initial expenditure figures by the level of provincial prices. Therefore, the inclusion of these five hospitals has to be taken with caution. (Note that, in as far as at least two thirds of the hospital expenditure is committed at the national level, the deflation of the full amount comes to exagerate the adjustment). This issue is moreover difficult to overtake given the problem of finding another appropriate index for the required adjustment.

TABLE V.10

Relatively high spending Hospital Group:	1979 % DI	1979 BOR	1979 LOS	1981 % DI	1981 BOR	1981 LOS
"Virgen de Montetoro" -Mahón, Baleares	13.4	66	*	16.5	67.5	*
"Virgen de los Lirios" -Alcoy, Alicante	23.5	*	9.05	12.8	77.6	11.1
"General Moscardó" -Lérida, Cataluña	8.9	*	9.21	7.4	-	-
"Gneral Alvarez Castro" -Gerona, Cataluña	27.4	*	*	18.3	-	-
"Ntra. Sra. del Rosario" -Riaño-Langreo, Oviedo	13.2	50	10.92	-	71.8	11.6
"San Agustín" -Avilés, Oviedo	11.6	54	*	-	57.7	*
"Ntra. Sra de los Alarcos" -Ciudad Real, Castilla	6.3	74	*	13.7	7.4	*
"Perpetuo Socorro" -Albacete	10.7	*	*	-	*	*
"General Sanjurjo" -Valencia	33.6	78.5	11.5	15.3	73.3	11.7
"Ntra. Sra. de los Pinos" -Las Palmas, Canarias	34.0	76.3	*	11.3	69.2	*
"J. Gómez Sabugo" -Gijón, Oviedo.	14.3	75.8	10.6	14.7	71.1	9.8
"Onésimo Redondo" -Valladolid, Castilla	6.4	*	11.1	5.3	79.3	11.13
......after the Relative Price Adjustment, the following hospitals should be incorporated:						
"Obispo Polanco" -Teruel, Aragón	-	*	11.5	-	*	10.3
"Virgen del Camino" -Ponferrada, León	-	*	10.8	-	77.8	10.4
"Virgen de Sonsoles" -Avila, Castilla	-	64.1	11.5	-	64.4	10.5
"Virgen Blanca" -León	-	*	9.4	-	*	9.5
"San Juan de la Cruz" -Ubeda, Jaén, Andalucía	-	57.3	*	-	69.2	

......correspondingly "Onésimo Redondo" - Valladolid, should be dropped from the initial group.

DI = Deviation index between predicted and actual expenditure (ratio in percentage terms)
* = BOR ≥ 80%; LOS ≤ 9 days.
- = Statistic unavailable.

(iii) A geographical inter relation of the former results shows that the group comprises the two hospitals located in the Spanish islands, 10 hospitals located in the capital of rather rural provinces and the other six observations correspond to old industrial areas - (Gijón, Ponferrada, Avilés,....)

The corresponding 'institutional' exploration of the results in higher spending groups reveals that five of the seventeen hospitals could argue some sort of reason for their poor performance, mainly referring to the extent of the health care provided and other related reasons (as in the case of the hospitals in the islands), such as their design in principle for a larger provision and being forced thereafter to close some wards for operational reasons, or being in urgent need of reform. But a more detailed exploration would also show that all hospitals, except one, record values for the average length of stay well above 9 days (up to 11.5) and bed occupancy ratios well below 80% (down to 66%). Comparing these values with those of the relatively "low-spending" hospital sample, we can note that, whereas for the former group the average length of stay is 10.46 days, for the latter group it was just 8.36.

An analysis of some basic work-load indicators is also presented in Table V.11. From these calculations we can clearly see the relevance of the deviations between the actual values and the sample mean for most of those measures in all except three hospitals. This result itself suggests a specific control policy for hospital manpower which can be explored further in the light of the Table presented in Subappendix - Two.

The analysis was completed with the estimation of the initial regressions over the period comprised between 1979 and 1981 and by predicting an individal figure for each of those years. In so doing, we had to assume that the specialty case-mix vector of the hospitals remained constant over the three year period. This is because hospital data on 'specialty' bases are collected in Spain only every four years, and the values corresponding to 1983 are not yet published. Nevertheless, the assumption required for our initial test, although not fully satisfactory, is not entirely inplausible in our context, at least for a short term view. With respect to the other tests,we proceed as indicated in the formulation of the simulation exercise.

TABLE V.11

Analysis of some work-load indicators with regard to the relatively high spending hospital group.

	Actual	Sample mean		Actual	Sample mean
2 -Mahón					
Beds per HPS	4.7	8.3	Beds per hospital physician	2.4	3.3
Patents assisted per HPS	177	269	Patients assisted per physician	92	107
7 -Alcoy (Alicante)					
Patients assisted per HPS	173	269	Beds per total hospital manpower	0.45	0.51
Beds per registrar	4.7	6.1	Patients per total hospital manpower	10	16
Patients assisted per registrar	111	196	Beds per ancillary workers	1.3	1.7
Beds per total no physicians	2.6	3.3	Patients per ancillary workers	29	54
Patients per total physicians	61	107	Patients per clinical auxiliar	49	68
Patients per nurse	47	56			
15-Lérida (Catalunya)					
Beds per HPS	6.7	8.3	Patients per ancillary worker	39	54
Patients per HPS	163	269	Patients per total hospital manpower	13	16
Patients per registrar	142	196			
Beds per physician	2.9	3.3			
Patients per physician	70	107			
19-Gerona (Catalunya)					
Beds per nurse	1.6	1.8	Beds per ancillary worker	1.4	1.7
" " clinical auxiliar	1.8	2.2			
21-Riaño-Langreo (Oviedo)					
Beds per registrar	5.5	6.1	Patients per nurse	51	56
Patients per "	137	196	" " clinical auxiliar	53	68
Beds per physician	3.0	3.3	Beds per ancillary worker	1.5	1.7
Patients per physician	74	107	Patients per ancillary worker	37	54
Beds per total hospital manpower	0.38	0.51	Patients per total hospital manpower	12	16

HPS = Highest Physician status ("Chief"of Department, Service or Clinic)

Source: Own calculation from the information Appendices to the decentralized Budget Requests. (1979)

	Actual	Sample mean		Actual	Sample Mean
23-Avilés (Oviedo)	*	*		*	*
28-Ciudad Real	*	*		*	*
36-Albacete	*	*		*	*
40-Valencia (Valencia)					
Patients per HPS	210	269	Patients per nurse	48	56
" " registrar	180	196	Patients per ancillary worker	46	54
Patients per physician	97	107			
41-Los Palmas (Canarias)					
Beds per HPS	6.7	8.3	Beds per nurse	1.27	1.75
Patients per HPS	230	269	Patients per nurse	44	56
Beds per registrar	3.8	6.1	Beds per clinical auxiliar	1.02	2.15
Patients per registrar	132	196	Patients per clinical auxiliar	35	68
Beds per physician	2.4	3.3	Beds per total auxiliar personnel	0.55	0.93
Patients per physician	84	107	Patients per total hospital manpower	12	16
Beds per total hospital manpower	0.36	0.51			
44-Gijón (Oviedo)					
Beds per registrar	5.8	6.1	Beds per total auxiliar personnel	0.79	0.93
Patients per registrar	152	196	Beds per ancillary worker	1.6	1.72
Patients per physician	89	107	Patients per ancillary worker	43	54
Beds per nurse	1.45	1.75	Beds per total hospital manpower	0.46	0.51
Patients per nurse	38	56			
Beds per clinical auxiliar	1.95	2.15	Patients per total hospital manpower	12	16
Patients per clinical auxiliar	51	68			
47-Valladolid (Castilla)					
Patients per physician	82	107	Beds per total auxiliar personnel	0.72	0.93
Beds per nurse	1.33	1.75	Beds per ancillary worker	1.06	1.72
Patients per nurse	32	56			
Beds per clinical auxiliar	1.66	2.15	Patients per ancillary worker	26	54
Patients per clinical auxiliar	40	68	Patients per total hospital manpower	9	16
Beds per total hospital manpower	0.38	0.51			

5.6. TESTS.

With respect to the hypothesis of self-correcting mechanisms in hospital performance, only 5 of the 12 relatively high spending hospitals proved some type of self-correction over time: that is, their over-expenditure centrally moved towards their predicted values according to our results from the estimation. Four of them followed this approximation with a stable character over the period.

With regard to the whole sample, 15 observations seem to lose over the period their initially close fit to the predicted values estimated from the regression. Amongst the rest, 6 remained as high-spending hospitals over the whole period, and 3 showed a cumulative causation effect, with values progressively moving away from their postulated targets.

However, half of the sample stayed in the inital classification group and a quarter of them kept their expenditure close to the predicted values over the three year period.

Finally, 5 hospitals (3 of them teaching hospitals) showed rather erratic patterns, fluctuating above and below the 5 and 3 per cent margins around the estimated values postulated as the benchmark figures.

The institutional analysis on the utilization variables over the three year period also gives support to the hypothesis of system inefficiency or absence of incentives for improving efficiency over time. In fact, with respect to the 'high spending' group, 6 of 7 observations with low bed occupancy ratios kept these values over the whole period, showing no effort to adjust to the capacity available. Similarly, 8 to 10 hospitals with relatively high average lengths of stay did not show any change in shortening the length of stay. With respect to the other two, only one represented an actual improvement.

With regard to the remaining sample, all except one observation with values for the bed occupancy rate below 80% recorded improvements in their ratios over the period. In similar terms, 7 over 10 hospitals with average lengths of stay greater than 9 days displayed a relative improvement at the end of the period, with substantially shorter length of stays. (The composition of both groups can be seen in table V. 10).

Finally, we supplied the results from the more accurate estimation performed for the 1979 period (Table V.7) as the raw data for the simulation of the PRAP. Having fixed as the relegation criterion the inclusion of hospitals with the 30% highest relative deviations between actual and predicted values (sequentially calculated over the three years in which the designed policy we postulated to be in operation), 20 hospitals were submitted to the incentive reimbursement policy. Only 4 amongst those observations showed,over the period, actual values moving close to those calculated according to the incentive reimbursement policy (although one of the hospitals did not seem to follow the adjustment on a permanent basis).

This result appears to be independent of the weighting factor established for defining the speed of the adjustment through which the revenue formula proposed was required to move towards the predicted targets. (In fact, a loop in the program enabled us to supply values for the weighting factor between 0.1 and 1).

In conclusion, the following points can be summarized:

(i) The presence of system inefficiency seems to be proved, particularly in relation to the results of the estimation of the hospital behavioural cost functions over time and in the hospitals performance in terms of the pattern followed by some relevant utilization variables over the period studied.

(2) The absence of self-correcting mechanisms is proved specifically with respect to the figures predicted on hospital expenditure, according to the estimated values for each corresponding year over the period considered,

and in view of the results of the simulation exercise.

(3) However, the hypothesis of cumulative causation seems rather doubtful since it is supported by only 3 observations amongst the 20 hospitals analysed.

SUMMARY OF THE CHAPTER.

In this fifth Chapter we have examined the actual construction of the proposed cost-base line for the analysis of the relative efficiency of the hospital expenditure in Spain.

According to the hospital data available today in Spain, we have selected an initially homogeneous sample for purposes of analysis, which has been further refined according to the results of the cluster exercise. From the application of the corresponding clustering techniques, the teaching characteristic of the hospitals has proved its major relevance amongst the derived 'classification' or 'diagnosis' variables.

On the basis of this latter result, we have run separate cost regression analysis for the overall sample and each of the derived subsamples. The residuals between the actual and fitted values have been then explored in order to assert some initial inferences on the relative efficiency of the hospital expenditure. This analysis has shown, moreover, that the assumption of an identical functional form amongst hospitals with different status and, likely, of an identical behavioural model, needs to be tested by the researcher. In fact, the results of the estimation of the overall sample and each of the subsamples considered shows, in our case, that those assumptions would impose illegitimate restrictions.

After adjusting the results of the estimation in view of the need to account for geographical price differences on institutional expenditure, we have examined the deviations between actual and predicted figures (according to the 'PRAP') in an institutional and provincial basis. Having performed some initial tests for the period 1979-81 under the unsatisfactory, but necessary assumption at this stage, that the specialty case-mix vector of the hospitals remained constant over the period, we have supplied the results of the more accurate estimation for the exercise 1979 as the input data for the simulation exercise, such as specified in Chapter Four, for implementing an incentive reimbursement policy.

In both cases, the results have been analysed in view of testing the hypotheses of system inefficiency and the absence of self-correcting mechanisms in the pattern followed by sequential hospital expenses. Both hypotheses can be widely accepted in light of the results provided.

However, those results would not seem to support the tentative hypothesis of a cumulative causation effect in absence of a 'PRAP' of the characteristics described in the thesis.

FOOTNOTES:-

(1) They are: "Información Económico-Funcional de las Instituciones Sanitarias" (INP), INSALUD; "Encuesta de Establecimientos Sanitarios en Régimen de Internado", INE; "Encuesta de Morbilidad Hospitalaria" INE; "Memorias provinciales del INSALUD" (INP); "Anexos estadísticos a la elaboración de los Anteproyectos Presupuesarios de las Instituciones cerradas de la Seguridad Social" (unpublished), and "Anuario Estadístico de España", INE.

(2) Körner Report - Working Group F on Health Services Financial Information (1983).

(3) For this subsample, a greater amount of information was available, in addition to those statistical data published on a provincial basis.

(4) We explored this question through some regression estimations. The following results are summarized here:

(i) The high dispersion in the variable collecting the interprovincial range of variation of the beds supplied per population covered (coef. of variation = 50.20) diminishes once we analyse public provision in regional basis. However, Canarias and Catalunya show still coefficients relatively low and Aragón and Extremadura relatively high (approximately, 60% over the average).

(ii) According to the results of the estimation, a 25% of the variation in the total number of beds per capita could be explained by variations in the crude level of provincial income.

(ii) Once the total number of beds (public and private) per capita are added, the utilization of the mean value produces the following relations at the provincial level:
a) provinces with lower than average public beds per capita ratio, and greater than the average ratio of private beds per capita. These provinces are mostly those with the highest per capita income.
b) Provinces with a higher than average ratio of public beds per capita and lower than average private beds per capita ratio. These provinces can be mostly identified as those with a per capita income less than the average.
c) Finally, those provinces with low values for both ratios, or provinces historically less well endowed in terms of inpatient care supply.

Finally, it could be noticed that almost all the provinces with only one public hospital covering the total provincial population showed a value of the bed per capita ratio higher than the average. This may suggest some problems on the hospital minimum size with respect to the provision of health care in territorial bases.

(5) We refer here to the "Encuesta de Establecimientos Sanitarios en Régimen de Internado", published in Spain every four years.

(6) This was due to the fact that, for example, the first Survey, but not the third, includes those expenditure items on disaggregated bases for clinical and laboratory inputs; the second, and not the third includes the specific components of the manpower wage bill; the third and not the first, presents the figures in a more homogeneous or standard manner.

(7) In order to validate a satisfactory adjustment for those exogenous aspects related to the provision of services, we need to assume that there are not unwanted effects between the planned and the available scope of services. This monothetic relation between actual and potential levels should be in fact proved rather than just assumed.

(8) Alternatively, it was introduced a Dummy variable accounting for whether or not the hospital had a complete refrigeration system (as the structural quality-proxy variable).

(9) Data on provincial population under 15 and over 65 years, even when they were regarded as particularly relevant for the purposes of our analysis, are not available. Data referring to the age of the patients can be taken in Spain only partly from the hospital morbidity statistics.

(10) From "Encuesta sobre Morbilidad Hospitalaria". This Report is published every four years. Since the last data available for those indicators are referred to 1976, we will need to assume that their relative importance for cross-sectional studies has not varied very much with respect to 1979.

(11) As in (vi), all these data refer to 1976, last year available.

(12) A more detailed analysis can be performed here with the utilization of data on the number of beds, patients, patient-days and outpatient visits per specialties. This is, in relation to Internal Medicine and Medical Specialties;

General Surgery and Surgical Specialties; Obstetrics; Gynaecology; Psychiatry; Paediatrics (New-borns' units, General Medicine and Surgery); Intensive care; Rehabilitation; Tuberculosis and 'others'. Due to the ultimate reference of most of these figures to the actual use of the existing facilities, it was thought that their utilization was not convenient for our purposes.

(13) Ideally, we should measure all these values as average figures over a certain number of years, in order to get a more accurate structural measure independent of any fortuitous practice. This was not possible, however, in our case due to lack of data.

(14) The variables selected in (ix) were chosen with the help of two physicians. Although it was recommended, the inclusion of the extent of the services within 24 hours coverage in the Casualties Departments - as considered in some systems of Hospital Accrediation - was not possible as no published information of this nature was available.

Other variables can be added to achieve the same objective. For instance, with regard to whether an "organ bank" exists, if nuclear medicine, plastic surgery, level of radium-therapy, etc. is provided. As a whole, they were thought not to add very much to what it was already collected by the included variables.

(15) Hierarchical cluster techniques have the general disadvantages of containing no provision for the reallocation of entities which may have been poorly classified at an early stage of the analysis.

(16) No specific treatment was, however, applied to tackle the problem of the bias towards finding spherical clusters. This is a common problem to most of the clustering techniques. No definitive procedure exists to deal with the problem of imposing, without empirical grounds, a particular shape on the structure of data.

(17) An additional problem could arise from the potential bias introduced by the group size. This bias appears when two groups fused at a given stage are very different with respect to size. The centroid of the small group under a cluster spherical shape can result that is very close to that of the large group, and even perhaps, remaining out of the group. However, and contrary to what happened before, median cluster analysis can be specifically used for dealing with the problem, since this method makes the clustering strategy dependent on the group size.

(18) See Klastorin, Watts and Trivedi: "A Study of the Classification of Hospitals for Prospective Reimbursement", prepared for by Health Care Financing Administrations Report no. 10, DHEW, Washington. Seattle U.S. Department of Commerce. National Technical Information Service. Jan. 1978.

(19) For most of the cost and functional variables, figures from 1977 to 1981 were available. However, for "structural" and "quality" related variables (basic factors for our purposes of cluster analysis) and for the specialty case-mix vector (number of cases, of patient days and number of beds devoted) only figures for 1979 were available. With regard to this data already available, we selected the measure of similarity or distance which suited most of the data. HIERARCHY performed the corresponding fusions according to the prior specified method (median, nearest neighbourhood, group average, Ward's method etc.). Finally, procedure RESULT printed the overall ouput in the form of classification arrays and PLINK produced the corresponding plots.

(20) In fact, this is the basis of most of the existing hospital classifications. For instance, in Great Britain, Acute hospitals are split according to the number of beds: less than 50 beds, between 51 and 100, 101 and 300 and over 300 beds. The teaching status attribute has also been commonly considered relevant in the literature,being incorporated in the estimation exercises of hospital cost functions. However, it has to be noted that when it is really the case that two groups of hospitals have different 'outputs', it is clearly unsatisfactory to pool the two groups and account for the apparent difference between them by inserting a dummy variable in the regression analysis.
This point is expounded upon later in the text.

(21) The third edition of David Wishart's version of CLUSTAN has been used, mainly through the conversational command generator CLUSCO. Procedure FILE was performed for the standarisation of the raw data, product moment correlation and principal component analysis. The similarity and distance matrix was constructed by CORREL. A maximum of three (i.e. size brackets) and a minimum of two groups (i.e. regarding the existence of a certain binary attribute) were requested in such a way that they could be used thereafter for further regression analysis (this is, having to keep hence some minimum degree of freedom).

(22) The MDS (X) package was not, at that moment, in operation in York.

(23) A spanning tree results from joining the n points (individuals) in a p dimensional space (numerical variables) by straight lines such that:
(i) none of the lines form a closed loop;
(ii) each point is visited by at least one line; and
(iii) the resulting collection of lines is connected
(that is, there is a path which possibly passes through a number of intermediate points, from one point to any other point).
The 'minimum spanning tree' is that in which the sum of the length of the lines which comprises it is minimized. This provides some evidence about the nature of the many dimensional scatter: Two individuals who appear close together in the two dimensional representation of a principal component analysis plot, may, in fact, be some distance from each other. This will be indicated by the fact that the minimum spanning tree does not link them directly.
A 'dendogram' is the representation of a hierarchical cluster in which the vertical scale of the diagram shows decreasing levels of similarity (if this is the measure selected) as one moves from downwards to upwards. The coefficient gives the similarity levels, say, to which each cluster fuses. (Cooper and Weekes, 1983)

(24) DIVIDE is a procedure for association analysis, which can be used only for binary data. It consists of the division of a set of data, depending on whether the units possess some specified attributes. On the other hand, RELOCATE proceeds from an initial classification from which it reconsiders each identity in turn for reassignment to another cluster, with this taking place if it causes an increase (or a decrease in the case of minimization) in the criterion measure selected in advance. The procedure continues until no further relocation of a single entity causes any improvement.

(25) They are (i) the case 33 - which is an hospital with 400 licenced beds but with only 130 beds available (this has been the case for some time), (ii) the observation 34 - which is a 403 hospital beds not yet in full operation; (iii) the observations 27 and 28 (with no clear explanation).

(26) These are, (i) hospital no. 16 - which is grouped in the National Catalogue within a different category to that of the "provincial level", given its extent of care - and (ii) no. 19 - without apparent reasons.

(27) This is the hospital located in the Canary Islands, which, for obvious reasons, must supply a more self-contained set of inpatient care.

(28) The direct validation of this test consists in reality of repeating the analysis, after deleting a small number of variables. If this does not greatly alter the clusters found, those clusters can in fact be thought as "real" and not merely artifacts of the particular technique used.

(29) We ran for this objective the iterative package (1980-version) of STATPACK, and we used the histograms produced for visualizing the distributions. The assumption that the elements of the population were distributed by a multivariate normal density function is required for our parametric tests of statistical validity (F and t ratios).

(30) For example, suppose that the criterion of similarity between X and any other cluster Q is given by S (Q,X). If S(Q,X) exceeds, say, S(P,X), the method moves X from cluster P to cluster Q. The centroids of clusters P and Q are recomputed to account for this change at the time that the switch occurs. The relocation phase is repeated to obtain a local optimum for (k-1) clusters and the fusion phase re-starts to yield k-2 clusters.

(31) In fact, for the initial equation in which total costs appear as the dependent variable and just hospital size an flow of cases per bed as the explanatory variables, the application of the Goldfield and Quandt test gave an F calculated value of 3.160. Consequently, at the 95% interval of confidence we could reject the null hypothesis of homoskedasticity (3.160 > F(0.05;16,16) = 2.34).

The application of the test for the average cost per case value for the same simple equation gave an F value inferior to the critical one at the conventional levels of significance. This was, 1.2932< F(0.01;17,17)=3.28 and < F(0.05; 17,17)=2.30. Therefore, we could not reject for this modified equation the null hypothesis of homoskedastic disturbances.

(32) From a simple regression analysis point of view, the case-mix vector adds 23% to the variance of the dependent variable already explained by the rest of variables, as seen from the improvement in the R^2 statistics for the goodness of the fit.

The utilization of the percentage figures in the total number of cases instead of the absolute ones for the specialty case-mix vector was due to the fact that there exists empirical evidence about the collinearities of "size" and other institutional and characteristic variables with respect to the specialty case-mix vector (absolute figures). For instance, with regard to size, as measured by the number of beds available, these were the correlation coefficients:

BA-MIEC= 0.694	BA-PENEC = 0.458
-CGEC= 0.730	-PEMEC = 0.476
-OBSC= 0.650	-CUIC = 0.564
-GINC= 0.625	-OTRC = 0.259

However, in using these variables in percentage terms, the correlation coefficients were:

BA-MIECST = -0.148	BA-CGECST = -0.072	BA-OBGIST = 0.027
BA-OTRCST = 0.104	BA-CUICST = 0.322	

(33) A conveniently modified 'bed-occupancy rate' benchmark could be also considered for these purposes. This modification could consist of the definition of a minimum occupational target - or some average one - from which the number of cases corresponding to each hospital size could be derived. This figure would be used then to adjust the average cost amount for purposes of performance evaluation. This system has, obviously, to allow for a controlled reduction in the number of beds according to a planning policy which could enable hospitals to face thereafter, in this manner, recurrent problems of overspending.

An approach of this type has been used in the New York formula-based Prospective Reimbursement System (1970). In this way, the New York Reimbursement formula set up certain lower limits for the bed occupancy rate variable (i.e. 60% for Maternity, 70% for Paediatrics and 80% for Medical and Surgical departments). Hospitals above these limits were reimbursed according to their actual number of patient-days (since this was, in that case, the reimbursing rate used). Hospitals below those percentages were reimbursed according to the number of days corresponding to that occupancy rate at the limit.

(34) In fact, after ordering the sample according to hospital size, as measured by the number of beds, we estimated a subsample of smaller hospital size (postulated as having lower variance) and a sample with larger hospital size(and higher variance group), leaving 20% of the observations between both subgroups our of the analysis. We then performed the Goldfeld and Quandt test in the way specified in Subappendix one.

(35) Given the small number of observations in each specific subsample and the total number of variables included in the regression, we could not test specifically for heteroskedasticity within each subsample, because of the lack of the appropriate degrees of freedom.

(36) This effect was partly visualized by drawing the bed size and the throughput variables, alternatively, against the dependent variable in a two dimensional space. (See Graphs Appendix I).

(37) In this sense, Pindyck and Rubinfeld (1976) - page 161 - point out that this is a point of choice. "A single equation regression model can have significant t statistics and a high R^2 and still forecast very badly. (...) Good forecasts, on the other hand, may come from regression models which have relatively low R^2's and only one or more significant regression coefficients. This may happen because there is little variation in the dependent variable, so that although it is not being explained well by the model, it is easy to forecast". See also in this subject Cox and Snell (1974), 'The Choice of Variables in Observational Studies', Applied Statistics. no. 23, no. 1 page 51.

(38) These were some of the correlation coefficients calculated:

(i) Appendicectomy: DVLH- LOS = -0.568; DVLH- AC =-0.132;DVLH- ACPD = 0.467.

(ii) Caesarean operation: DVLH- LOS = -0.350; DVLH- AC =0.597; DVLH- ACPD = 0.731.

(iii) Phimosis: DVLH- LOS = -0.574; DVLH- AC= 0.098; DVLH-ACPD = 0.483.

(iv) Hysterectomy: DVLH- LOS = -0.498; DVLH- AC = 0.454; DVLH- ACPD = 0.734.

(v) D and C (post abortion) DVLH- LOS = -0.293; DVLH-AC=0.129; DVLH- ACPD = 0.551.

(vi) Normal child delivery: DVLH-LOS = -0.423; DVLH- AC=0.642; DVLH- ACPD = 0.748.

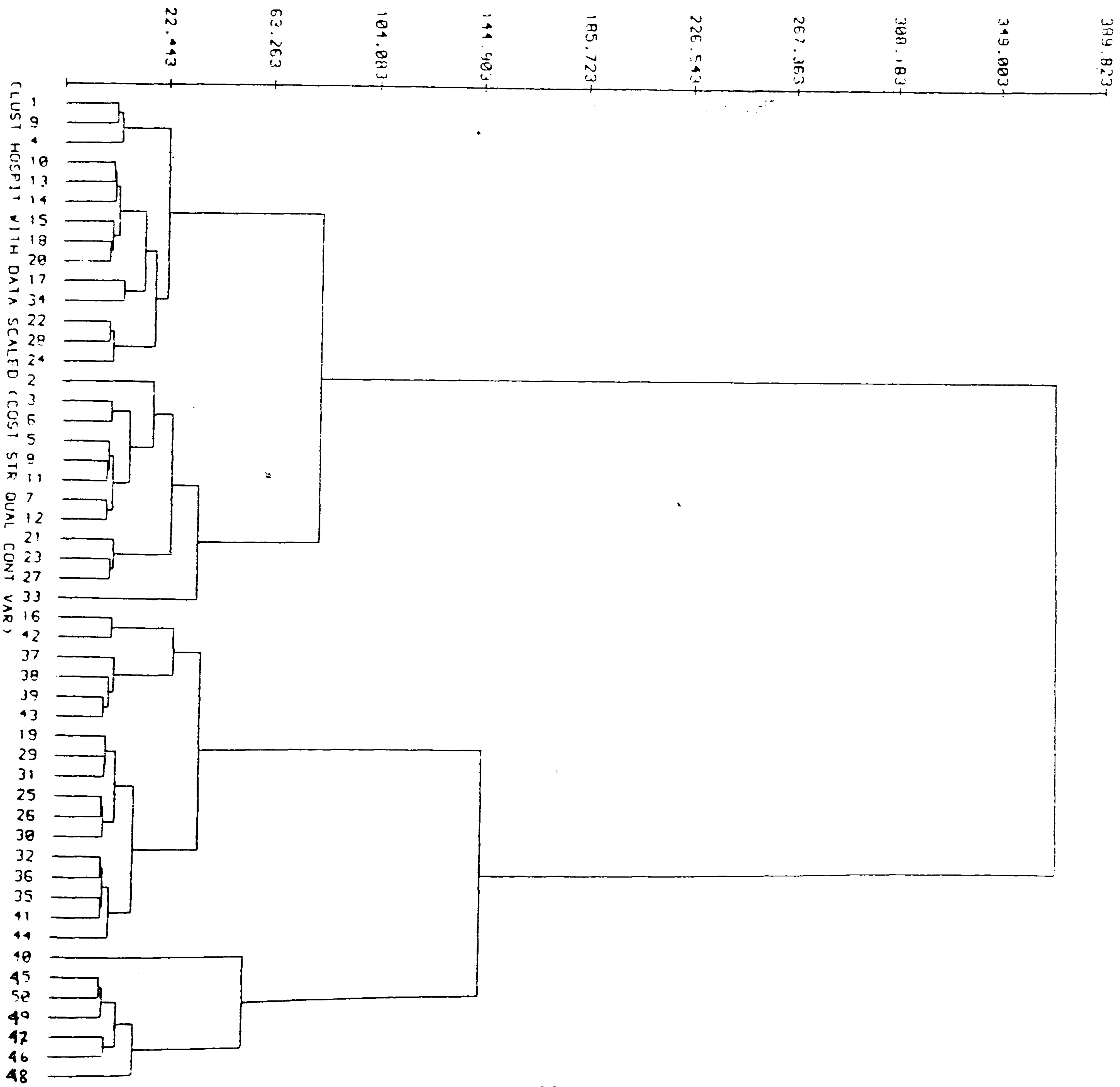
(vii) All the cases: DVLH- LOS = -0.269; DVLH-AC = 0.046; DVLH- ACPD = 0.508.

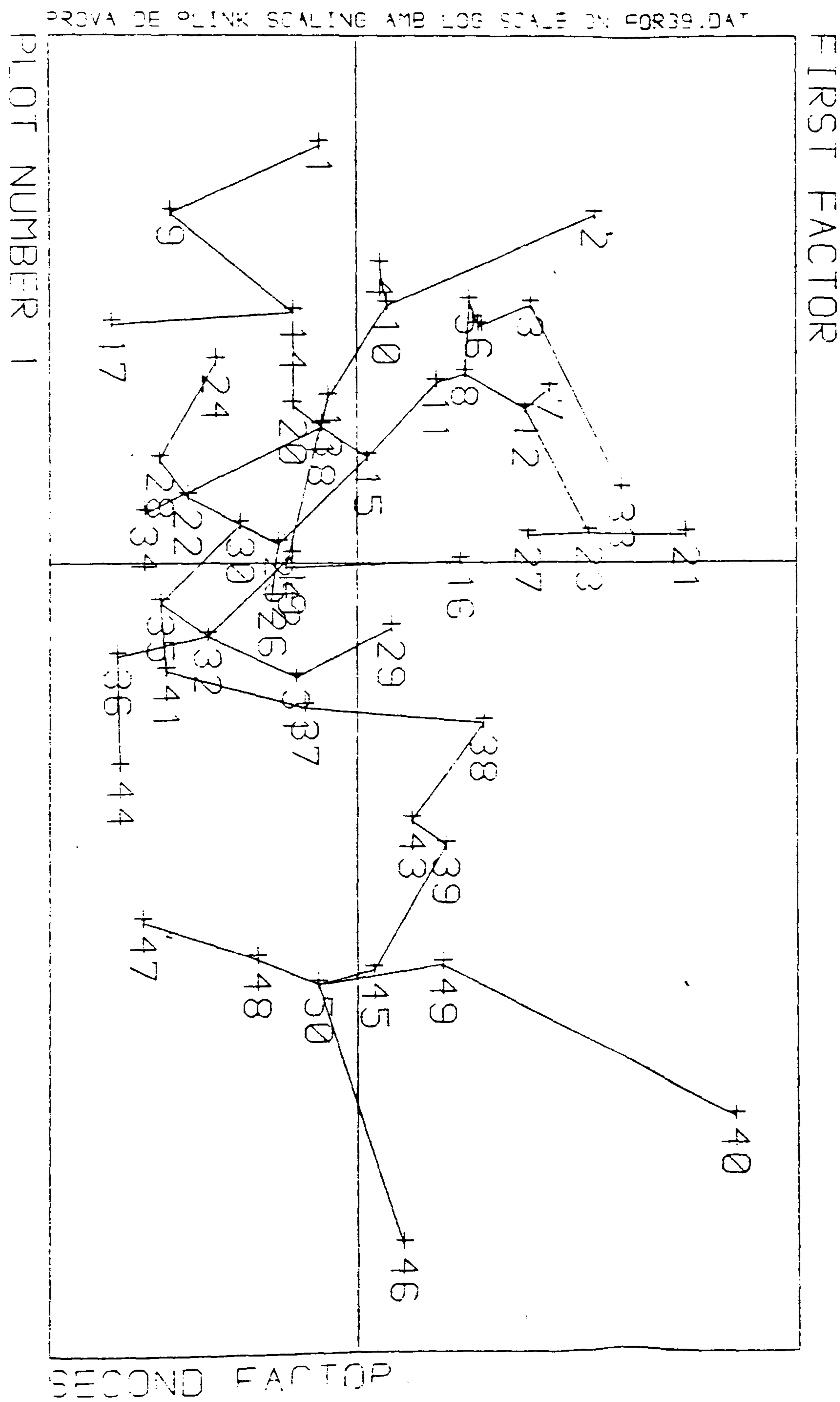
A P P E N D I C E S

A dendogram is the representation of a hierarchical cluster in which the vertical scale of the diagram shows decreasing levels of similarity (if this is the measure selected) as one moves from downwards to upwards.

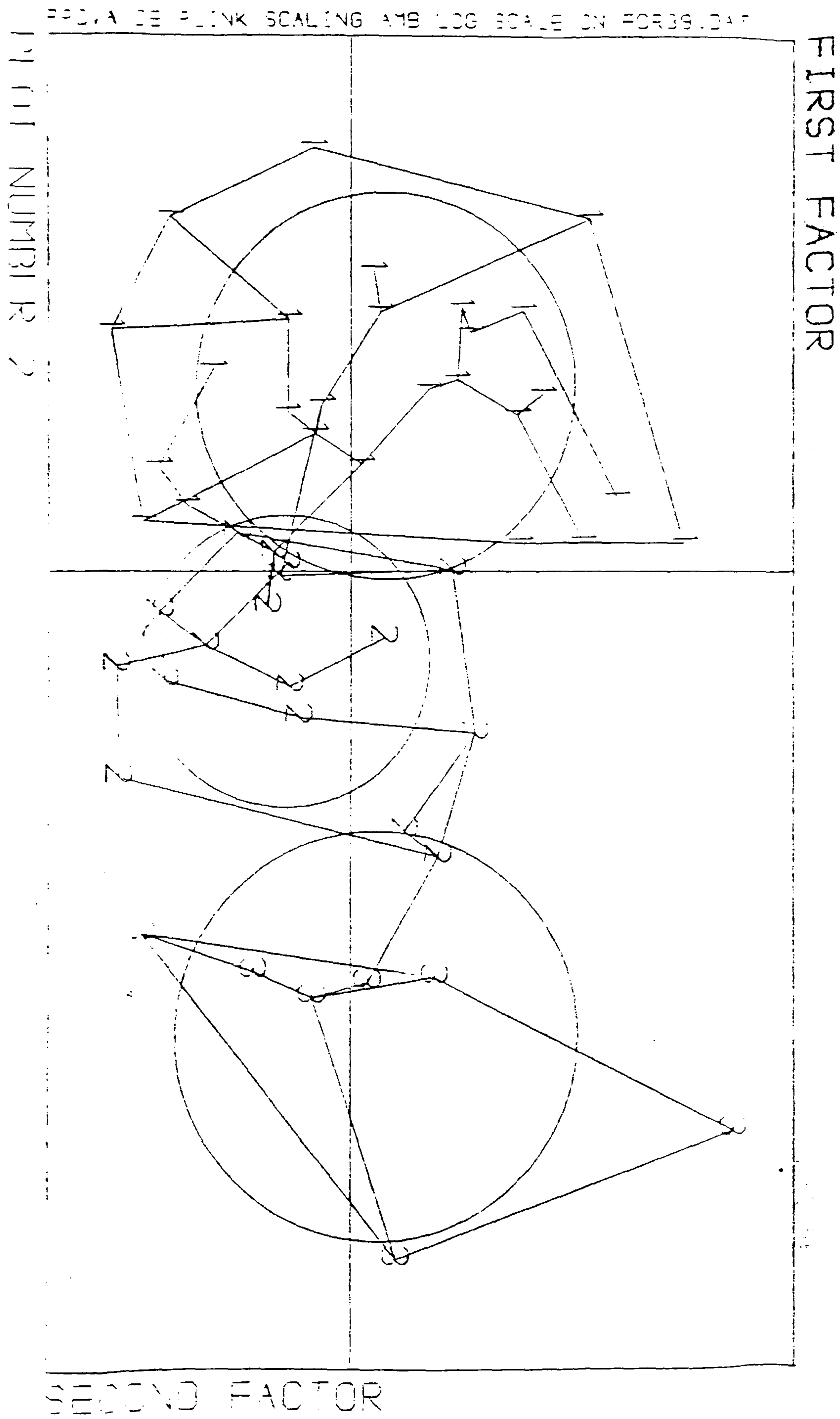
The dendogram below shows the result of clustering 59 continuous variables, refering structural, cost-related and quality proxies of hospital units.

Note that "size" proves to be relevant.(Numbers below the dendogram involve the ordering of the hospitals according to number of beds). The first cluster groups mostly hospitals between 100 and 320 beds. The second one comprises hospitals between 320 and 500. A third cluster may be also pointed out, grouping the six largest hospitals.



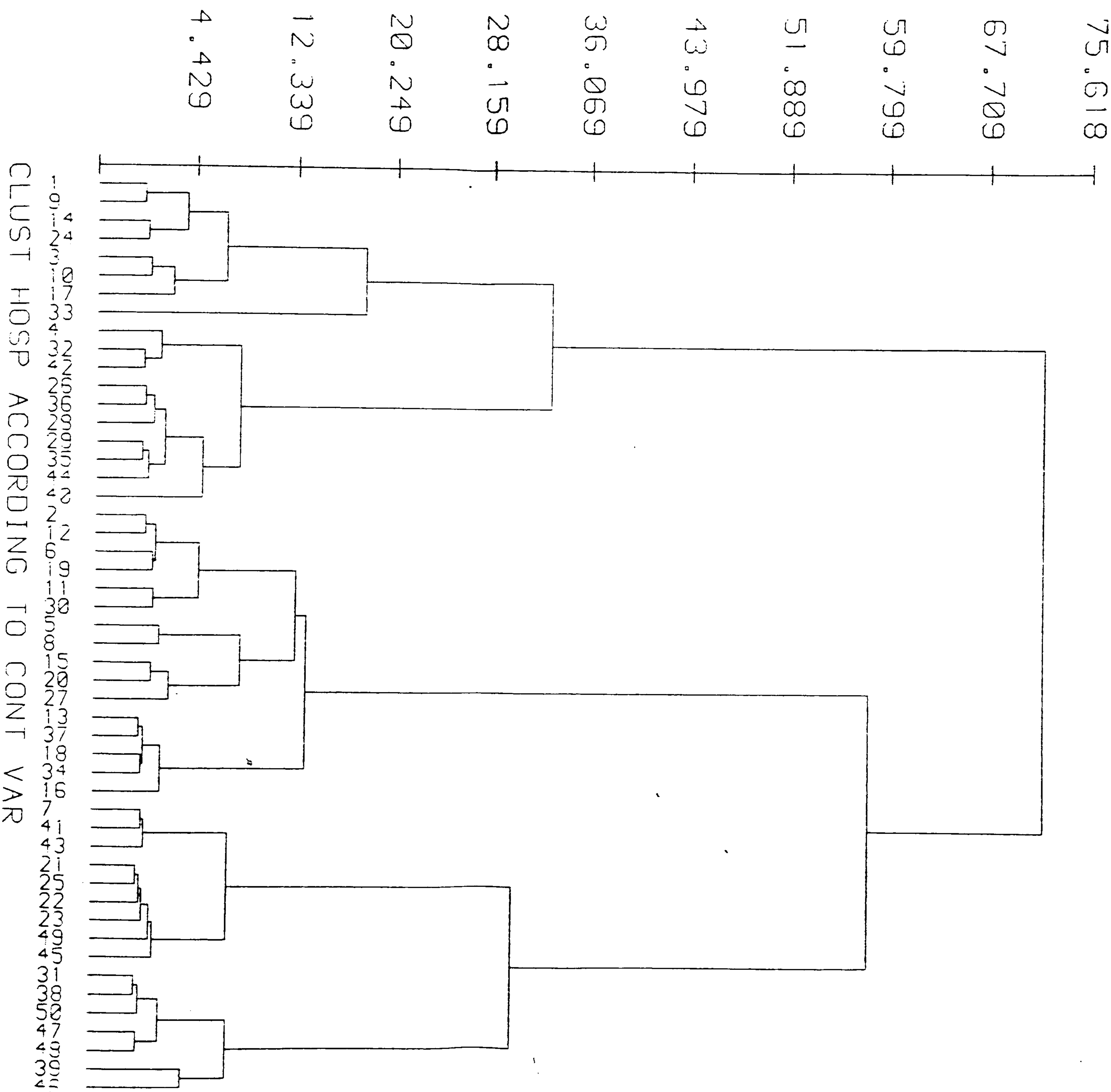


<u>Plot of the first against the second principle component</u> factors as derived from the cluster exercise pointed out in A 1.
The spanning tree is the result from joining the <u>n</u> points in a <u>p</u> dimensional space by straight lines (see footnote 23)



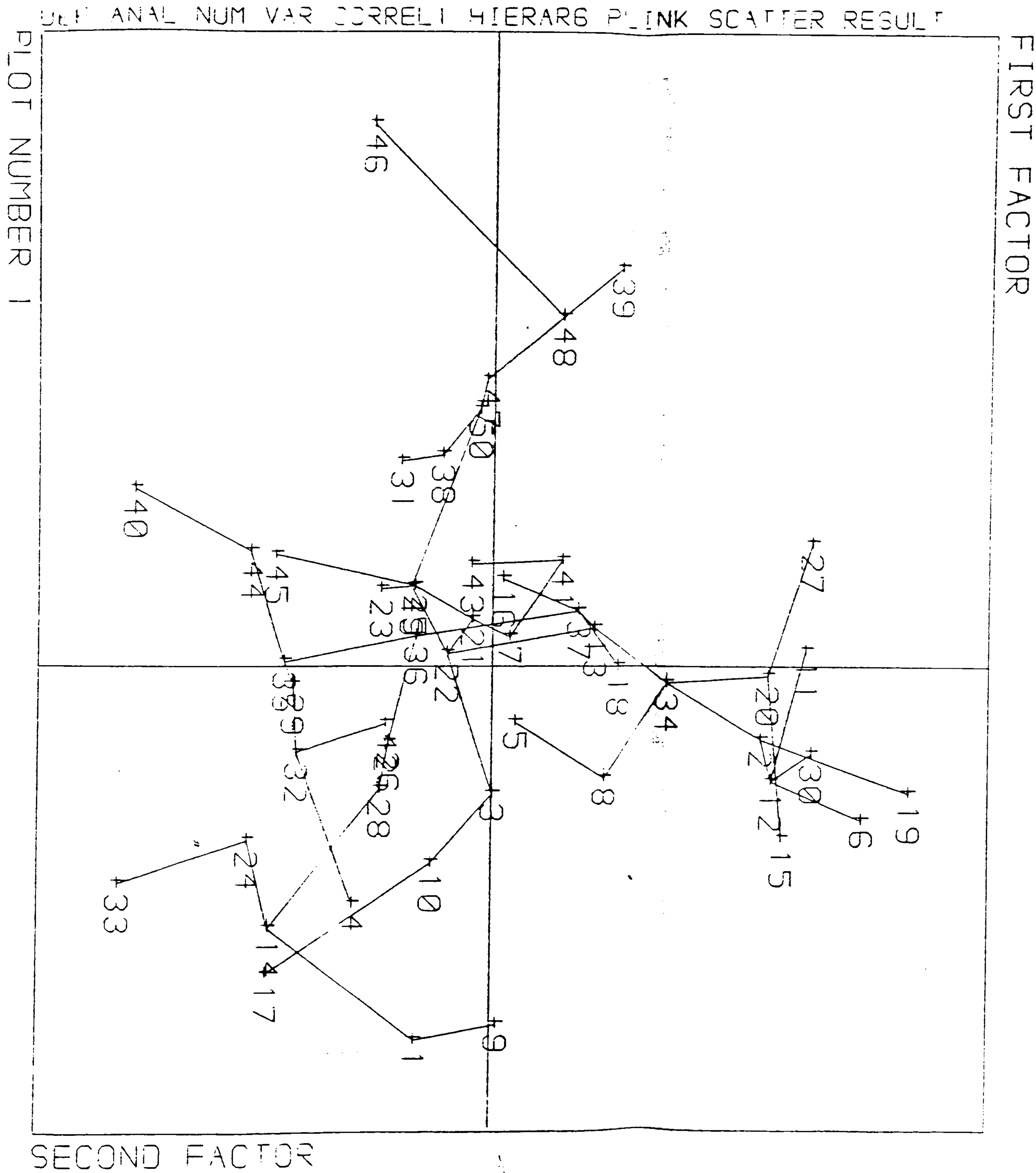
Minimum Spanning Tree derived from the clusters performed in A 1.
It results from the minimimization of the sum of length of the lines, and it provides some evidence about the nature of the p dimensional scatter (see footnote 23).

The dendogram below (see footnote 23) shows the result of clustering 29 continuous variables. This second set of variables excludes any cost-related observation, since they may reflect factual differencies in inpatient care provision making for 'unlegitimate' differences in hospital output.

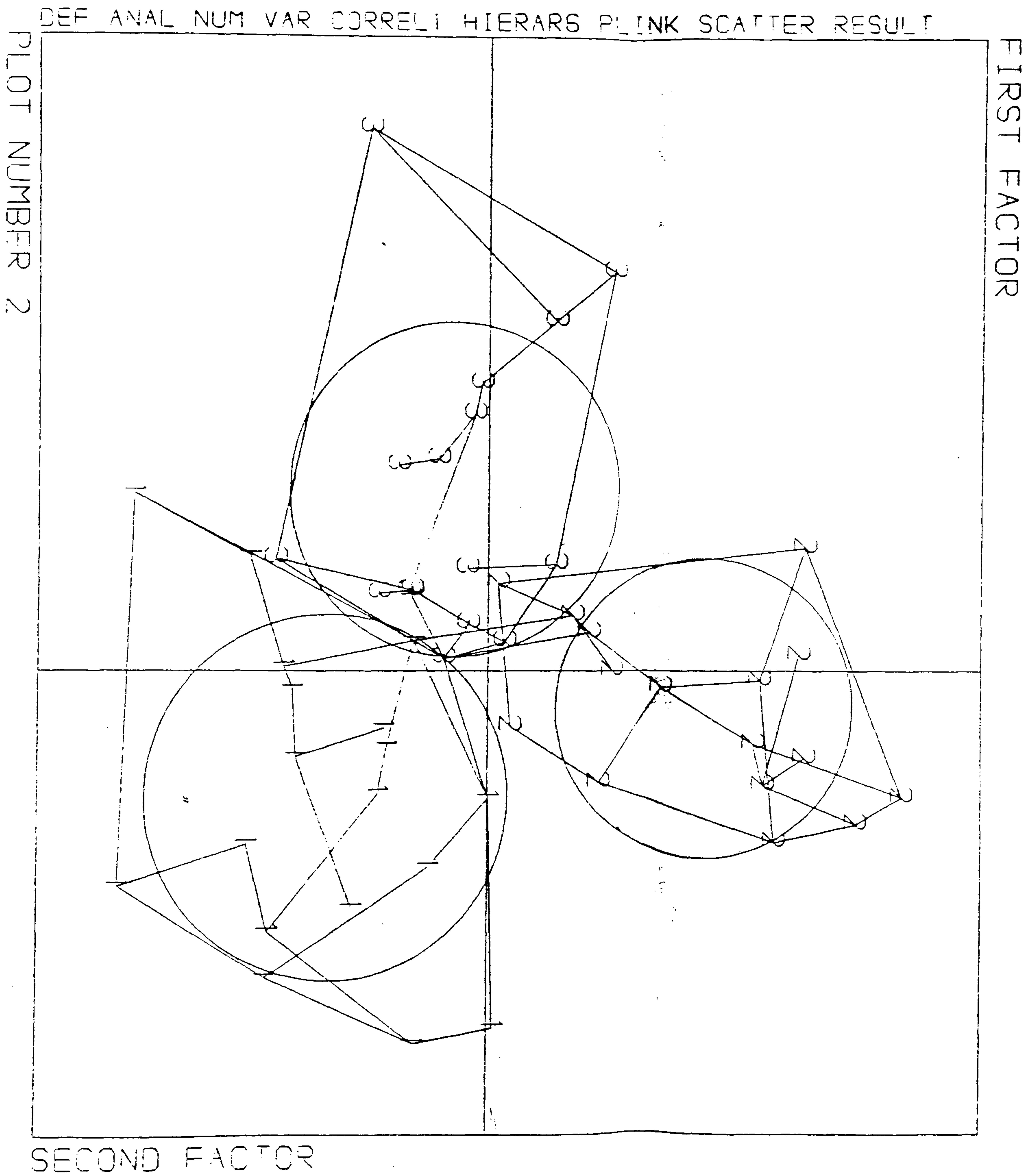


Note that at least three clusters can be considered.
The ratio of inpatient to outpatient cases (and mainly whether the hospital provides outpatient services) and the proportion of higher medical 'hierarchies' over the total number of physicians appear as the two best classification variables for that grouping.

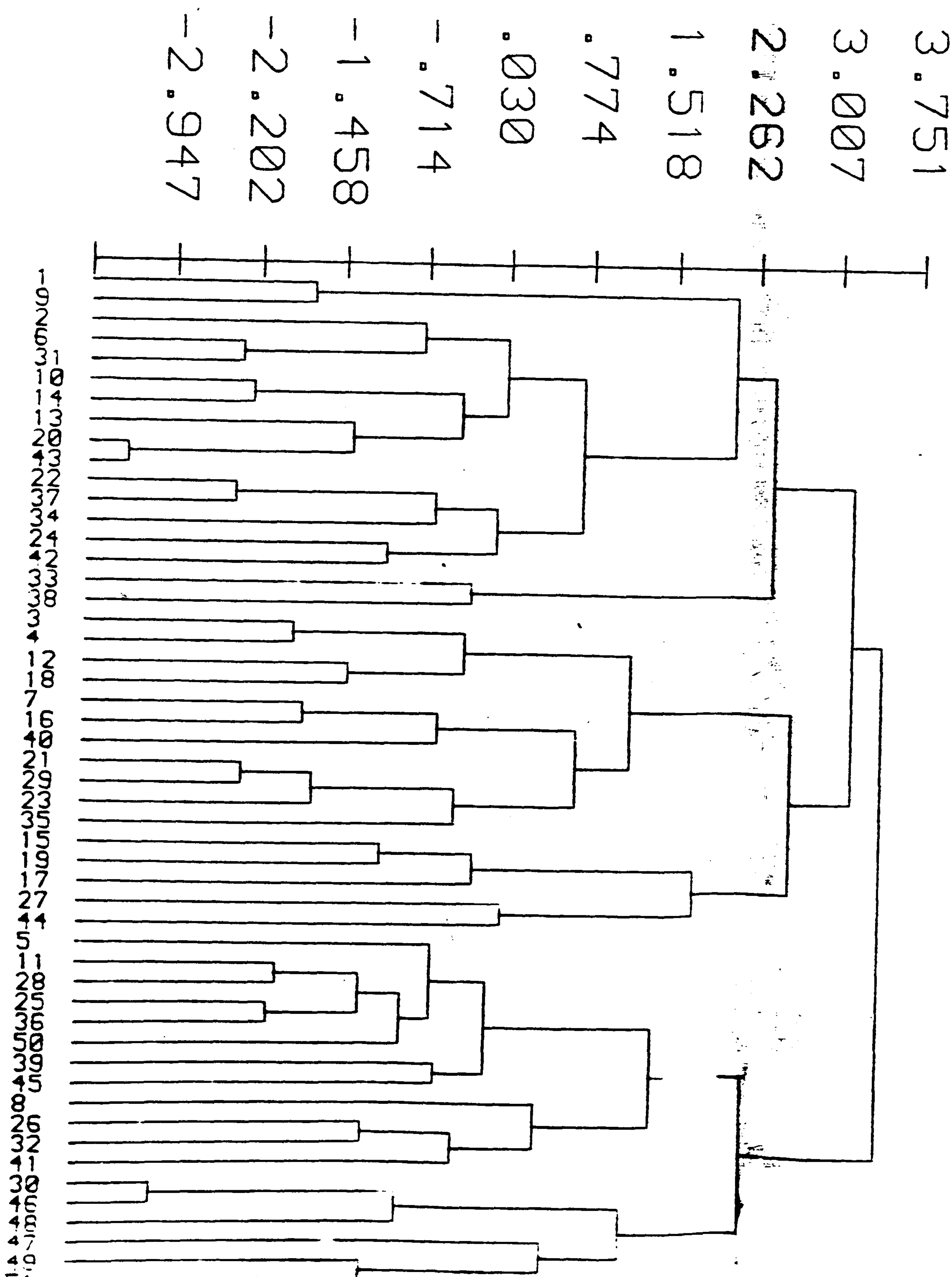
<u>Plot of the first principal component factor against the second one</u>, and minimum spanning tree as derived from the former cluster (A 4). (See footnote 23).



Minimum Spanning Tree (see footnote 23) of the three clusters, as plotted in A 5.

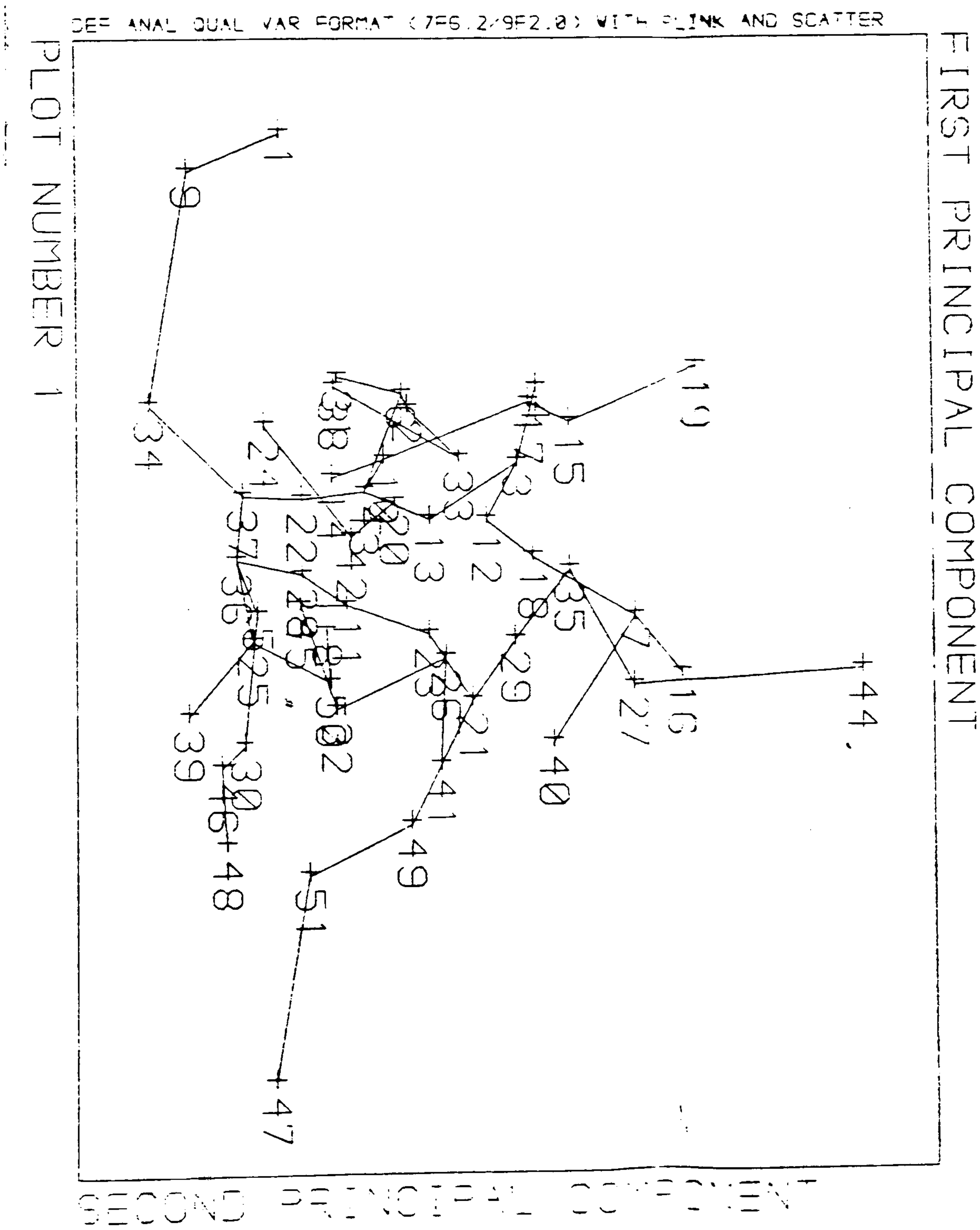


The dendogram below results from applying hierarchical procedures on 7 selected variables, taken as quality-proxies for hospital output. The result seems to suggest that, from the use of this information, a single overall cluster can be considered.

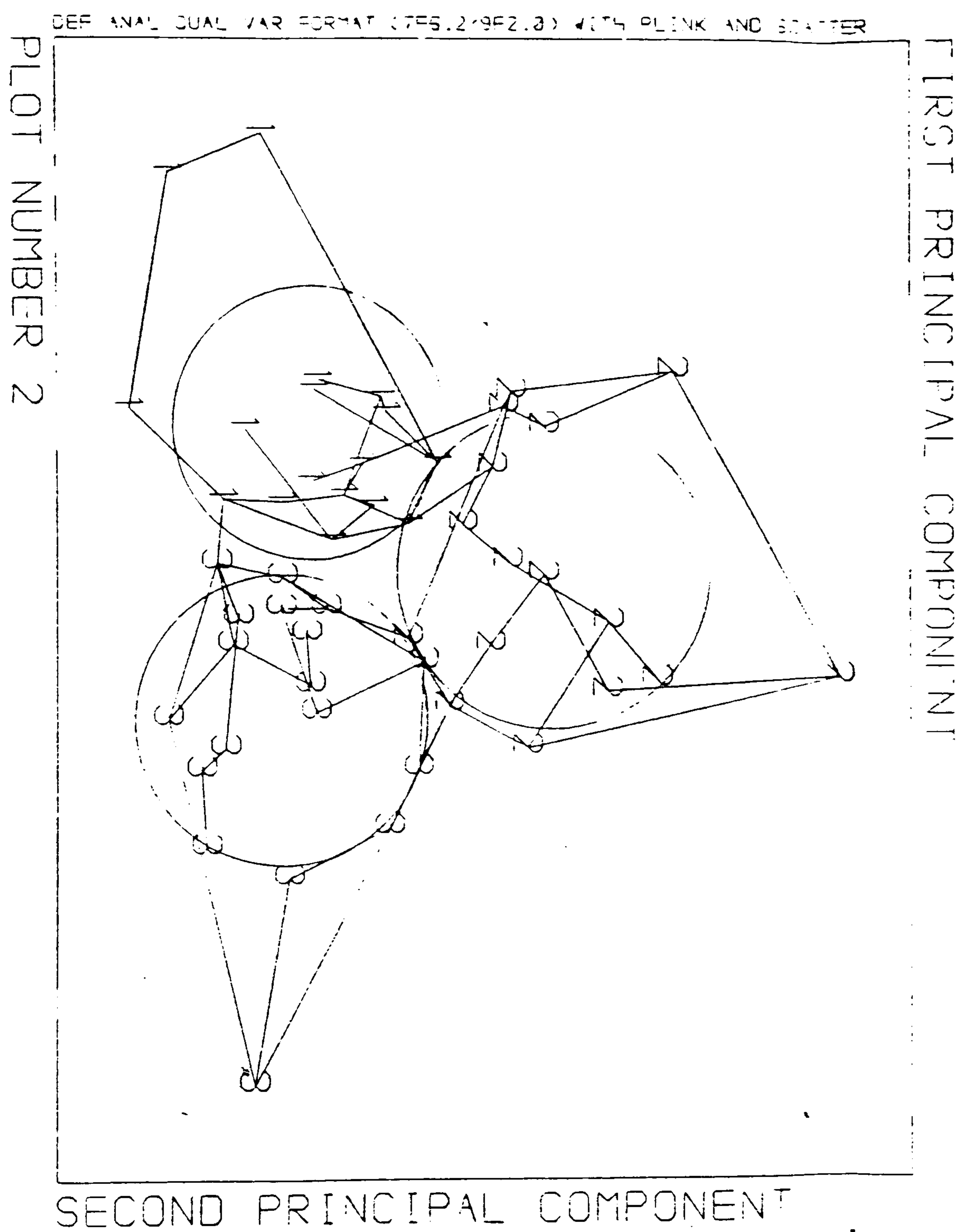


CLUSTERING HOSPITALS REGARDING CONTINUOUS QUALITY PROXY VARIABLES

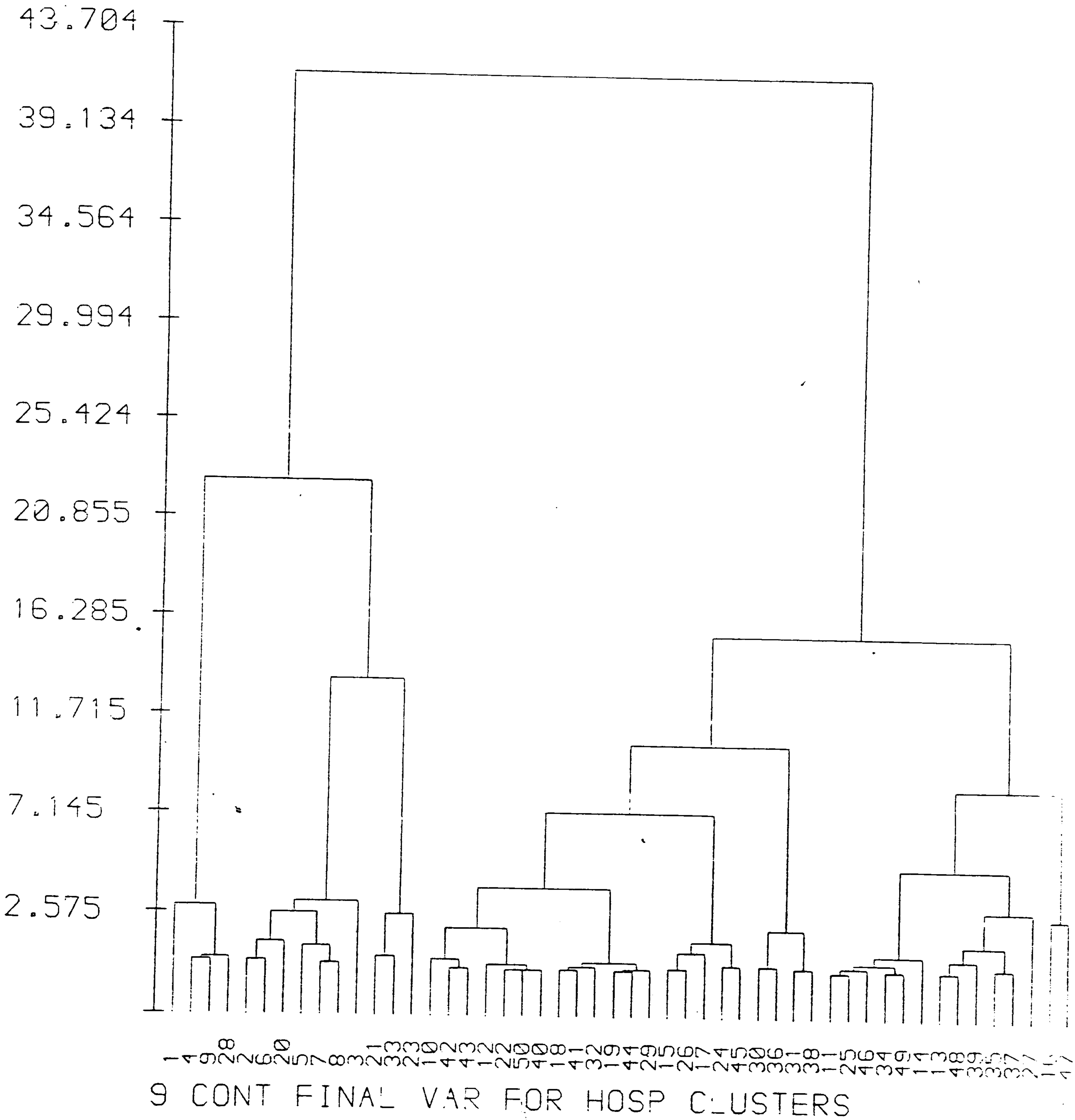
<u>Plot of the two first principal components</u>, accounting for 52% of the cluster variance (see footnote 23), from the former cluster exercise A 7.



Minimum Spanning Tree (see footnote 23) corresponding to A 7.

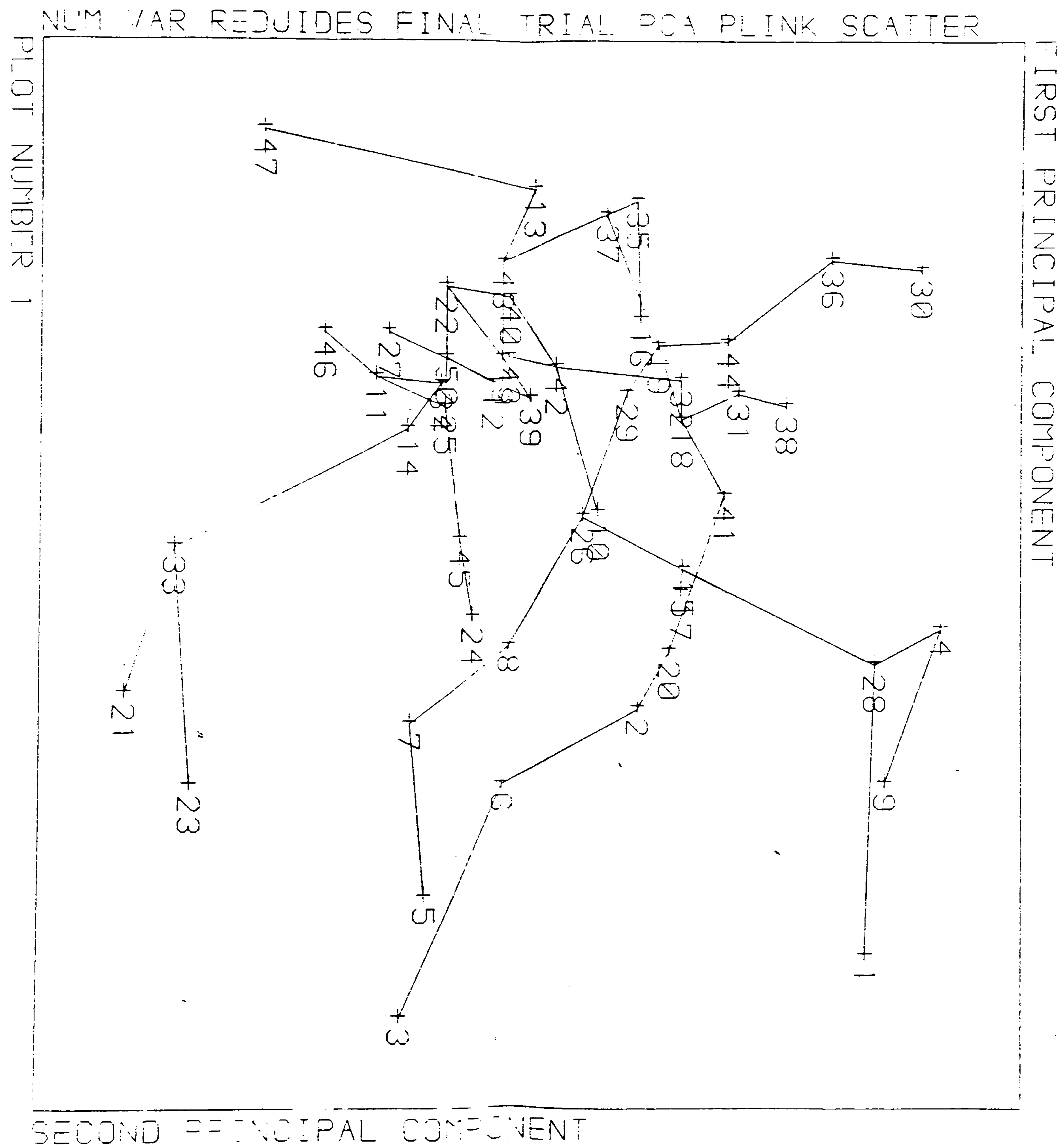


<u>Dendogram</u> (see footnote 23) corresponding to 9 continuous variables finally selected as a result of restricting information to those variables which were <u>thought to be most relevant for defining homogeneity of hospital outputs.</u>

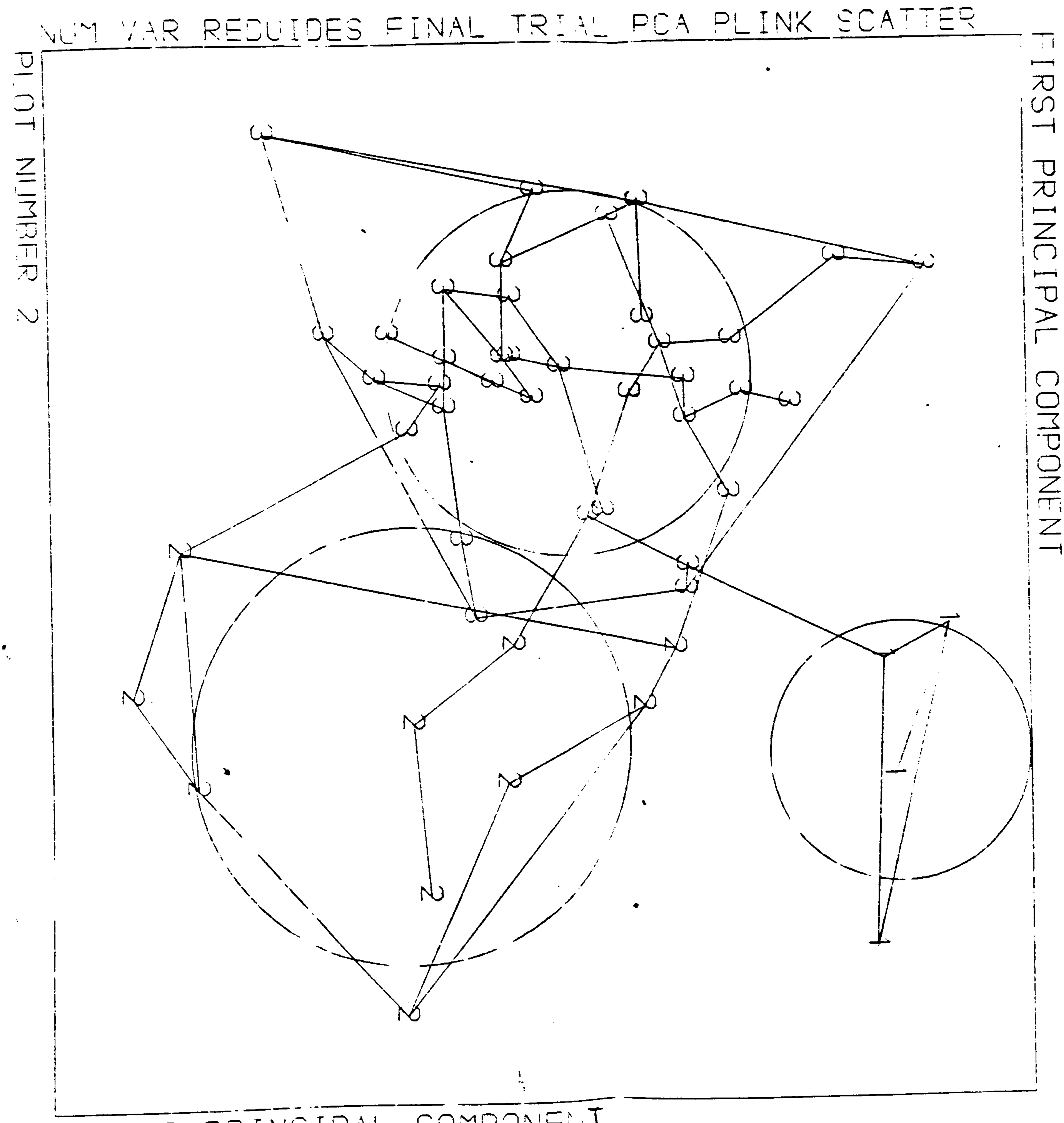


Note that two clusters appear, with the teaching status and the relative importance of the outpatient work-load on inpatient activity as the two diagnostic variables.

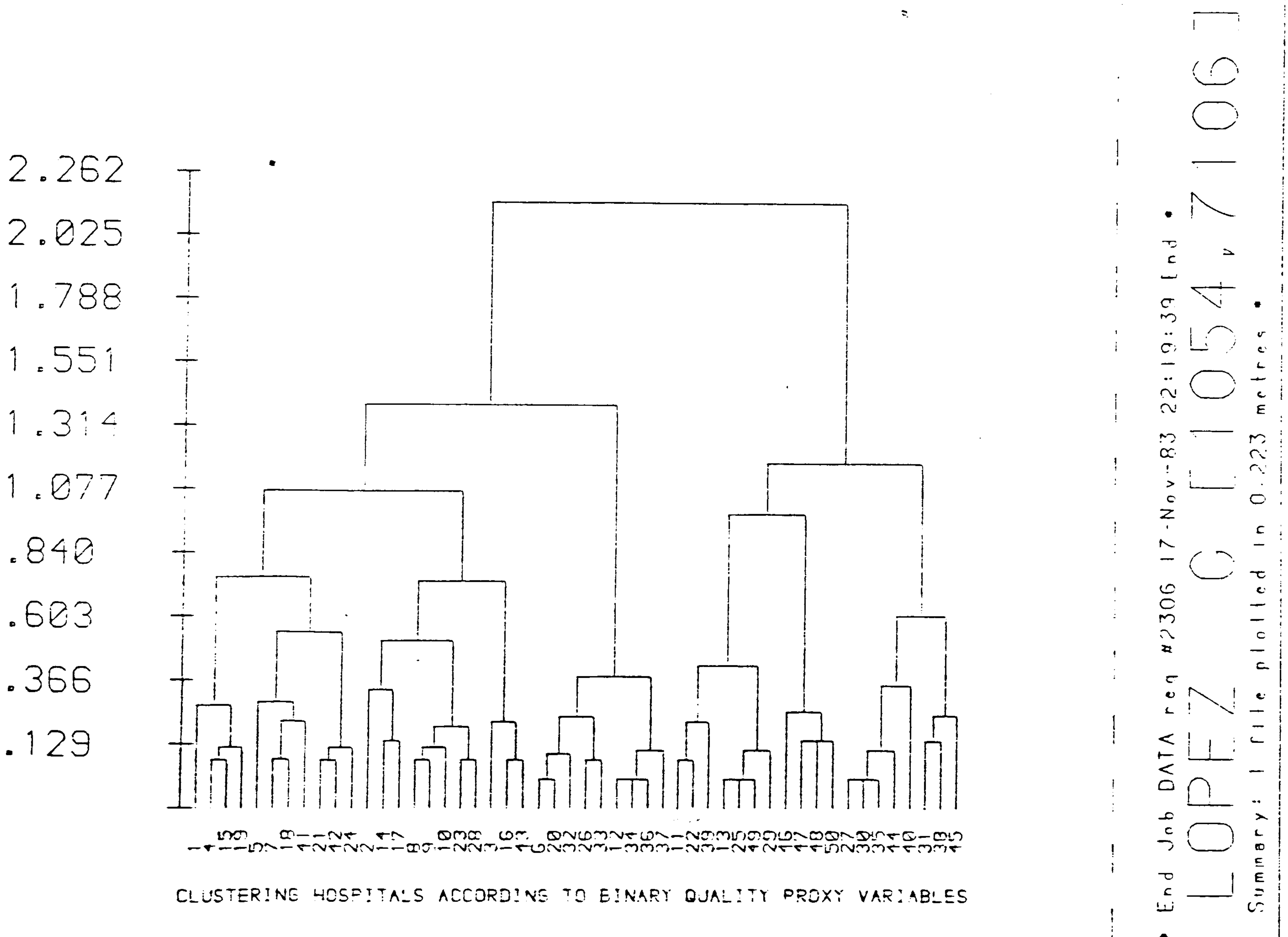
Plot of the two principal components and spanning tree (see footnote 23) corresponding to the result of clustering the 9 continuous variables finally selected.



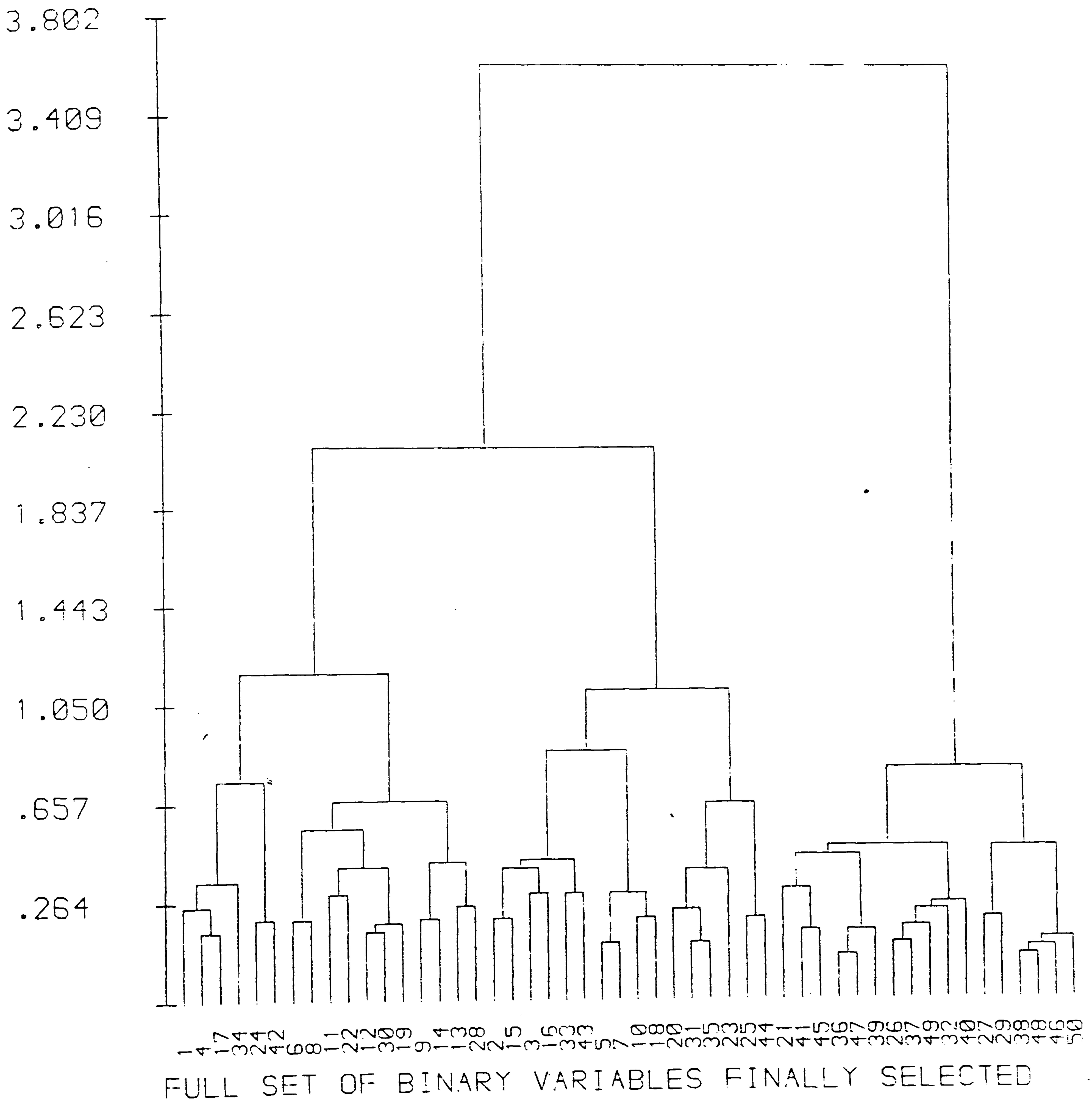
Minimum Spanning tree (see footnote 23) pointing out the two main clusters which appear from the cluster exercise of A 9.



The Dendogram below (see footnote 23) shows the result of applying hierarchical procedures on the selected set of binary variables. The results seem to support the grouping of observations into two main clusters according to the presence or absence of the Dummy variable for the 'medical teaching' status.



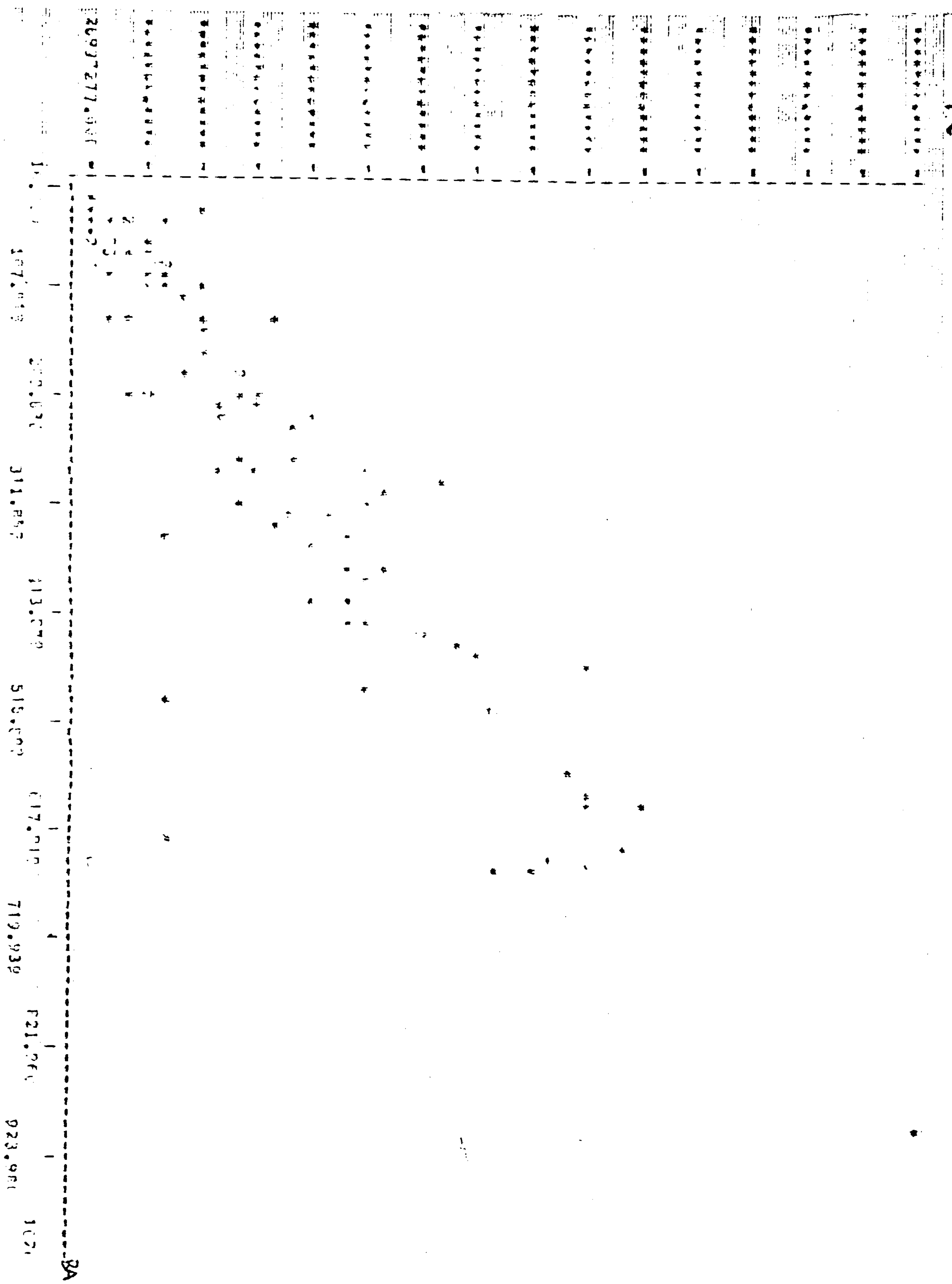
Dendogram corresponding to a second set of binary variables including the transformation of those continuous variables judged as relevant in the previous grouping exercises. (See footnote 23).



This result is similar to the former one, validating the relevance of the 'teaching status' as the best 'classificatory' variable in grouping subsamples with higher internal homogeneity in terms of hospital output (in the way approached by the variables selected).

- GRAPHS -

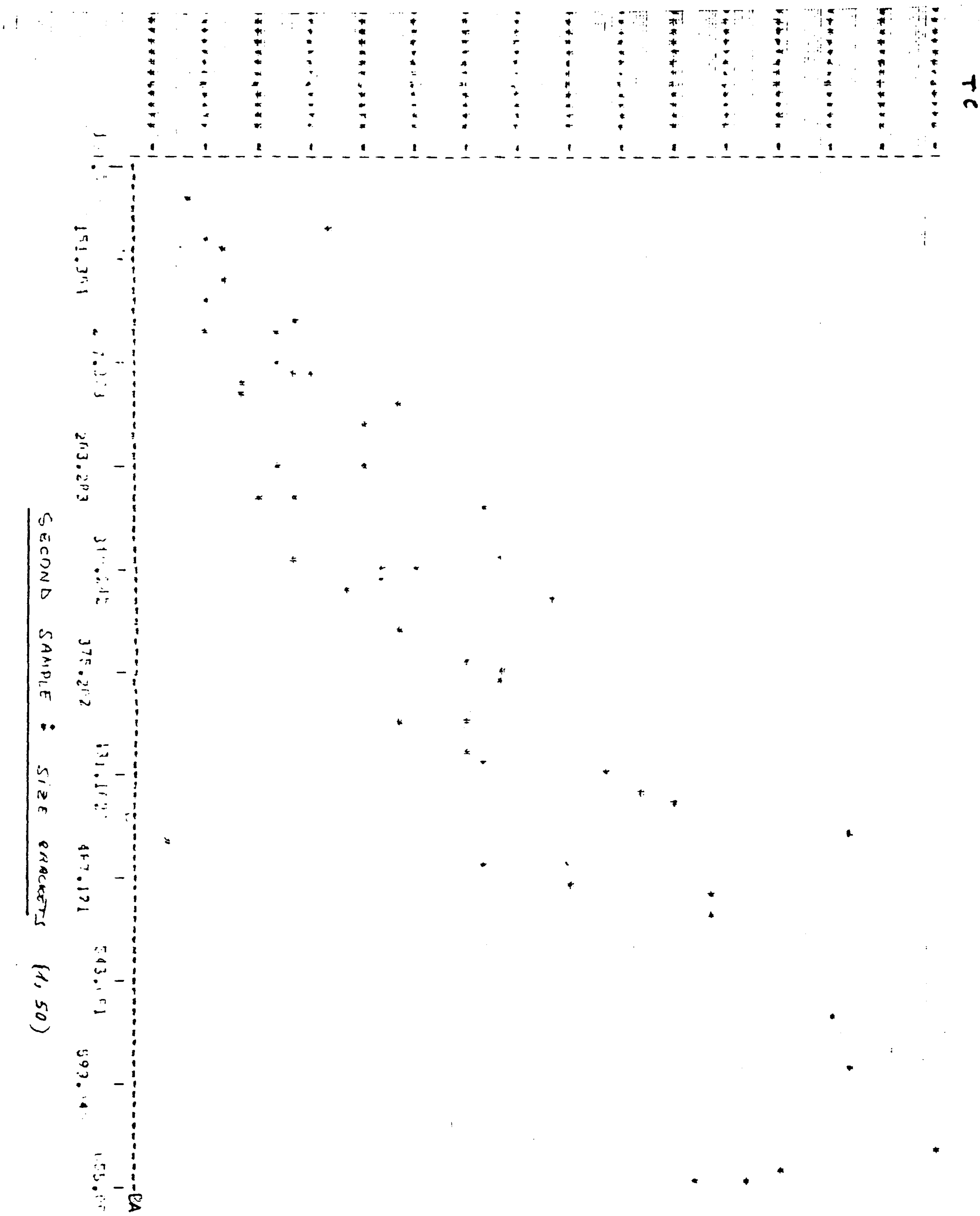
The Graph below plots total operating cost against hospital size (number of beds) for the ninety observations initially selected.



FIRST SAMPLE : GEOGRAPHICAL LOCATION (1, 90)

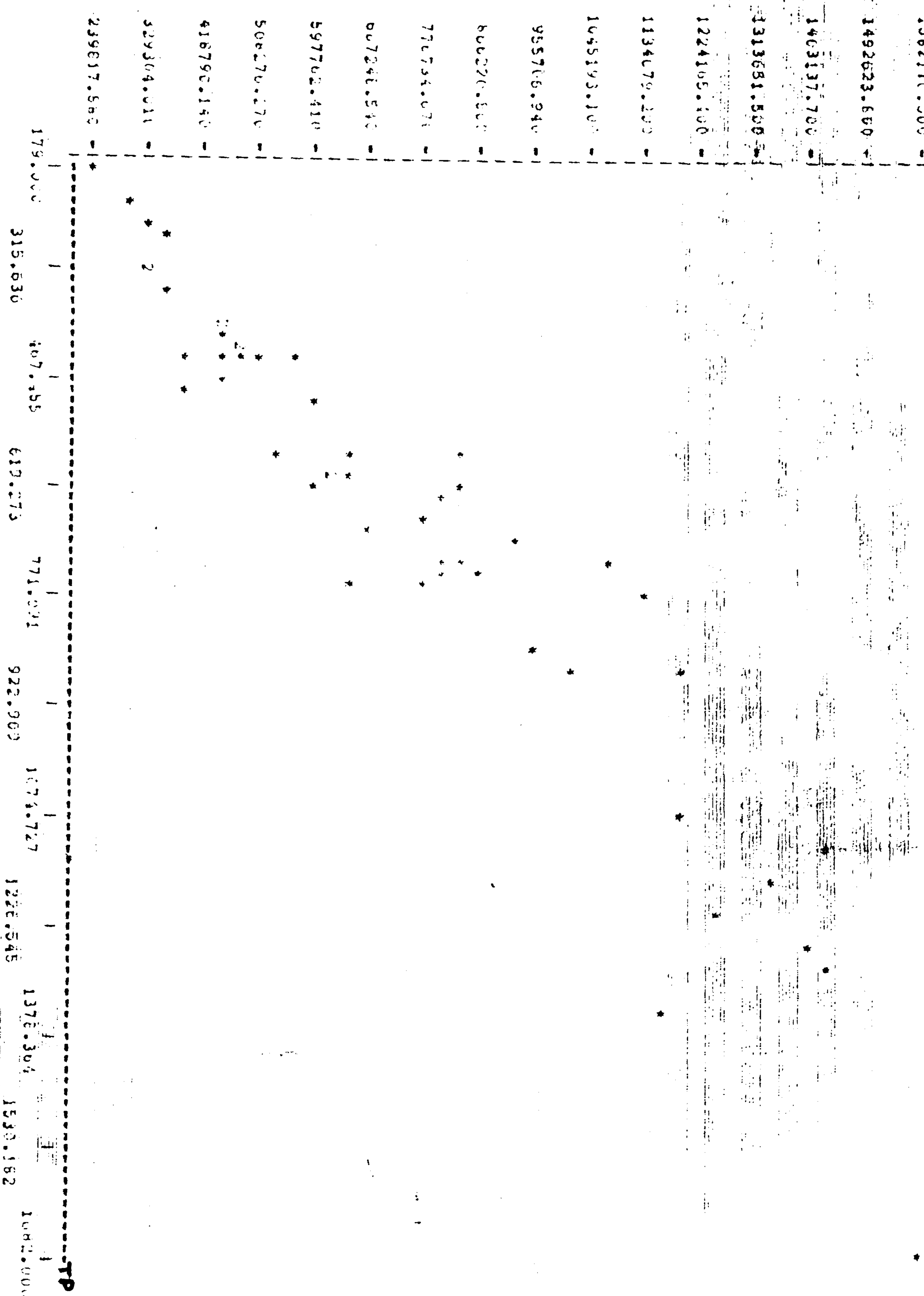
FIGURE I

The Graph below plots total operating cost against hospital size (number of beds) for the fifty hospitals finally selected.



The Graph below plots total operating costs against total hospital manpower for the sample of fifty hospitals.

FIGURE III



SAMPL 1.50. TOTAL OPERATING COST vs TOTAL HOSPITAL MANPOWER

The Graph plots the average cost per case against hospital size for our sample of fifty observations.

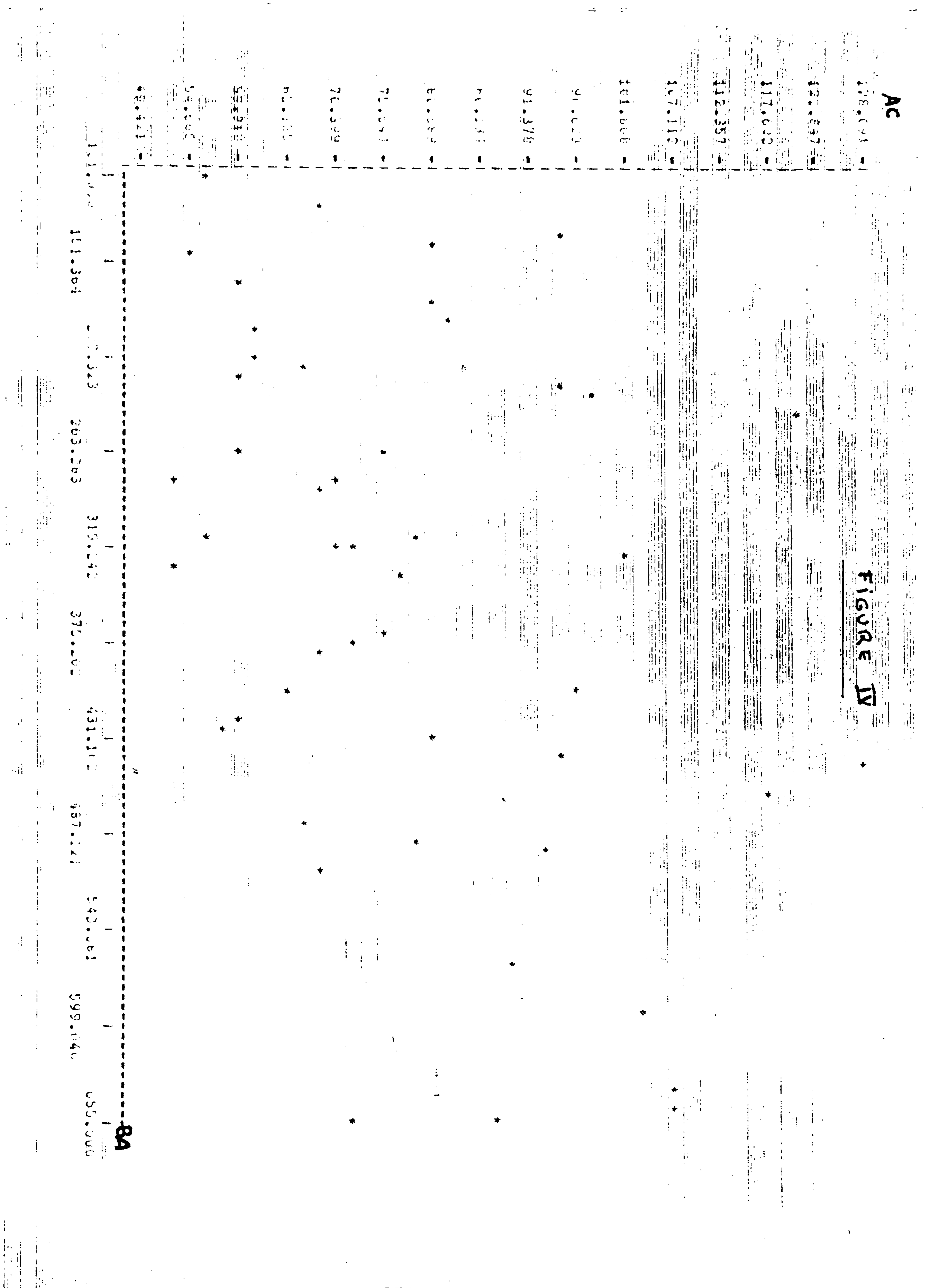


FIGURE IV

The Graph below plots the average cost per case against the flow of cases per bed year for our sample of fifty hospitals.

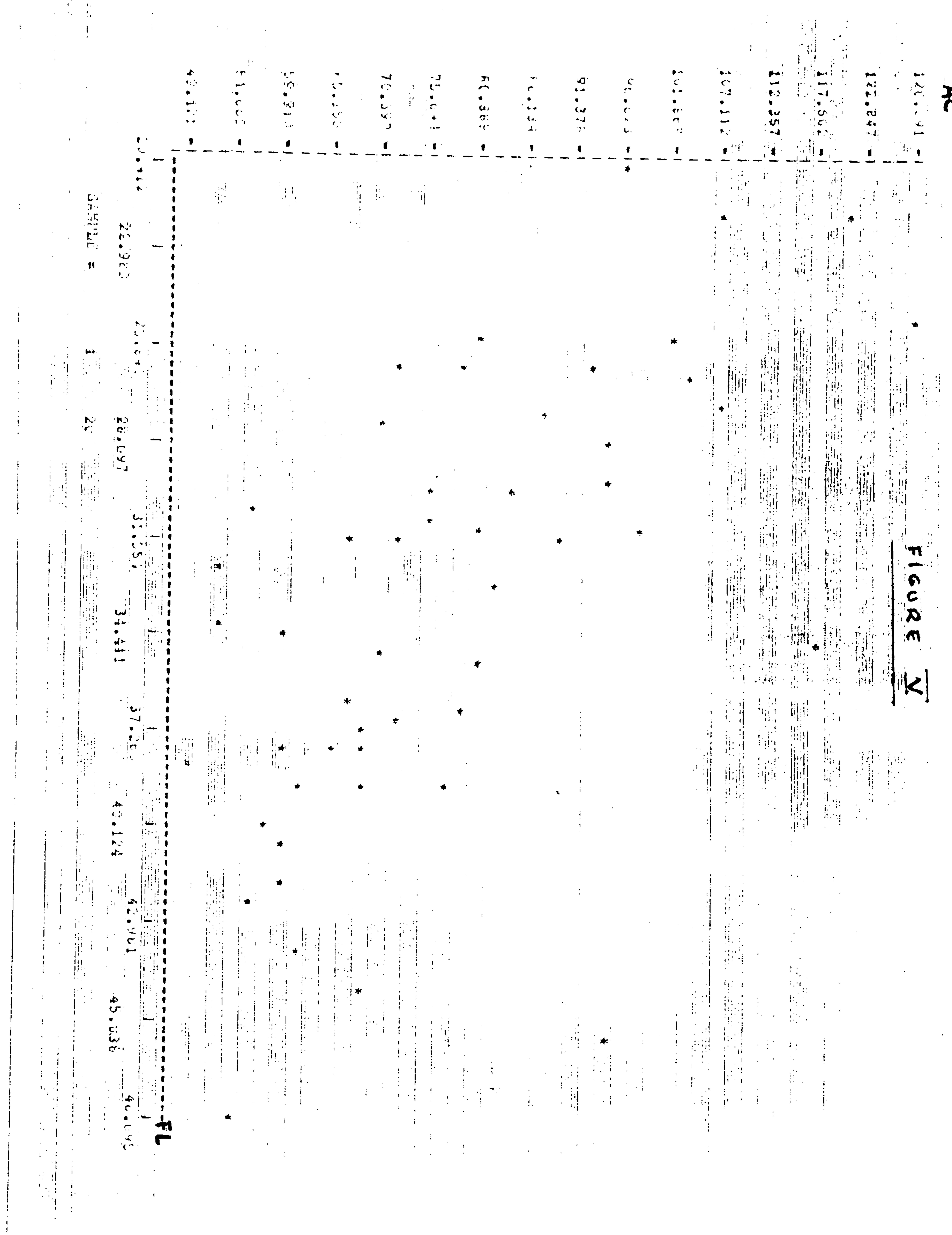
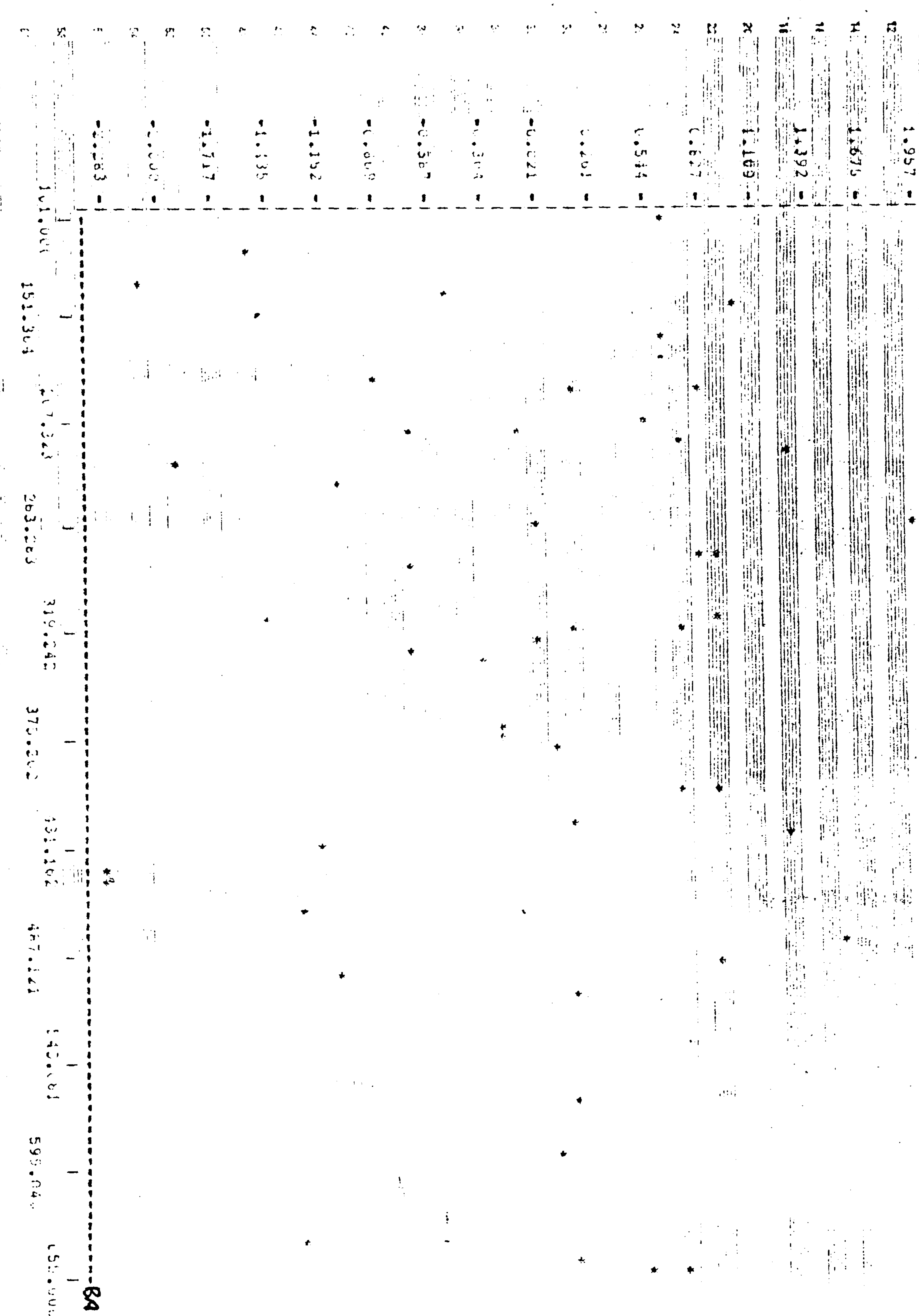


FIGURE V

This Graph plots the residuals from our estimation exercise against the hospital number of beds. No pattern can be derived unambiguously for the overall sample.

FIGURE VI



ANALYSIS OF THE RESIDUALS . OVERALL SAMPLE

This Graph is identical to the former one but it only refers to the subsample of teaching hospitals. Note that a positive association seems to be suggested between 'residuals' and hospital size.

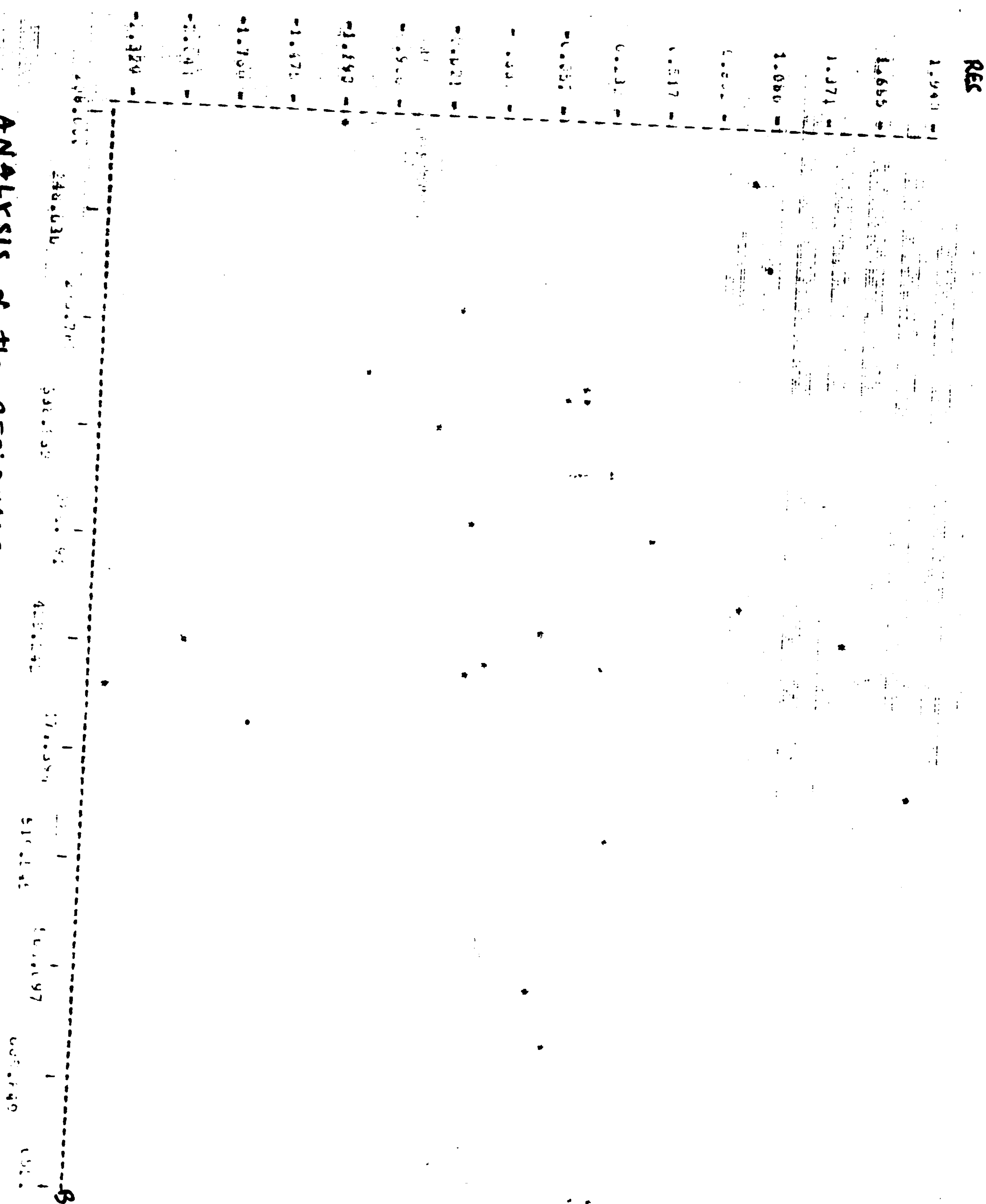
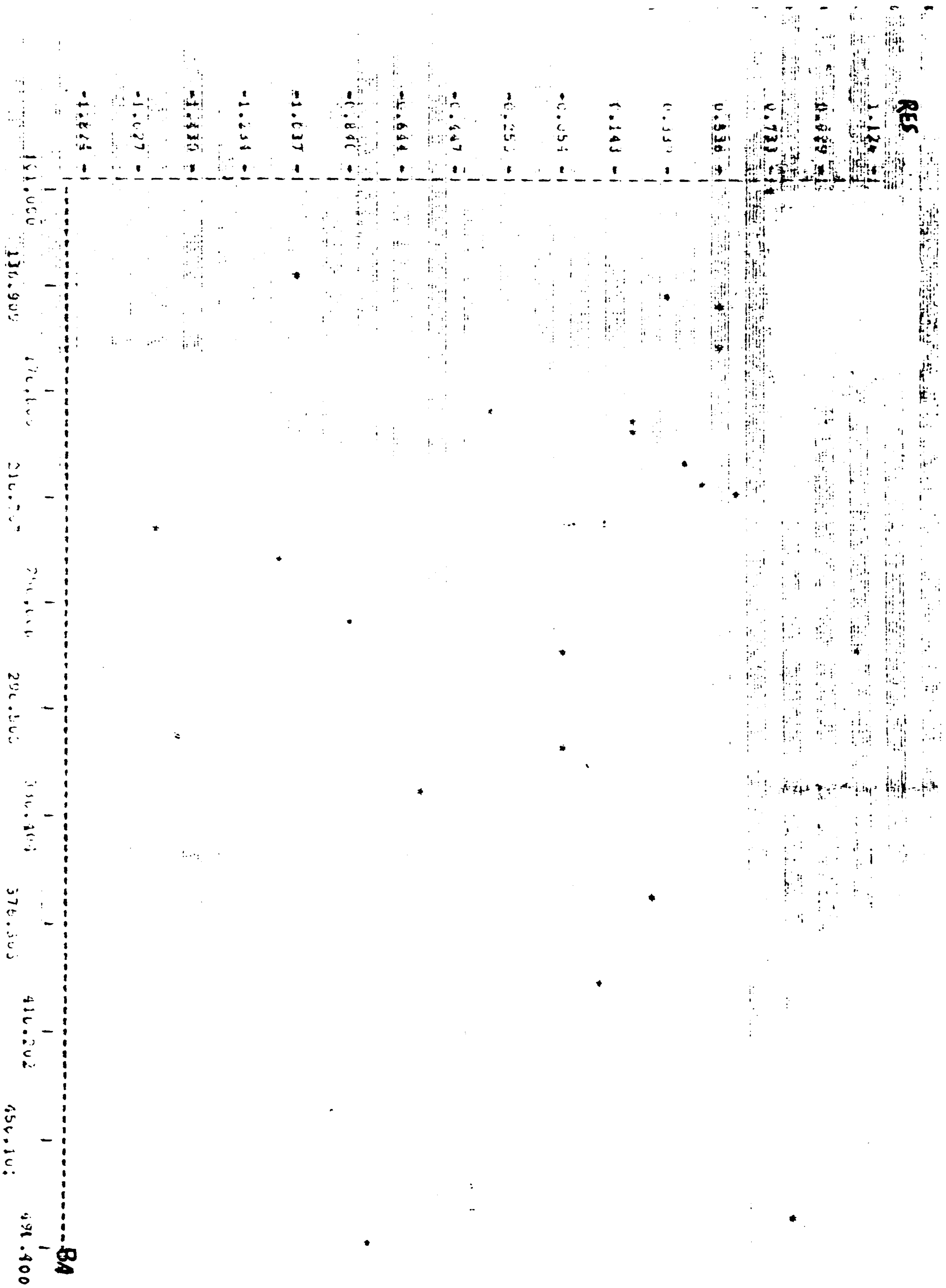


FIGURE VII

ANALYSIS of the RESIDUALS. SAMPLE of TEACHING HOSPITALS

Ditto, for the subsample of non-teaching hospitals.
Note that the pattern suggested from the former Graph seems to vanish.

FIGURE VIII



ANALYSIS OF THE RESIDUALS, SAMPLE of NON-TEACHING SUBSAMPLES

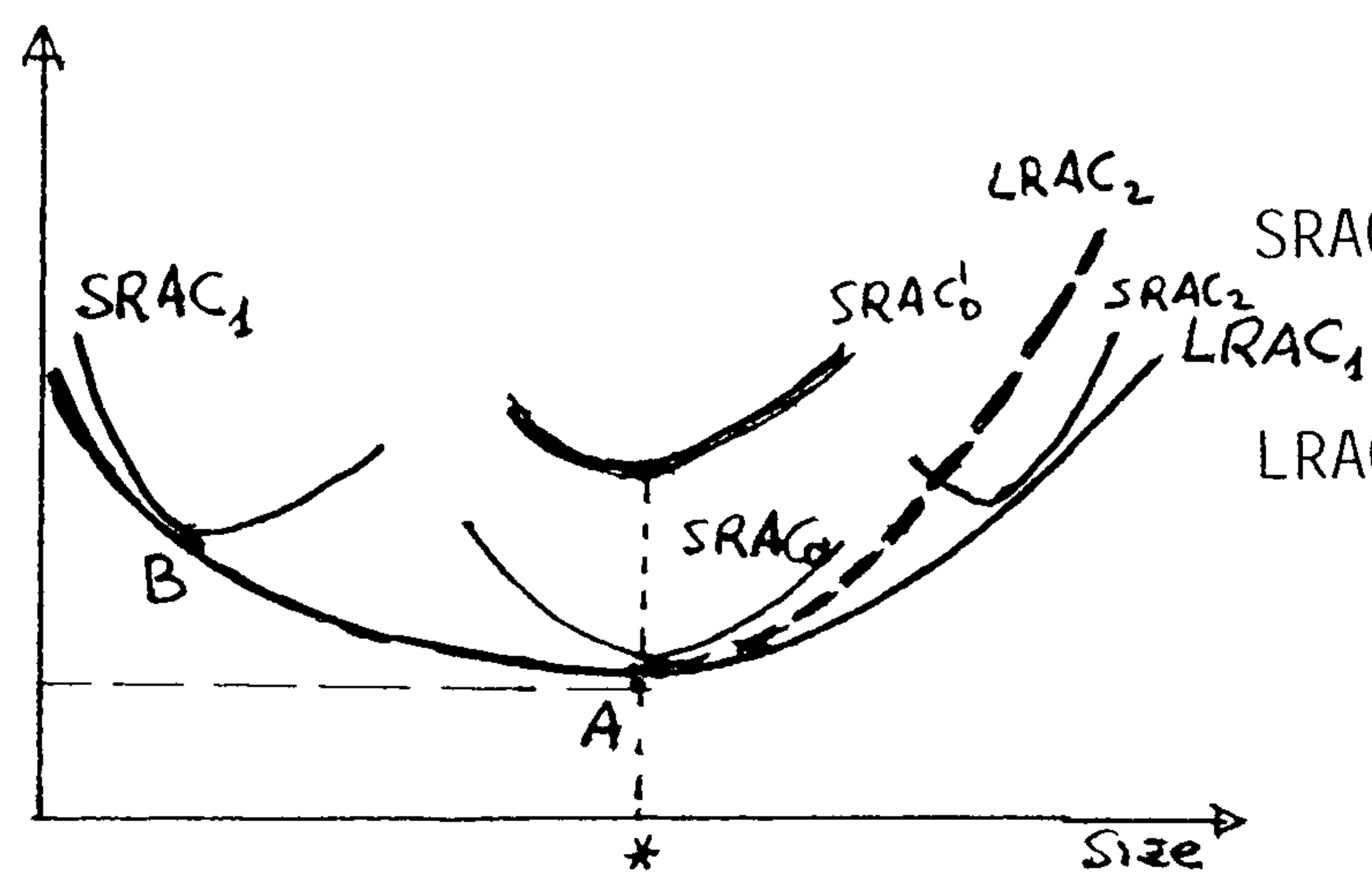
A P P E N D I X   V.2

A research on hospital economies of scale

The question of whether economies of scale could be reached from alternative variations in the hospital capacity has attracted the attention of the literature on hospital costs. The search for a turning point from the results of the estimation of a quadratic 'U-shaped' cost function involves conventionally in setting equal to zero the first derivative, and checking for a positive value for the second derivative, given the estimated parameters of the regression equation.

Figure 1-A

Average cost



SRAC = short run average cost.

LRAC = Long run average cost function.

Optimum plant and scale economies

With respect to the former approach, some points need to be emphasized:

(i) In principle, the use of the standard multiple regression techniques comes to determine the parameters of the empirical cost (production) function so that the deviations of the actual from the predicted values tend to sum to zero. It follows that one obtains (at best) an average cost function, instead of the locus of a technically efficient input-output combination. Its results are, therefore, a simple statistical summary of the 'cost-experience' of the hospitals in the sample. The predicitions about "optimal size" can be then related just to <u>this</u> 'experience'.

(ii) It has to be noted that, in reality, economies of scale are studied commonly with regard to the signs and coefficients of a quadratic function in which hospital size, approached by the number of beds, is regressed on average costs. That is, we are looking at the impact, in the long run, of hospital capacity changes on costs, assuming that the rest of variables remain constant. Therefore, we have to fully account for these latter factors, since some evidence seems to exist in the literature that the 'bed occupancy rate' is positively related to the size of the hospitals, that larger hospitals commonly offer a greater variety of services, keep their patients longer and have a higher throughput, at the same time as they offer medical training programs. Otherwise, if the claim that larger hospitals have to tackle a more complex output is substantiated, this may imply that the right hand side of the average cost per unit of output has a more steeper U shape once reached a certain minimum size (i.e. $LRAC_2$ with respect to $LRAC_1$ in figure 1A), and then it may result in invalidating the estimation procedure.

(iii) In addition, the operation at full capacity with an adequate size on the minimum of an observed average cost curve may not necessarily mean "full efficiency", because, for instance, inefficiency in actual resource management may prevent the hospital from operating at the lowest possible cost level. This can be seen in the case of A' of figure 1-A.

As a result, the examination of efficiency aspects of hospital operation from the single point of view of scale economies offers in itself some caveats.

In addition, for some cases, even when no evidence may appear of a "turning point" on the long run average cost curve, we cannot rule out categorically a 'U-shaped' cost curve for hospitals of all sizes. This may be due to the fact that after a certain hospital size, the presence of managerial diseconomies, and/or a decreasing labour efficiency, may violate the "ceteris paribus" assumption which supports the type of regression analysis performed. In this way, those factors may indeed outweight the economies which would stem otherwise from the postulated optimal size of the hospital.

Bearing these points in mind, we can proceed now to examine the presence of scale economies for Spanish General Acute Hospitals.

As explained in the Chapter, we will take "size" as the plant scale variable, such as defined by the number of beds in hospital. In addition, we will introduce the output-proxy variable, as reflected from the use of the total number of patient-days and the number of cases (with respect to the average cost per patient-day and per-case respectively), as the output scale variable.

For the overall 90 observations (ranging from 50 to 1050 beds), we run an initial regression with number of beds as the independent variable and cost per patient-day as the dependent one. These were the results of this sample estimation.
(values in pesetas):

$$\begin{array}{l} CE = 10397.6 - 9.85\ BA + 0.0084\ BA^2 \\ \qquad\ \ (17.253 \quad (-2.689) \qquad (1.87) \end{array}$$

$$\bar{R}^2 = 0.09 \qquad F(2.87) = 5.174$$

Thus, the coefficients of the estimate show the expected signs for the existence of scale economies. The 'U shape' appeared for all the subsamples specified according to the geographical criteria with a turning point around the 586 beds.

Although the F test enable us to reject the null hypothesis of zero estimated coefficients, the low value of the adjusted R^2 suggest that, by omitting other variables in the equation, we left unexplained the major part of the variation of the changes in the dependent variable.

We tested for the structural stability of the systematic part of the regression through the CHOW's test (see subappendix 1), and for the specified functional form of the equation formulated through the RAMSEY's modified RESET test, by introducing the square value of the predicted dependent variable as an additional independent variable in the initial equation.(*)

In addition, we tested against the violation of the 'homoskedastic' assumption through the Goldfeld and Quandt test, according to certain initial expectations. These were, that the varience of the disturbances among the different subsamples decreased once we moved from the first subsample - for which the observations included a larger variation in hospital size - to the third one - with a lower range. Arbitrarily, we left out in the middle 20% of the observations in order to get an appropriate power for the test. This exclusion affected most of the second and more homogeneous subsample.(**)

The use of the average cost per case as the dependent variable gave some weaker results. Although the parameters appeared again with the expected signs, the value of the adjusted R^2 was still rather

---

(*) - The t statistic for the squared value of the predicted endogenous variable was 0.945. The calculated value for the F statistic was 0.606 which was inferior to 3.05, or critical value for F(0.05;2,86).
In both cases, these results allowed for the acceptance of the null hypothesis such as specified.

(**) - As we explain in the chapter, we later tested more appropriately for heteroskedasticity by ordering our second selected sample by the number of beds, thereby validating this test for the hypothesis of homoskedastic disturbances. For this second exercise, F showed a value equal to 1.448 which was inferior to the critical value of F(0.05; 33,33) $\sim$ 1.739. We could not reject, then, the null hypothesis of homoskedasticity.

low. Moreover, we could not reject the hypothesis of a null coefficient vector in the regression estimation.(*)

Finally, although there existed an obvious risk of mis-specification (**), we used some measure of output as the proxy-variable for the hospital scale.

We took first the cost per patient day (CE) as the dependent variable, on which we regressed the total number of patient-days (TE) as the explanatory variable. These were the results (values in pesetas):

$$CE = \underset{(19.73)}{10514.2} - \underset{(-3.261)}{0.0386}\,TE + \underset{(2.225)}{0.00000012}\,TE^2$$

with $\bar{R}^2 = 0.14$ and $F(2,87) = 7.93$ $\quad (>F^*(99;\ 2,\ 87) \sim 4.95)$

As was the case before, the 'U shape' appears again unambiguously, with a turning point around the 161,000 days. Some caveats can be expressed, however, about the exact amount of this figure, given the low adjusted R^2 statistic - even when we could reject the null hypothesis of zero estimated coefficients at the 99% level of confidence.

(*) - We could reject the null hypothesis for the presence of homoskedasticity, but not the hypothesis of an alternative functional form. Although the power of the RESET test is lower in presence of heteroskedastic disturbances, the t ratio for the square value of the predicted dependent variable was 2.497.

(**) - Since a case-mix adjustment was not possible for the overall sample, we had to assume that there was no systematic influence of the case-mix on average costs.

The result of the test for the functional form allowed for the non-rejection of the already specified functional form, at the usual levels of significance.(*) In similar terms, the null hypothesis of structural stability across the two subsamples was **not** rejected.(**)

We regressed thereafter on the average cost per case (AC), the number of cases (C). A similar 'U-shaped' short run cost function appeared with significant t - statistic values despite the low adjusted R^2 for the equation.

$$AC = 329{,}941 - 79.44\ C + 0.00032\ C^2$$
$$(5.902) \qquad (-4.153) \qquad (4.146)$$

$$\bar{R}^2 = 0.15 \qquad\qquad F\ (2,87) = 8.623$$

(Values in pesetas)

However, for this specified regression, problems of heteroskedasticity could be present according to the results of Goldfeld and Quandt's test (***). It did not pass either the test for the functional form not for structural stability, except for the second and more homogeneous subsample (****). (However, it has to be noticed once again that the power of these tests was already reduced by the presence of

---

(*) - The value of the t statistic for $\hat{CE}^2$ was 1.116.

(**) - The values for the calculated F were 0.901, 0.728 and 1.102 for each of the subsamples considered; all of them result inferior to the critical value $F(0.05;\ 2,\ 86) \sim 3.65$.

(***) - The value of the F calculated was 65.04 ($> F\ (0.05;\ 33{,}33) \sim 1.739$).

(****) - The value of the t ratio for $\hat{AC}^2$ was 2.292 and the F calculated values for the Ramsey's modified Reset test were 16.67 and 15.91 for the first and third subsample ($> F^*(0.05;\ 2,\ 86) \sim 2.51$) and 0.308 for the second one ($< F^*\ (0.05;\ 2,\ 86) \sim 2.51$).

heteroskedasticity. This is due to the lack of efficient properties of the OLS estimates.

In sum, the results pointed out above for the overall sample have to be taken with care. Theoretical and econometric difficulties prevent a clear-cut evidence of scale economies through the conventional regression estimation approach. In addition, lack of data availability does not allow for pursuing a more accurate adjustment. However, some evidence for economies of size seem to appear, particularly with regards to the pattern of the 'U-shaped' average cost function for the overall sample of 90 observations, which includes the largest range of variation in hospital size. This seems to be especially the case when the number of patient-days is the proxy-variable considered for the hospital-output scale. (Whether this is or not a good measure for the output capacity of the hospital has been already discussed in earlier sections of this chapter).

For the _second and final sample selected_ for increasing hospital output homogeneity according to some external factors of provision - sample which records a smaller range of variation in hospital size - and for the subsamples resulting from the clustering exercise, scale economies can be seen from the examination of equations (1'), (2'), (6') and (9') in table V.6.

The questions of functional forms has been already analysed and they will not be discussed further here.

The impact on costs of the utilization variables.

Two lines of research have been followed with regard to the study of the impact on costs of the utilization variables:

a) The first one considers the impact on the average cost per case of an extra day of a patient in hospital;

b) The second line looks for the impact on total costs from increasing the number of cases treated in the hospital.

We will examine first a).

The study of a) responds to the question of whether the marginal rather than average cost per day provides a more accurate measure of changes in hospital utilization, as measured by the number of stays. (In fact, some indications seem to point out that the pace of the average cost over time involves a sequential decrease as length of stay increases).

The result of the estimation of the equation [1] indicates that for 1979, the average cost per patient day is above the marginal cost for the average LOS of the "sample" hospital.

This is, $\partial AC / \partial LOS = 7358$ ptas., a 82.3% of the average cost per patient day (8940 ptas.).

Diagramatically, this effect can be seen through the following figures:

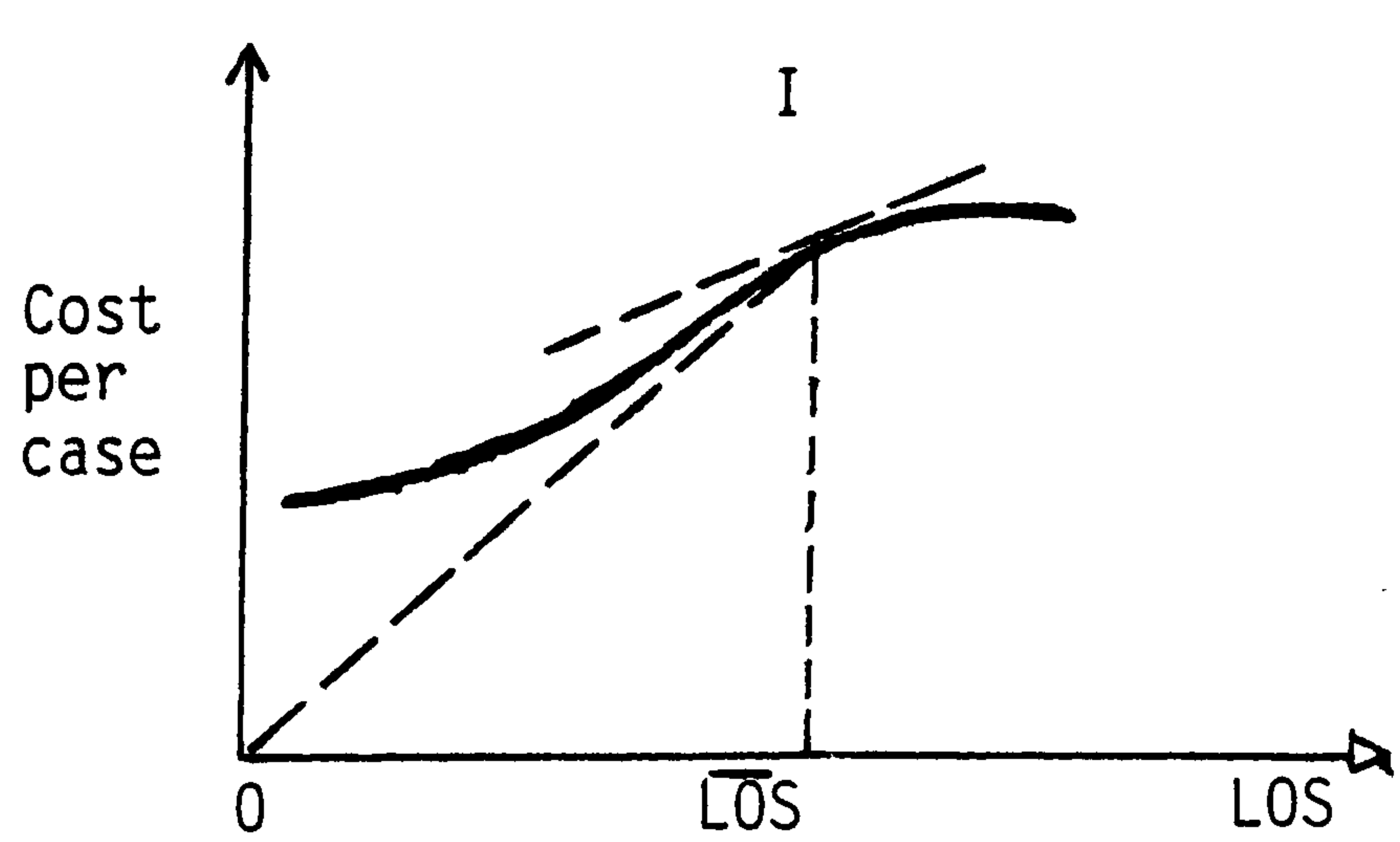
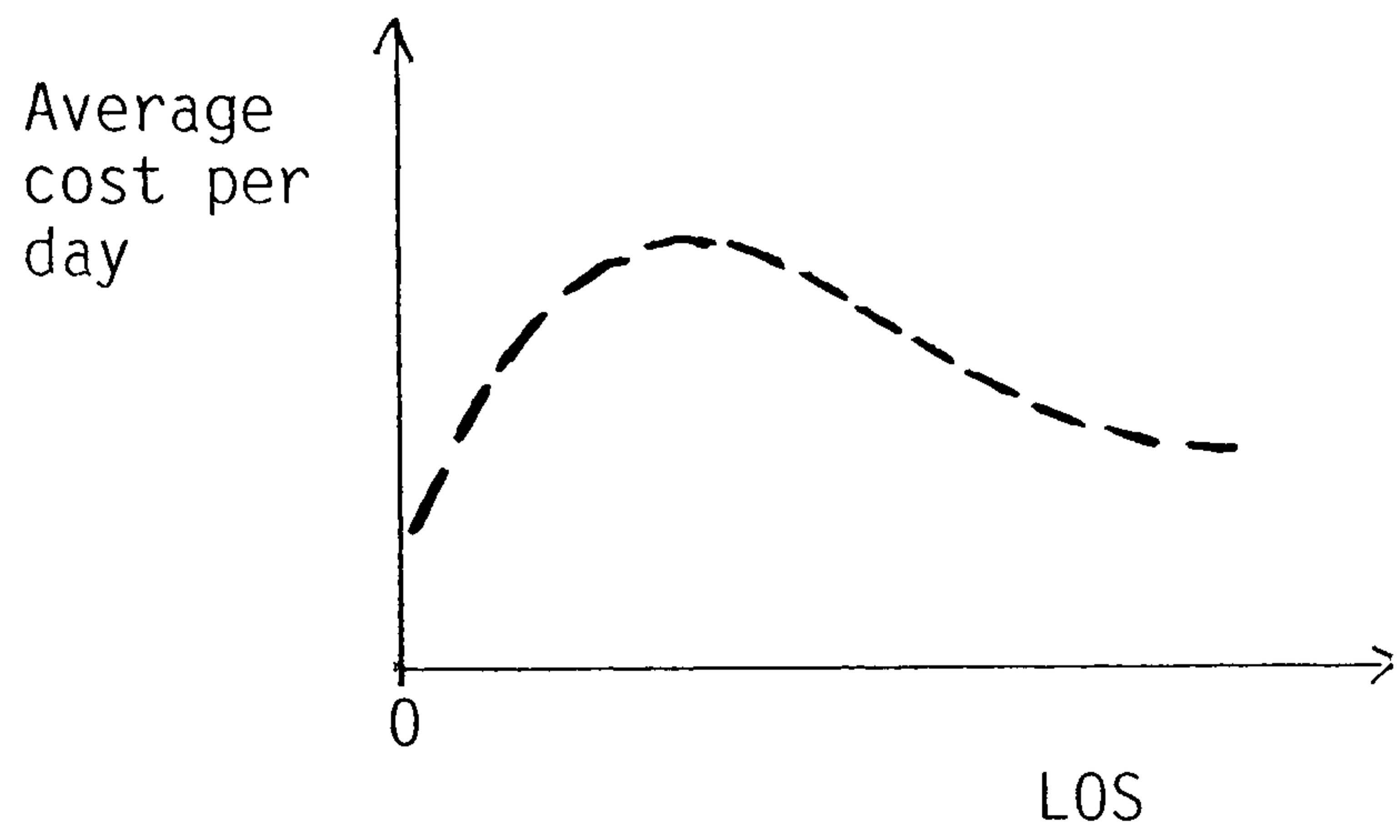


TABLE 1

	Thousand pesetas $AC^{WF=BA}$	C	FL	BOR	LOS	CAS	DVMIR	DVCE	MIECST	CGECST	OBGIST	PECST	CUICST
[1]	$R^2=0.66$ $\bar{R}^2=0.573$ $F_{(10,39)}=7.571$	24.26 (0.65)	-0.173 (-0.23)	-0.396 (-1.55)	7.358 (2.87)	– –	0.572 (0.15)	-10.23 (-1.82)	-1.478 (-0.04)	6.811 (0.25)	70.05 (2.39)	14.005 (0.38)	-9.139 (-0.09)
[2]	$R^2=0.59$ $\bar{R}^2=0.496$ $F(9,40)=6.348$	102.3 (3.67)	-2.085 (-5.75)	0.152 (0.83)	–	–	0.537 (0.13)	-9.47 (-1.55)	10.193 (0.26)	13.473 (0.46)	93.79 (3.07)	-10.13 (-0.26)	-40.51 (-0.37)
[3]	$R^2=0.64$ $\bar{R}^2=0.56$ $F(9,40)=7.868$	53.32 (1.62)	-1.165 (-2.99)	–	4.38 (2.53)	–	-0.721 (-0.192)	-10.02 (-1.753)	-0.458 (-0.012)	1.532 (0.057)	77.56 (2.63)	10.693 (0.29)	-27.64 (-0.27)
[1']	$R^2=0.88$ $\bar{R}^2=0.85$ $F(8,41)=36.49$	$1002.7E^3$ (2.97)	–	-4587.4 (-2.62)	91778.1 (6.16)	69.41 (12.43)	–	–	590443.0 (1.21)	104759.0 (0.285)	109216.0 (2.938)	376801.0 (0.793)	664541.0 (0.51)
[2']	$R^2=0.76$ $\bar{R}^2=0.72$ $F(7,42)=19.30$	543.87 (1.2)	–	-5424.7 (-2.27)	–	77.93 (10.50)	–	–	1108880.0 (1.68)	848053.0 (1.782)	1345130.0 (2.655)	-180274.0 (-0.28)	1377130.0 (0.768)
[3']	$R^2=0.86$ $\bar{R}^2=0.83$ $F(7,42)=35.73$	$1240E^3$ (3.57)	–	–	94808.4 (5.98)	65.87 (11.38)	–	–	484426.0 (0.934)	-124088.0 (-0.33)	932218.0 (2.38)	656340.0 (1.33)	732268.0 (0.522)
	Thousand pesetas $TOC^{WF=BA}$	C	FL	BOR	LOS	CAS	DVMIR	DVCE	MIECST	CGECST	OBGIST	PECST	CUICST

this being likely to be due to the following relationship:

II



(Note that for any particular stay, the average cost is shown in the graph by a ray from 0 up to the point of stay on the distribution, whereas the cost of an additional day, by the slope of the distribution at that point on the stay distribution).

Therefore, as HURST (*) points out, average costs fall with stay, approaching marginal costs as stay increases. (In this sense, whereas average costs are not a close measure for marginal costs in an acute short-term hospital, they may be in long stay hospitals).

Equation [1] also allows to calculate the effect on average cost per case of an increase, say, in the bed occupancy rate (BOR) - due to an increase in admissions per bed (FL) - holding constant the average length of stay (LOS), or resulting from variations in LOS, pushing down BOR, but holding FL constant.

In fact, according to our results, if LOS increases by half a day and, as a result, occupancy goes up from 77.7 to 80.4%, holding FL constant, the average cost per case would raise according to:

(*) - See J.W. HURST: "Saving hospital expenditure by reducing inpatient stay". HMSO, 1977. London

$AC = 0.5 \times 7358 - 0.396 \times (80.4-77.7) = 2610$ ptas.

In similar terms, the effect on the average cost per case of an increase of BOR from 80 to 85%, due to an increase in admisions (FL), holding LOS constant, would involve a decrease of total costs amounting 317.700 ptas.

$$TOC = -0.396 \times (85-80) + (-0.173) \times 5 \times 365 \quad \text{(thousand ptas.)}$$

(Note that, in this case, since we do not know the exact amount in the variation of FL, we infer its value from the relationship $FL=(BOR/LOS) \times 365$, taking into account the constancy of LOS).

Equations [1] , [2] , [3] , imply different types of interpretations of marginal changes in utilization variables on the average cost per case.

Thus, [2] has to be interpreted as the effect of an increase in FL, once we account for changes in BOR (this is, being accompanied by variations in LOS, since the number of bed days must remain constant) and [3] can be interpreted as the effect of an increase in FL with LOS constant (this is, through an increase in BOR).

In sum, we can conclude that, in general, the use of the average cost per day of staying in hospital as the basis for estimating the likely magnitude of savings in hospital expenditure would overestimate the true marginal costs a 17.74%(this is the type of effect indicated in figures I and II).

The line b) of analysis tries to answer the question of whether actual average costs could be reduced by the increase of the admission rates, in making a better use of the hospital capacity and resources available.

The short run analysis is based on the specification of the following equation:

$$\text{Total Cost} = \text{constant} + \beta_1 \text{Number of beds} + \beta_2 \text{Number of cases} + \beta_i \text{Case-mix adjustment}$$

The first derivative of total costs with respect to the number of cases (β_2) gives us the marginal cost of treating an extra case with a given number of beds. This can be done in two ways:

a) By increasing the percentage occupancy of the available number of beds (resulting in an increase in BOR, holding LOS constant, e.g. 'filling spare capacity'), or

b) by increasing the number of cases treated with a given number of patient days (decreasing the average duration of stay per case, holding BOR constant).

For deriving the a) type of result, our empirical estimation should account for the LOS variable if efficiency implications have to be derived from changes in hospital utilization. In the case of b) the specification can be stated as in [2'] .

As it can be seen from comparing [2'] and [3'] , the marginal cost per case of a new admission represents a higher value in [2'] than in [3'] , as it could be expected from an efficient allocation point of view. This last marginal cost represents a 93% of the actual average cost per case. Therefore, we can conclude from this result that average costs overstate slightly the marginal costs per case treated.

SUBAPPENDIX ONE

The tests referred here are linked to two crucial assumptions in an Ordinary Least Squares type of estimation. They are:

(i) The absence of heteroskedasticity. That is, that the variance of the disturbances in the regression does not vary among the observations, and hence OLS estimators are the best linear unbiased estimators, being efficient in small and large samples.

(ii) The absence of misspecification in the systematic part of the functional form and/or with regard to the structural stability of the model specified for the overall sample.

(i) - We have seen in this Chapter that there are reasons for expecting the violation of the assumption made in (i) in a cross-section analysis, and, particularly, with respect to the estimation of the total cost function.

We test for the presence of heteroskedastic disturbances following the Goldfeld and Quandt's test.

This test requires the calculation of the ratio between the squared sum of errors between the first and third subsample, once we have made explicit our hypothesis about the order in the variance of the disturbances and we have left out some observations in the middle (say, 20% of the observations of the sample), in order to get an appropriate power for the test(*).

(*) - In fact, the power of the test depends on the value of these f observations. For a very large f it will be small. On the other hand, as f is reduced, the residual variance will likely move close together and this will tend to offset the added power that comes from the extra observations.

This test has otherwise the problem that because no restrictions are

This is,

$$F = \frac{\text{SSE (higher variance sample)}}{\text{SSE (lower variance sample)}} \sim F^*((n-f-2k)/2)\ ,\ (n-f-2k)/2)$$

where $\underline{n}$ represents the total number of observations;
$\underline{f}$, the number of observations considered as the central values;
$\underline{k}$, the number of parameters estimated;
SSE, the Error sum of squares.

Afterwards, the sample value of F is compared to its critical value. The null hypothesis of the presence of homoskedasticity will be rejected if the value of the F calculated is largely greater than the critical value.

(ii) - The test used in this thesis for checking each of the specified functional forms is the RAMSEY's modified RESET test.

This test consists of setting up a fitted value for the dependent variable which has to appear thereafter as a separate independent variable in its second or higher degree (depending on which is the alternative form otherwise thought as more acceptable). If the value of the $\underline{t}$ ratio for this new variable proves to be insignificant at the conventional levels of significance for a certain interval of confidence, we can accept the original functional form; otherwise we can reject it.

Given the characteristics of the test, it will require an initial check for the presence of heteroskedasticity, since, if this is the case, the power of the test will be minorated.

- Finally, in order to detect any potential structural change in the

---

(...) made on the regression parameters, as well as on the error variances, in each of the two regressions run, statistical power is lost. A more powerful test would then take into account the information that the regression parameters are identical for both sets of data and that only the error variance has changed.

specified regression equation or, more in general, the structural stability of the systematic part of the equation, we have used the CHOW's test.

As it is known, in the absence of structural stability Ordinary Least Squares will provide biased and inconsistent estimators.

For purposes of the CHOW's test, the overall sample is split according to the postulated factor which is though to attempt against the structural stability of the model. If, for example, two groups can be suggested,we need to calculate the Error Sum of Squares for each of the groups $\underline{i}$ (ESS_1) for the n_1 observations, say, ESS_2 for the n_2 observations (being $n_1 + n_2 = n$), and ESS_0 for the total number $\underline{n}$ of observations.

The statistic is then calculated as:

$$F = \frac{ESS_0 - (ESS_1 + ESS_2)}{(ESS_1 - ESS_2)} \cdot \frac{n - 2k}{k} \sim F^*(k, n-2k),$$

where $\underline{k}$ is the number of coefficients in the model.

If the F sample value is greater than the critical value of F^* with (k, n-2k) degrees of freedom, we can reject the hypothesis of a stable structure for the $\underline{n}$ observations and the subsample of n_i observations considered. If it is less than that value, we cannot reject the null hypothesis.

This test is also affected by the presence of heteroskedasticity, since the F statistic will be calculated in this case on the basis of inefficient estimators even for large samples.

S U B A P P E N D I X   T W O

Other alternatives of hospital grouping were also considered,
1 - The actual Spanish system of Hospital Classification (or "Catálogo Nacional de Hospitales") - 1980. This system is based on cross-tabulation techniques, using some variables in pre-determined discrete intervals in order to group all the hospitals (public and private) in the following terms:

- -Local, Provincial - or District - and Regional Hospitals.
- -Level of health care (A,B,C).
- -In functional terms as: General, Surgical, Maternities, Psychiatric, Paediatric and 'Other' hospitals.
- -Status and ownership (Central or Local Public Administration, 'Institutional' Administration, private-for-profit and private-for-non-profit).

Since our sample is in itself the result of an initial scrutiny among all the existing hospital units, we came to identify a 90% of the observations within the category of 'District Hospitals', 'Generals' in the functional sense, belonging to the Central Public Sector and with 'A' levels of health care provision.

In relation to the 'A','B','C' levels of health care provision, the "National Catalogue of Hospitals" follows again a cross-tabulation technique, without any additional validation of the size brackets used for the variables considered. These were the selected variables and their value intervals chosen for classification purposes:

- -the number of medical and ancillary personnel per bed: intervals of 1.0, 1.20 and 1.40.
- -regarding some services judged as relevant for the definition of the levels of health care, and mainly those referred to the existence of medical heart services, digestive system, nephrology, rheumatology, hematology, premature new-borns unit, levels of radiotherapy, and the existence of laboratory services dealing with microbiology, biochemistry, hematology and pathological anatomy.

2. - The present system of Hospital's accreditation in Spain.
The system is commonly applied to private hospitals which provide services on behalf of the public sector, in order to determine the appropriate 'per diem' rates to be reimbursed.

In principle, after 1980 they should be based on similar criteria to those stated for the "Catálogo". These are, the extent of the provision, function and level of the inpatient care provided. But in reality, a special body of "Auditors" is encharged with monitoring those aspects so that some variations may exist between the two sources.

3. - Finally, and mainly for planning purposes and scrutiny of the existing manpower requests, hospitals are classified by INSALUD according to bed size and other throughput measures of personnel per bed. These are the figures used at present:

	up to 250	250-500	500-1000	> 1000
Physicians	0.23	0.25	0.28	0.30
Nurses	0.45	0.50	0.55	0.58
Auxiliars	0.40	0.43	0.45	0.48
-subtotal (1)	1.08	1.18	1.28	1.36
Clerks	0.18	0.19	0.19	0.20
Ancillary workers (qualified)	0.09	0.09	0.09	0.10
Ancillary workers (unqualified)	0.33	0.34	0.34	0.34
-Subtotal (2)	0.61	0.62	0.63	0.64
-Total personnel	1.68	1.80	1.90	2.00
Junior doctors	-	0.05	0.10	0.10

TABLE 2-A.

Geographic location of HR:	Classification		Nº of public HR beds	Coverage Population/public beds	Nº of private beds	Coverage Population/private beds
ALAVA (2)	S_1^1	S_2^2	847	312.3	273	984.4
ALBACETE	S_1^1	S_2^2	420	780.9	106	3094.3
ALICANTE	S_1^1		971	711.0	⊤	2286.5
ALCOY	S_1^1	S_2^1	184	"	511	"
ELCHE	S_1^1	S_2^1	488	"	⊥	-
ALMERIA	S_1^1	S_2^1	400	985.5	137	2877.3
AVILA	S_1^1	S_2^2	403	438.6	112	1578.2
BADAJOZ	S_1^3		358	634.0	⊤ 47	1294.5
D. BENITO VILLAN	S_1^2	S_2^1	202	"	⊥	
MALLORCA	S_1^2		820	70.0	⊤	
MAHON	S_1^1	S_2^2	117	475.9	688	1002.0
IBIZA	S_1^1		46	1181.0	⊥	
STA COLOMA	S_1^1		40	-	-	-
BURGOS	S_1^2	S_2^2	654	480.0	⊤ 171	2002.6
MIRANDA	S_1^1		64	"	⊥	
CACERES	S_1^1	S_2^1	418	636.0	⊤ 201	2007.3
PLASENCIA	S_1^1	S_2^1	216	"	⊥	
CADIZ	S_1^1		1026	583.0	⊤	
JEREZ	S_1^1	S_2^2	334	"	773	1292.9
LA LINEA	S_1^2	S_2^1	141	"		
ALGECIRAS	S_1^1	S_2^1	212	"	⊥	
CASTELLON (2)	S_1^2	S_2^2	682	628.6	559	766.9
CIUDAD REAL	S_1^2	S_2^2	367	614.0	⊤	
MANZANARES	S_1^1		96	.	171	2703.5
PUERTOLLANO	S_1^1	S_2^1	189	"		
VALDEPEÑAS	S_1^1	S_2^1	101	"	⊥	
FERROL	S_1^1	S_2^1	316	279.0	880	1213.7

CUENCA	S_1^1	S_2^2	186	1106.8	76	2350.7
GERONA	S_1^1	S_2^1	281	1641.1	853	540.6
GUADALAJARA	S_1^1	S_2^2	224	596.8	61	2191.7
S. SEBASTIAN	S_1^2		950	757.0	772	912.5
HUELVA	S_1^1	S_2^1	581	689.7	148	2707.7
HUESCA	S_1^1	S_2^2	309	685.3	185	1144.6
JAEN	S_1^3		574	696.0	⊤	
UBEDA	S_1^2	S_2^1	275	"	151	4192.8
LINARES	S_1^2		60	"	⊥	
LEON	S_1^2	S_2^1	377	781	⊤ 816	625.1
PONFERRADA	S_1^1	S_2^2	276	"	⊥	
LERIDA	S_1^1	S_2^2	262	1335.3	536	625.3
LOGROÑO	S_1^1	S_2^2	430	566.7	286	852.0
LUGO	S_1^1	S_2^2	368	1080.3	323	1230.8
EL ESCORIAL	S_1^3		45	-	-	-
ALCORCON	S_1^3		97	1229.9	-	-
CARTAGENA	S_1^3	S_2^1	311	522.9	⊤ 625	1471.6
LORCA	S_1^3		84	770.4	⊥	
PAMPLONA	S_1^1	S_2^2	655	757.7	630	787.2
ORENSE	S_1^3		676	603.1	205	1986.9
MIERES	S_1^3	S_2^1	138	453.0	⊤	
GIJON	S_1^3	S_2^2	496	510.7	585	1941.3
AVILES	S_1^3	S_2^2	227	393.3		
RIAÑO-LANGREO	S_1^3	S_2^1	237	265.7	⊥	
PALENCIA	S_1^1	S_2^1	160	1114.0	51	3949.7
LAS PALMAS	S_1^2		460	1762.9	680	1192.5
LA PALMA	S_1^2		87	"	-	
PONTEVEDRA	S_1^2	S_2^1	319	932	⊤ 921	955.7
VIGO	S_1^1	S_2^2	625	"	⊥	

SALAMANCA	S_1^1	S_2^2	438	763.6	27	12[illegible]
TENERIFE	S_1^1		569	1138.8	548	13[illegible]
STA CRUZ LA PALMA	S_1^1		104	-	-	
SEGOVIA	S_1^2	S_2^2	416	345.9	24	59[illegible]
TARRAGONA	S_1^2	S_2^2	373	326.4	⊤ 3032	
TORTOSA	S_1^2	S_2^1	208	"	⊥	
TERUEL	S_1^3	S_2^2	203	716.5	75	[illegible]
TALAVERA	S_1^2	S_2^2	262	227.2	-	
GANDIA	S_1^3		68		-	
VALLADOLID	S_1^1	S_2^1	638	394.4	78	61[illegible]
ZAMORA	S_1^1	S_2^2	130	1665.2	52	41[illegible]
CALATAYUD	S_1^3		73	"	-	
MADRID	S_1^3		3803	1204.7	3581	1204.[illegible]
BARCELONA	S_1^3		3472	1361.4	6719	1361.4
VALENCIA (2)	S_1^3	S_2^1	2816	173.0	991	173.0
BILBAO	S_1^3		1598	641.7	1229	641.7
SEVILLA	S_1^3		1970	518.6	564	518.6
ZARAGOZA	S_1^3		1731	458.5	694	458.5
OVIEDO	S_1^3		1321	-	585	[illegible]
CORUÑA	S_1^3		925	1359.6	580	1359.[illegible]
GRANADA	S_1^3		929	906.2	398	906.2
MURCIA	S_1^3		947	1241.2	625	1241.2
MALAGA	S_1^3		1077	973.7	592	973.7
CORDOBA	S_1^3		1265	749.8	60	749.8
TOLEDO		S_2^2	588	-	-	-

S_1^1 = Sample 1, subsample 1: 1 Hospital HR in one province
S_1^2 = " 2: More than 1 HR Hospital in one province
S_1^3 = " 3: One or more hospitals in a province with Health City
S_2^2 = Sample 2, subsample 2: Hospitals over the average age
S_2^1 = [illegible] under [illegible]

MAP 3A

● = General Hospitals (Residencias Sanitarias) ✱= Regional Hospitals (Ciudades Sanitarias) ● = Maternity Hospitals

CODE of the variables used.

AC = Average cost per case.

FL = Flow of cases per number of beds available.

LOS = Average length of stay.

BA = Number of beds available.

BOR = Bed occupancy rate.

CE = Average cost per patient-day

C = Total number of cases.

TE = Total number of hospital patient-days.

TC = Total costs.

MIEC = 'Internal Medicine' and medical specialties. Number of cases.

MIECST = idem, but in percentage terms over the hospital number of cases.

DVTD = Dummy variable per teaching doctors.

DVCS = Dummy variable for the provision of outpatient care in the hospital.

DVHP = Dummy variable equal to one if the proportion of physicians with the higher hierarchical positions over the total number of physicians in the hospital was greater than the sample average; 0 otherwise.

DVNCET = Idem as before, but for the proportion of outpatients over the number of inpatient cases treated in the hospital.

CGEC = General Surgery and surgical specialties . Number of cases.

CGECST = idem, but in percentage terms over the total number or cases.

OBGINC = Obstetric and Ginaecology cases . Absolute figures.

OBGIST = idem, but in percentage terms over the number of cases.

PED = Paediatrics (New borns, internal medicine and surgery): no. of cases.

PECST = idem, but in relative terms over the total number of cases.

CUIC = Number of cases in the Intensive Care Units. Absolute figures.

CUICST = idem, but as relative values in percentage terms.

OTRC = "Other cases", as a residual category, in absolute figures.

OTRCST = idem in percentage terms.

OVERALL SAMPLE OF SPANISH PUBLIC HOSPITALS

('HEALTH RESIDENCES'. 1979)

1	Valdepeñas (C. Real) 'G. Ortega'.	NT
2	Mahón. (Baleares) 'V. Montetoro'.	NT
3	Mieres (Oviedo).	NT
4	La Línea. (Cádiz)	NT
5	Palencia 'L. Ramírez'	NT
6	Teruel 'Obispo Polanco'.	NT
7	Alcoy (Alicante) 'V. Lirios'	NT
8	Cuenca 'V. Luz'.	NT
9	Puertollanco (C. Real) 'Sta. Bárbara'.	NT
10	D. Benito-Villanueva (Babajoz)	NT
11	Tortosa (Tarragona) 'V. de la Cinta'.	T
12	Guadalajara 'F.P. Rivera'.	T
13	Talavera Reina (Toledo) 'N.S. Prado'.	T
14	Plasencia (Cáceres) 'V. Puerto'.	NT
15	Lérida 'Gral. Moscardó'.	NT
16	Cartagena (Murcia) 'N.S.Rossell'.	T
17	Ubeda (Jaén) 'S. J. de la Cruz'.	NT
18	Ponferrada (León) 'V.Camino'.	NT
19	Gerona 'Gral. Alvarez Castro'.	T
20	Huesca 'San Jorge'.	NT
21	Riaño-Langreo (Oviedo) 'N.S.Rosario'.	NT
22	El Ferrol (Coruña) 'Arq. Marcide'.	T
23	Avilés (Oviedo) 'S. Agustín'.	NT
24	Algeciras (Cádiz) 'Punta Europa'.	NT
25	Pontevedra. 'Montecelo'.	T
26	Lugo 'Hnos. Pedrosa'.	NT
27	Segovia 'Licinio de la Fuente'.	T
28	Ciudad Real 'Ntra. S. Alarcos'.	NT
29	Jerez (Cádiz) 'Gral.Pr. Rivera'.	T
31	Tarragona 'Juan XXIII'	T
32	León 'Virgen Blanca'.	T
33	Almería 'Virgen del Mar'	NT
34	Zamora 'R. Ledesma Ramos'.	NT
36	Cáceres 'S.P. Alcántara'	T
35	Avila 'Virgen Sonsoles'.	T
37	Albacete 'Perpetuo Socorro'.	T
38	Logroño 'A. Coello Cuadrado'.	T
39	Salamanca. 'Virgen Vega'	T
40	Valencia 'Gral. Sanjurjo'.	T
41	Las Palmas 'N.S. del Pino'.	T
42	Castellón. 'N.S. Sagr. Corazón'.	T
43	Elche (Alicante)	NT
44	Gijón (Oviedo) 'J. Gómez Sabugo'.	NT
45	Huelva 'Manuel Lois'.	T
46	Vigo 'Almirante Vierna'.	T
47	Valladolid 'Onésimo Redondo'.	T
48	Alava 'Ortiz Zárate'.	T
49	Pamplona (Navarra) 'V. Camino'.	T
50	Toledo 'Virgen Salud'.	T
51	Burgos 'Gral. Yague'.	T

T = Teaching hospitals

NT = Non-Teaching hospitals

CHAPTER VI CONCLUSION

6.1 CONCLUSION

CHAPTER VI

CONCLUSION

Conclusion

Regarding the ultimate aims of this thesis and the particular intents advanced in the introduction to this research, and completing those partial conclusions already put forward at the end of each Chapter, the following conclusion can be now outlined. To this can be referred, therefore, the contribution of this thesis:

(i) If health care targets are to be achieved in an age of financial restraint - as in Spain and in some other European countries at present - there exists the need to improve the efficiency with which resources are allocated in the health sector.

(ii) The effort for a more cost-conscious management of resources, which is particularly recognized in the public hospital sector in Spain, requires a change in the present budgetary arrangements set up between the reimburser and the spending units, if efficiency-promoting incentives are to be generated.

(iii) In addition, our analysis shows that an alternative budget policy needs to link, in some way, differentiated levels of reimbursement to the observed levels of differential 'performance', instead of the application of across-the-board increases on individual hospital past expenses.

(iv) This strategy needs to recognize, moreover, that an Agency type of relationship, rather than a hierarchically structured network, is the appropriate setting for monitoring and control of hospitals' operation, and that the allocation of global budgets and attribution of overall value responsibility are, consequently, more reallistic policies.

(v) In this context, a budget-based contract, prospectively advanced under verifiable conditions by both parties of the relation, with respect to the prior commitment of the resources available for distribution within the sector, can prove to be 'optimal' for our purpose; that is, capable of inserting incentives for the implementation of efficiency-guides in the internal allocation of resources and for improving hospital performance over time.

(vi) The operationalisation in the Spanish context of a Prospective Revenue Allocation Policy (PRAP) of the former characteristics, need to be accompanied by some other institutional changes at some more disaggregated steps in the hospital decision-making process, in order to make those financial arrangements not only necessary but also sufficient conditions for a behavioural change.

(vii) The specific design of PRAP for a selected sample of Spanish hospitals on the basis of the prediction of a 'relevant expenditure' line from a multiple regression analysis, enables us to point to some relevant factors for the study of hospital efficiency -such as the need to increase 'within subsample hospitals' homogeneity', and to undertake separate cost analysis according to those cluster attributes.

(viii) Moreover, the actual construction of this "expenditure line" allowed us to validate in the empirical arena some hypotheses advanced on 'system-inefficiency', and mainly, the absence of self-correcting mechanisms in hospital performance. However, the results of the test for the hypothesis of a cumulative causation effect on hospital operation over time did not seem to suport it unambiguously.

Finally, in implementing a PRAP of the nature designed in this thesis, some additional research should likely need to be undertaken in future, mainly (a) in the theoretical field, by modelling in a more comprehensive manner, the hospital decision-making process and hospital behaviour, and (b) in the empirical arena, by working a more appropiate data set, mostly orientated to refine the techniques for comparing hospital performance, in view of a more satisfactory definition of health and hospital output.

BIBLIOGRAPHY

REFERENCES : CHAPTER I

BACON, R. et al, *The Dilemas of Government Expenditure,* The Institute of Economic Affairs (I.E.A.), 1976.

BACON, R. & ELTIS, E. 'Britain's economic problem: The growth of the non-market sector', *The Economic Journal, 89*, June 1979.

BANYARD, R. 'Efficiency savings in the NHS', *Hospital and Health Services Review,* Nov./Dec., 1982.

BANTA, H,D. & KEMP, K. B. (eds.) *The Management of Health Care Technology in Nine Countries,* Springer Publish, New York, 1982.

BECK, M. 'The expanding Public Sector: Some contrary evidence', *National Tax Journal,* vol 2, 1, 1976.

BECK, M. 'Public Sector Growth: A real perspective', *Public Finance,* no. 3, 1979

BEVAN, R. G. 'Cash Limits', *Fiscal Studies,* 1980.

BLUME, S. 'Explanation and Social Policy: The problem of social inequalities in Health', *Journal of Social Policy,* II, 1 (7-32), Jan., 1982.

BOADWAY, R. *Public Sector Economics,* Winthrop Publishers, Massachusetts, 1979.

BOLTHO, A. 'Course and Causes of Collective Consumption Trends in the West', in R.C.O. MATTHEWS & G. B. STAFFORD (joint eds.) *The Grants Economy and Collective Consumption*, International Economic Association London, 1982.

BROWN, L. D. 'Competition and Health Cost Containment: Cautions and Conjectures', Milbank Memorial Fund Quarterly, *Health and Society*, vol 59, no. 2, 1981.

CHASSIN, M. R. 'The Containment of Hospital Costs: A strategic Assessment', *Medical Care,* vol 16, 10, supplement, 1978.

CONFEDERACION ESPAÑOLA DE CAJAS DE AHORRO, 'Seguridad Social' (Tema monográfico), *Papeles de Economía Española,* Madrid, 1982.

CULYER, A. J., MAYNARD, A. & WILLIAMS, A. 'Alternative Systems of Health Care Provision: An essay on motes and beams', in M. OLSON (ed.) *A New Approach to the Economics of Health Care*, AEI, 1981, Washington.

DAHLBERG L. & JAKOBSON V. 'On the effects of different patterns of Public Consumption Expenditure', Review of Income and Wealth, 1977.

DHSS (Department of Health and Social Security), Black Report: Working Group on Inequalities in Health, HMSO, London, 1980.

DEMPSTER, M. & WILDAVSKY, A. 'Modelling the US Federal Spending Process: Overview and implications'. in R.C.O. MATTHEWS & G. B. STAFFORD (joint eds.), The Grants Economy and Collective Consumption, (op.cit)

EUROPEAN ECONOMIC COMMUNITY, Organization, financing and cost of health care in the EEC, EEC, Social Policy Series, 1979. (German, France, U.K., Netherlands).

FOSTER, R. W. 'Cost based reimbursement and prospective payment: Reassessing the incentives', Journal of Health Politics, Policy and Law, 1982.

FUCHS, V. R. 'The output of the health industry', in M. H. COOPER & A. J. CULYER in Health Economics, Penguin, London, 1973.

FUCHS, V. R. 'The economics of health in a post-industrial society', The Public Interest, no. 56, Summer 1979, (pp. 3-20).

FUCHS, V. R. The Service Economy, National Bureau Economic Research, New York, 1968.

GALBRAITH, J. K. 'The Conservative Onslaught', The New York Review of Books, 22 Jan., 1981.

GEORGESCU-ROEGEN, N. 'The economics of production', American Economic Review, Papers and Proceedings, May, 1970.

GONZALEZ-FINAT A. & DIEZ-GONZALEZ A. 'Racionalización Presupuestaria: Algunas Reflexiones', Revista de Hacienda Pública Española, 1982.

HELLER, P. S. 'Diverging Trends in the shares of nominal and real government expenditure in GPD: implications for policy', National Tax Journal, vol XXXIV, 1980.

HOLMES, R. G. 'Control of Government Expenditure: Some institutional considerations (I)', International Monetary Fund, Fiscal Affairs Department, Occasional Paper, June 1982.

HUGHES, E. F. Hospital Cost Containment Programs. A policy Analysis, Bellinger Publishing, Cambridge, Massachusets, 1978.

JACKSON, P. M. 'The Public Expenditure Cuts: Rationale and Consequences', Fiscal Studies, 1982.

JOSKOW, P. J. 'Alternative Regulatory Mechanisms for controlling Hospital Costs', in M. OLSON (ed) A new Approach to the Economics of Health Care, AEI, Washington, 1981 (op. cit.).

JUDGE, K. (ed.) Pricing the Social Services, Martin Robertson, Oxford, 1981.

KLEIN, R. & BUXTON,M. 'Allocating Health Resources. A commentary on the Report of the R A W P', Research Paper no. 3, London, HMSO, August 1978.

MARTIN LOPEZ, E. 'Pluriempleo y conciencia de paro en la profesión médica', Revista Española de la Seguridad Social, 1982.

MAYNARD, A. & LUDBROOK, A. 'Budget Allocation in the National Health Service', York Reprinted Series, University of York, 1979.

MAYNARD, A. & LUDBROOK, A. 'Inequality, the NHS and Health Policy', York University Discussion Paper, July, 1981.

MINISTERIO de TRABAJO & I.N.P. Investigación sobre la asistencia farmacèutica en España, Madrid, 1977.

MINISTERIO de SANIDAD y SEGURIDAD SOCIAL, Necesidades sanitarias y Recursos asistenciales. Instituto Nacional de la Salud.MADRID, 1979.

MINISTERIO de ECONOMIA - INE,Estadísticas de establecimientos sanitarios en régimen de internado, 1976. Atalaya, Madrid, 1981.

MINISTERIO de SANIDAD y CONSUMO INSTITUTO NACIONAL de la SALUD, La farmacia hospitalaria en 1980, Coloquios de Farmacia Hospitalaria,Siguenza, Nov. 1980, Madrid, 1982.

MOONEY, G. H., RUSSELL, E. M. & WEIR, R. D. Choices for Health Care, The Macmillan Press Ltd., London 1980.

MULLER, C. F. 'Hospital Cost Containment and the Health Economy', Health Care Management Review, vol 6, 2, 1981.

OECD, 'The rise in Public Expenditure', The Observer, OECD, 1979.

OECD, Direct job creation in the Public Sector, mimeo, OECD, Paris, 1980.

OLSON , M. (ed.) A New Approach to the Economics of Health Care, A.E.I. Washington, 1981.

ORZECHOWSKI, W. P. 'Labour intensity, productivity and the growth of the Federal Sector', Public Choice, 1974.

PAULY, M. V. Medical Care at Public Expense, Praeger Publishers, London, 1972.

PEACOCK, A. 'Public X-inefficiency: Informational and institutional constraints', U.C.B., (Univeristy College of Buckingham), Discussion Papers in Economics, no. 2, 1981.

'Reducing Government Expenditure Growth: A British View', U.C.B. (University College of Buckingham) Discussion Paper no. 15, 1982.

PEACOCK, A. T. & WISEMAN, J. 'Approaches to the analysis of Government Growth', Public Finance Quarterly, 1979.

PESKING, H. M. & PESKIN, J. 'The valuation of 'non marketed' activities in income accounting', Review of Income and Wealth, 1979.

PLUTA, J. E. 'Real Public Sector Growth and decline in Developing Countries', Public Finance no. 3, 1981.

POYNTON, D., ALLAN, R. & RIDLEY, M. 'Health Service Feature: Managing in a time of Financial Restraint', Public Finance and Accountancy, October, 1983.

PRICE, R. W. 'Public expenditure: Policy and Control', National Institute Economic Review, vol 90, Nov. 1979.

ROYAL COMMISSION on the NATIONAL HEALTH SERVICE, Management of Financial Resources in the National Health Service, London H.M.S.O., 1978.

SELDON, A. 'Microeconomic controls - Disciplining by price', in The Taming of government, Institute of Economic Affairs, 1979.

SHANNON, R. 'Incentives for containment of Health care costs in OECD memeber countries. Theoretical and empirical issues', mimeo, Newcastle University, Secretariate of OECD, 1982.

SKOLKA, J. V. 'The substitution of self-service activities for marketed services', Review of Income and Wealth, 1976.

SKOLKA, J. V. 'Unbalanced Productivity growth and the growth of Public Services', Journal of Public Economics, 7, 1977.

TAIT, A. A. & HELLER, P. S. 'International Comparisons of Government Expenditure', International Monetary Fund, Occasional Paper no. 10, Washington D. C.

TAYLOR, C. T. & GODLEY, W. A. H. 'Prinicples and Problems in the measurement of Public Expenditure', University of Cambridge, Discussion Paper, 1979.

TOWNSEND, P. & DAVIDSON, N. Inequalities in Health (The Black Report), Penguin Book, London, 1982.

VAN DEN BELD, C. A. 'The Growth of Employment in the Public Sector and in the Private Sector', mimeo (paper presented to the Expert Conference on the structural Determinants of Employment and Unemployment OECD, 1977).

WATSON, W. G. 'Bacon and Eltis on Growth, Government and Welfare', Journal of Comparative Economics, Vol 2, 1, March, 1978.

WORLD HEALTH ORGANIZATION, 'Control of Health Care Costs in Social Security systems', Report on a Workshop - Regional office for Europe, 1981 - no. 55, Copenhagen.

WORTHINGTON, P. N. 'Capital-Labour ratios in short-term Voluntary Hospitals', Inquiry, vol XI, June, 1974.

WYSZEWIANSKI, L., WHEELER, J. R. & DONABEDIAN, A. 'Market oriented Cost-containment Strategies and Quality of Care', Milbank Memorial Fund Quarterly, Health and Society, vol 60, 4, Fall 1982.

YUSTE-GRIJALBA, J. (ed.) Apuntes para un libro negro de la Seguridad Social, Akal edit, Madrid, 1982.

REFERENCES: CHAPTER II - 1st PART (2.1 - 2.3)

AKEHURST, R. L. & HAYCOK, A. R., 'Factor Inputs and Hospital Outputs in the Constituent Authorities of North Western Regional Health Authority', mimeo, Paper presented to the Meeting of the SSRC Health Economists' Study Group, Glasgow, January, 1982.

ASHFORD, D. E., BERNE, R. & SCHRAMM, R. 'The expenditure-financing decision in British Local Government', Policy and Politics, no. 5, 1976.

BARR, J. L. & DAVIS, O. A. 'An elementary Political and Economic theory of the Expenditure of Local Governments',The Southern Economic Journal, no. 2, Oct. 1966.

BAUMOL, W. J. 'Macroeconomics of Unbalanced Growth: The anatomy of urban crisis', American Economic Review, no. 57, June 1967.

BAUMOL, W. J. 'Urban Services: Interactions of Public and Private decisions', in Public Expenditure Decisions in the Urban Community, H. G. SCHALLER (ed.), John Hopkins Press, 1963.

BECK, M. 'Public Sector Growth: A Real Perspective', Public Finance, vol XXXVI, no. 3, 1979.

BERGSTROM, T. C. & GOODMAN, R. P. 'Private Demands for Public goods', American Economic Review, vol 63, no. 3, June 1973.

BORCHERDING, T. E. & DEACON, R. T. 'The Demand for the Services of Non-Federal Governments', American Economic Review, 1973.

BORCHERDING, T. E. (ed.) Budgets and Bureaucrats, Durham, Duke University, 1975.

BORCHERDING, T. E., BUSH, W. C. & SPANN, R. M. 'The effects on Public Spending of the Divisibility of Public Outputs in Consumption, Bureaucratic Power, and the Size of the Tax-sharing group', in T. E. Borcherding, Budgets and Bureaucrats, (op. cit).

BOULDING, K. E. 'Pathologies of the Public Grants Economy' in R.C.O. Matthews & G.B. Stafford, The Grants Economy and Collective Consumption (op. cit).

BRADFORD, D. F. & OATES, W. E. 'The analysis of revenue sharing in a new approach to collective fiscal decisions', Quarterly Journal of Economics, no. 85, 1971.

BRETON, A. & WINTROBE, R. The logic of bureaucratic conduct. An economic analysis of competition, exchange and efficiency in private and public organizations, Cambridge University Press, 1982.

BRUECKNER, J. K. 'A test for allocative efficiency in the local public sector', Journal of Public Economics, 19, 1982.

BUSH, W. C. & MACKAY, R. J. 'Private versus Public Sector Growth: A Collective Choice Approach', in T. E. BORCHERDING (ed.), Budgets and Bureaucrats (op. cit).

COCHRANE, A. L. Effectiveness and Efficiency, Nuffield Provincial Hospitals Trust, 1971.

COURANT, P. N., GRAMLICH, E. M. & RUBINFELD, D. L. 'The stimulative effects of intergovernmental grants: or why money sticks where it hits', in Fiscal Federalism and Grants-in-aid, P. MIESZKOWSKI & W. H. OAKLAND, Compe Papers on Public Economics 1, Urban Institute, 1979.

CULYER, A. J. 'Health Care and the Market: A British Lament', Journal of Health Economics, no. 3, 1982.

CULYER, A. J. 'Public or Private Health Services? A skeptic's view', Journal of Policy Analysis and Management, vol 2, 3, Spring, 1983.

CULYER, A. J. (ed) 'Health and Health Indicators: Proceedings of a European Workshop', 1981.

CULYER, A. J., MAYNARD, A. & WILLIAMS, A. 'Alternative Systems of Health Care Provision: An Essay on Motes and Beams', in M. OLSON (ed). A new approach to the economics of Health Care, A. E. I. Washington, 1981.

CUTHBERTSON, K. FOREMAN-PECK, J. S. & GRIPAIOS, P. 'A model of local authority fiscal behaviour', Public Finance, no. 2, 1981.

CYERT, R. M. & HEDRICK, C. L. 'Theory of the Firm: Past, present, and future; An Interpretation', Journal of Economic Literature, 1973.

DEATON, A. & MUELLBAUER, J. Economics and consumer behaviour, Cambridge University Press, 1980.

DENISON, E. F. 'The shift to services and the rate of productivity change', The Brookings Institution, Washington D.C., 1972.

DENISON, E. F.'Explanations of declining productivity growth' Brooking General Services, Reprint 354, 1979, Washington.

DENNY, M. & FUSS, M. 'Intertemporal and interspatial comparisons of cost-efficiency and productivity', University of Toronto, Working Papers, nos. 8202,8131,8018, (1980-82).

DENZAU, A. T. & MACKAY, R. J. 'Benefit Shares and Majority Voting', American Economic Review, vol 66, 1, March 1976.

DILLON, R. D. Zero-base budgeting for Health Care Institutions, Aspen, London 1979.

FERGUSON, C. E. The Neoclassical Theory of Production and Distribution, Cambridge University Press, London, 1969.

FOTHERGILL, S. & GUDGIN, G. Unequal growth. Urban and Regional Employment Change in the U.K. Heinemann Educational Books, 1982.

FREY, B. S. 'The Public Choice Approach to the explanation of Collective Consumption', in R.C.O. MATTHEWS & C. B. STAFFORD, The Grants Economy and Collective Consumption, (op. cit).

GINZBERG, E. 'The Competitive solution: Two views', The New England Journal of Medicine, Nov. 6, 1980.

FUCHS, V. R. 'The Output of the Health Industry' in M.H. COOPER & A. J. CULYER (eds) 'Health Economics' Penguin Books, 1973.

GROSSKOPH, S. & HAYES, K. 'Do local Governments maximize anything?' Public Finance Quarterly, vol 11, no. 2, April 1983.

GRAMLICH, E. M. 'State and Local Governments and their Budget Constraint', International Economic Review, vol 10, 2, June 1969.

GRAMLICH, E. M. 'Intergovernmental Grants: A Review of the Empirical Literature', in W. E. OATES, The Policitcal Economy of Fiscal Federalism, Lexington Books, 1977.

GRAMLICH, E. M. & CALPER, H. 'State and Local fiscal behaviour and Federal Grant Policy', Brookings Papers on Economic Activity, 1973.

GREENWOOD, R., HININGS, C. R. & RANSON, S. 'The politics of the budgetary process in the English Local Government', Political Studies, vol 25, 1977.

McGUIRE, M. 'Notes on Grants-in-aid and economic interactions among Governments', Canadian Journal of Economics, VI, 2, 1973.

McGUIRE, M. 'A method for estimating the effect of a subsidy on the receiver's resource contraint: with an application to U.S. Local Governments 1964-71', Journal of Public Economics, no. 10, 1978.

HARRIS, R.J.D. 'Estimates of inter-regional differences in production in the U.K., Oxford Bulletin of Economics and Statistics, vol 44, no. 3, 1982.

HARSTAD, R. M. 'Behavioural explanations of efficienct public good allocations', Journal of Public Economics, 19, 1982.

HEALD, D. Public Expenditure. Its Defence and Reform, Martin Robertson, Oxford, 1983.

HENDERSON, P. D. 'Political and Budgetary constraints: Some characteristics and implications', in J. MARGOLIS (ed.), Public Economics, McMillan, 1970.

HODGSON, G. M. 'Theoretical and policy implications of variable productivity', Cambridge Journal of Economics, no. 6, 1982.

HUTTON, J. P. 'Measuring efficiency in Government Services', in A. PEACOCK, Quantitative Analysis in Public Finance, 1969.

INMANN, R. P. 'Testing political economy's "as if" proposition: is the median income voter really decisive', Public Choice, 33, 1978.

JACKSON, P. M. _The Political Economy of Bureaucracy,_ Philip Allan publisher, Oxford, 1982.

JACKSON, P. M. & ULPH, D. T. 'The relative prices of public sector and private sector goods', _Discussion Paper,_ no. 15 in Economics, Finance and Investment. _University of Stirling_, March 1973.

JHUN, V. J. & YOO, J. H. 'The public good demand function: Tax share as a dependent variable', _Public Finance Quarterly_, vol 6,3, July, 1978.

JORGENSON, D. W. & GRILICHES, Z. 'The explanation of Productivity change', _Review of Economic Studies_, Survey, part II, May 1969.

JORGENSON, D. W., GRILICHES, Z & DENISON, E. F. 'The Measurement of Productivity', _The Brookings Institution_, Washington D.C. 1972.

JORGENSON, D. W. & NISHIMIZU, M. 'U.S. and Japanese economic growth 1952-74: An international comparison' , _The Economic Journal_, 88 Dec. 1978.

KALDOR, N. _Causes of the slow rate of growth of the United Kingdom. An inaugural lecture,_ Cambridge University Press, 1966.

KRAVIS, I. B. 'A survey of International Comparisons of productivity', _The Economic Journal_, 86, March 1976.

LANCASTER, K.'The measurement of changes in quality', _Review of Income and Wealth,_ (page 157), 1977.

LAWLER, E. E. 'Job design and employee motivation', _Personnel Psychology,_ vol 22, 4, 1969.

LAYARD, P.R.G. & WALTERS, A. A. _Micro-economic Theory,_ McGraw-Hill Book Company, New York, 1978.

LEIBENSTEIN, H. 'Allocative Efficiency versus-X-Efficiency', _American Economic Review_, 1966.

LEIBENSTEIN, H. _Beyond the Economic man,_ Cambridge University Press, 1975.

LEIBENSTEIN, H. 'X-inefficiency Xists - Reply to an Xorcist', _American Economic Review,_ vol 68, 1, March 1978.

LEIBENSTEIN, H. 'On the basic proposition of X-efficiency theory', _American Economic Review_, vol 68, 2, 1978.

LEIBENSTEIN, H. 'A branch of Economics is missing: Micro-micro theory', Journal of Economic Literature, vol XVII, June 1979.

LEIBENSTEIN, H. 'Notes on X-efficiency and Bureaucracy', in R.C.O. MATTHEWS & G. B. STAFFORD, The Grants Economy and Collective Consumption (op. cit).

LE GRAND, J. 'Fiscal Equity and Grants to Local Authorities', The Economic Journal, Sept, 1975. (pag. 531-47).

LINDSAY, C. M. 'A Theory of Government Enterprise', Journal of Political Economy, vol 84, 5, 1976.

LIPSCOMB, J. 'The Willingness-to-pay criterion and Public Program Evaluation in Health', in S. J. MUSHKIN & D. W. DUNLOP (eds.) Health: What is it worth? Measures of health benefits, Pergamon Press, Washington, 1977.

LOASBY, B. 'Comment to "Beyond Economic Man" of H. Leibenstein', The Economic Journal, Dec. 1976.

LOPEZ-CASASNOVAS, G. 'Un test sobre las proposiciones de W. Baumol acerca del "Productivity Lag" en el Sector Público', Revista de Hacienda Pública Española, Ministerio de Hacienda, Instituto de Estudios Fiscales, no. 72, 1981.

LUSK E. J. & LUSK J. G. Financial and Managerial Control: A Health Care Perspective, Aspen, 1979.

LYNCH, B. & PERLMAN, M. 'Predicting Local Authority Balances: can it be done?' Centre for Enviromental Studies, 1980.

MARRIS, R. & MUELLER, D. C. 'The Corporation, Competition and the invisible hand', Journal of Economic Literature, vol XVIII, March 1980.

MARTIN, J. P. 'X-inefficiency, managerial effort and protection', Economica, no. 45, 1978.

MATTHEWS, R. C. O. & STAFFORD, G. B. (joint eds.) The Grants Economy and Collective Consumption, International Economic Association, London, 1982.

MILWARD, H. B. & RAINEY, H. G. 'Don't Blame the Bureaucracy!', Journal of Public Policy, vol 3, 2, 1983.

MOONEY, G., RUSSELL, E. & WEIR, R. D. Choices for Health Care, McMillan Press, London, 1980.

MUELLER, D. C. Public Choice, Cambridge University Press, 1979.

NADIRI, M. I. 'Some approaches to the theory and measurement of the Total Factor Productivity: A Survey', Journal of Economic Literature, vol 8, 1970.

NATIONAL INSTITUTE ECONOMIC REVIEW, 'Productivity comparisons', N.I.E.R., no. 101, August, 1982.

NISHIMIZU, M. & PAGE Jr. J. M. 'Total factor productivity growth, Technological Papers and Technical efficiency: Dimensions of Productivity change in Yugoslavia (1965-78)', The Economic Journal, 92, Dec. 1982.

NISKANEN, W. A. Bureaucracy and Representative Government, Aldine-Atherton, Chicago, 1971.

OATES, G. 'Financial Management in Central Government', Public Finance and Accountancy, Oct. 1983.

OLSON, M. The Logic of Collective Action. Public Goods and the Theory of Groups, Harvard University Press, Cambridge, Massachusetts, 1965, Ninth Printing 1982 .

POMMEREHNE, W. W. 'Institutional Approaches to Public Expenditure. Empirical evidence from Swiss municipalities', Journal of Public Economics, 9, 1978.

PREMCHAND, A. 'Government Budgetary Reforms: An Overview', Departmental Memoranda, International Monetary Fund, Fiscal Affairs Department, April, 1981.

PRIMEAUX, W. J. 'An assessment of X-inefficiency gained through competition', Review of Economic and Statistics, no. 59, 1977.

REINHARDT, V. Physician Productivity and the Demand for Health Manpower, Ballinger, 1975.

REDWOOD, J. A. 'Local Government Finances', National Westminster Bank Quarterly Review, February 1977.

ROMER, T. & ROSENTHAL, H. 'Political resource allocation, controlled agendas, and the status quo', Public Choice, no. 33, 1978.

ROMER, T. & ROSSENTHAL, H. 'The elusive median voter', Journal of Public Economics, 12, 1979.

ROSS, J. P. & BURKEHEAD, J. Productivity in the Local Government Sector, Lexington, Massachusetts, D. C., Heath, 1974.

ROSSER, R. 'Issues of Measurement in the Design of Health Indicators: A Review' in CULYER, A.J. 'Health and Health Indicators' (op cit). 1981.

ROTHBERG, D. L. Regional Variations in Hospital Use, Lexington Books, Massachusetts, 1982.

SCHOTT, K. 'The Economic behaviour of local authorities: An econometric examination of their actions in the relatively economically unconstrained period 1948-74', Fiscal Studies,no. 18, March, 1982.

SHEN, T. Y. 'Technology diffusion, substitution and X-efficiency', Econometrica, vol 41, no. 2, 1973.

SILKMAN, R. & YOUNG, D. R. 'X-Efficiency and state formula grants', National Tax Journal, vol XXXV, no. 3, 1982.

SMITH, D. L. 'Response to Federal Grants of State and Local Governments', National Tax Journal, vol XX1, 1968.

SMITH, G. 'NHS performance indicators', Public Finance and Accountancy, April, 1983.

SPANN, R. M. 'Rates of Productivity change and the growth of State and Local Government expenditures', in T.E. BORCHERDING, "Budgets and Bureaucrats", (op. cit).

SPANN, R. M. 'The Macroeconomics of unbalanced growth and the expanding public sector', Journal of Public Economics, 8, 1977.

STACEY, M. 'Concepts of Health and Illness: A working Paper on the Concepts and their relevance for research in SSRC' in 'Health and Health Policy. Priorities for Research', June 1977.

STAFFORD, G. B. The End of Economic Growth? (Growth and Decline in the U.K. since 1945), Martin Robertson, Oxford, 1981.

STIGLER, G. J. 'Director's Law of Public Income Distribution', Journal of Law and Economics, April, 1970.

STIGLER, G. J. 'The Xistence of X-efficiency', American Economic Review, vol 66, March 1976.

STRAUSS, R. P. 'The impact of block grants on local expenditure and property tax rates', Journal of Public Economics, 3, 1974.

TOWNSEND, H. Price Theory, Penguin Book Ltd., London 1971.

TWADDLE, A. C. 'The Concept of Health Status', Social Science & Medicine. Penguin Press 1974, vol 8 (pg. 29-38).

Van VINDEN, F. & Van PRAAG, B. 'A dynamic model of the interaction between the State and the private sector', Journal of Public Economics, vol 16, no. 1, August 1981.

WADLEY, J. 'Research on Health Manpower productivity: A general overview', in J. RAFFERTY (ed.) Health Manpower and Productivity, Heath, 1974.

WAGNER, R. E. 'Revenue Structure, Fiscal Illusion and Budgetary Choice', Public Choice, vol XXV, Spring, 1976.

WILLIAMS, A. 'El rol económico de los "Indicadores de salud", in Primeres Jornades d'Economia dels Serveis Socials' de Barcelona I.E.S. 1983.

WILLIAMS, A. & ANDERSON, R. Efficiency in the Social Services. Blackwell, Oxford, 1975.

WRIGHT, M. (ed.) Public Spending Decisions, Allen and Unwin, London 1980.

REFERENCES: CHAPTER II - 2nd PART (2.4 - 2.6)

BAROCCI, T. A. Non Profit Hospitals. Auburn House, MIT, 1981.

BAYS, C. W. 'Case-mix differences between non-profit and for profit Hospitals' Inquiry, vol 1, March 1977.

BONGAERTS, J. C. & CARRIN, G. 'A Model of Hospital manager behaviour', Studiecentrum yoor Economisch en Social onderzoek, Rapport 82/124, April 1982.

CLARKSON, K. W., 'Some implications of property rights in Hospital Management', The Journal of Law and Economics, vol XV, 2, Oct. 1972.

DAVIS, R. 'Economic Theories of behaviour in a non profit private hospital', Economic and Business Bulletin, vol 24, 2, 1972.

DUSANSKY, R. & KALMAN, P. J. 'Toward an economic model of the teaching hospital', Journal of Economic Theory, 7, 1974.

EVANS, R. G. 'Modelling the Economic Objectives of the Physician', in R. D. FRASER (ed.) Health Economics Symposium, Proceedings of the first Canadian Conference, Sept. 4,5,6, 1974.

FELDSTEIN, M. 'The Rising Price of Physician's Services', Review of Economics and Statistics 52, May 1970.

FRISS, L. & WHITE, M. 'Productivity in Nursing', Topics in Health Care Financing, vol 4, 2, Winter 1977.

GOLDFARB, M. HORNBROOK, M.H. & RAFFERTY, J. 'Behaviour of the Multi-product Firm. A Model of the non-profit Hospital System'. Medical Care February 1980, vol XVIII, 2.

GOSS, M. E. 'Patterns of Bureaucracy among Hospital Staff Physicians' in E. FREIDSON (ed.) The Hospital in Modern Society, McMillan, 1963.

HARRIS, J. E. 'The internal organization of hospitals: some economic implications', The Bell Journal of Economics, Aut. 1977.

JACOBS, P. 'A Survey of Economic Models of Hospitals', Inquiry, vol XI, June 1974.

LAUNOIS, R. J. 'La theorie de la Bureaucratie a L'Hopital', Communication for Colloque INSERM, Paris, Oct. 1981.

LEE, M. L. 'A conspicuous production theory of hospital behaviour', Southern Economic Journal, 38, July 1971.

MACCHIA, A. F. 'An economic theory of the hospital and industrial organization of the hospital market: A study of the economic modelling of decision process, internal organization behaviour and strategic interaction', Disertation in Public Policy Analysis, DPhil University of Pennsylvania, 1979, Microfilm.

MIGUÉ, J. L. & BÉLANGER, G. The Price of Health,McMillan, Canada, 1974.

'Toward a general Theory of Managerial Discretion', Public Choice, vol XVII, Spring 1977.

MURRAY, J. A. 'The behaviour of physicians, key to the Health Care Sector', The Journal of Behavioural Economics, vol III, no. 1, 1974, (pages 245-272).

NEWHOUSE, J. 'Toward a theory of non-profit institutions. An Economic model of a Hospital', American Economic Review, no. 60, 1970.

PAULY, M. V. 'The behaviour of non profit hospital monopolies: Alternative models of the Hospital', in C. C. HAVIGHURST, Regulating Health facilities constructions, A. E. I., 1974.

PAULY, M. V. 'Doctors and their Workshops. Economic models of Physician behaviour', University of Chicago Press, 1980.

PAULY, M. V. 'Hospital characteristics and physician productivity, and fee levels', in R. M. SCHEFFLER & L. F. ROSSITER, Advances in Health Economics and Health Services Research, Jai Press, Inc, 1982.

PAULY, M. & REDISCH 'The not for profit hospital as a Physician's cooperative', American Economic Review, 63, March 1973.

PEŔEZ-DÍAZ, V. 'Médicos, Administradores y Enfermos: La calidad de la Asistencia Sanitaria', in Confederación Española de Cajas de Ahorro, Papeles de Economía Española, no. 12/13, 1982.

SALTMAN, R. G. & YOUNG, D. W. 'The hospital power equilibrium: An alternative view of the cost containment dilemma', Journal of Health Politics, Policy and Law, vol 6, no. 3, Fall, 1981.

SLOAN, F. A. 'A microanalysis of physicians' hours of work-decisions' in M. PERLMAN (ed.) The Economics of Health and Medical Care, I.E.A. (International Economic Association), London McMillan, 1974.

SPICER, M. W. 'The economics of the bureaucracy in the British NHS', Milbank Memorial Fund Quarterly, Health and Society, vol 60, 4, Fall 1982.

STEVENS, C. M. 'Hospital Market Efficiency: The anatomy of the supply response', in H. E. KLARMAN (ed.), Empirical Studies in Health Economics, John Hopkins Press, 1970.

REFERENCES: CHAPTER III

ALCHIAN, A. A. & DEMSETZ, H. 'Production, Information Costs and Economic Organization', American Economic Review, vol 61, 1972.

ANTHONY, R. N. & HERZLINGER, R. E. Management Control in Non-profit Organizations, R.D. IRWIN, 1975.

ATKINSON, A. B. & STIGLITZ, J. E. Lectures on Public Economics, McGraw-Hill Book Company (U.K.), 1980.

BAUERSCHMIDT, A. D. et al, 'Changing Health Care Delivery through a prospective-incentive reimbursement system', Topics in Financing Health Care, 1981.

BERGSON, A. 'Managerial risks and rewards in public bureaucracies', Journal of Comparative Economics, vol 2, nº 3, Sept. 1978, (pp. 211-226).

BREEN, W. 'Financial Management and Third Party Reimbursement: An overview', Topics in Health Care Financing, vol 3, 1, Fall, 1976.

CHIANG, A. C. Fundamental Methods of Mathematical Economics, McGraw-Hill, 1967.

CODDINGTON, A. Theories of the bargaining process, Allen and Unwin, 1968.

CYERT, R. M. The Management of Non-profit Organizations, with Emphasis on Universities, Lexington Books, 1975.

DEMSKI, J. S. & FELTHAM, G. A. 'Economic Incentives in Budgetary Control Systems', The Accounting Review, vol LIII, no. 2, April 1978.

DHSS (Department of Health and Social Security), Resource Allocation Working Party (for England), DHSS, H.M.S.O., 1976.

DOUGLAS, J. 'Towards a rationale for private non-profit organizations: A Review of current theory', Discussion Paper, I S P S, Yale University, 1980.

EVANS, R. G. 'Efficiency Incentives for Hospital Reimbursement', Harvard University, Massachusetts, Cambridge 1970, Dissertation for the degree of Ph.D.,Microfilm.

FAMA, E. F. 'Agency problems and the theory of the firm', Journal of Political Economy, vol 88, 2, 1980.

FAN, L. S. 'On the Reward System', American Economic Review, vol 65, no. 1, 1975.

GERWIN, D. 'Towards a theory of public budgetary decision-making', in R. BYRNE et al (eds.) Studies in Budgeting, North Holland publishers, 1971.

GLASER, W. A. 'Paying the Hospital: Foreign Lessons for the United States', Health Care Financing Review, Summer, vol, 4, 1983.

GREEN, J. R. & LAFFONT, J. J. Incentives in Public Decision-making, North Holland, 1979.

HARRIS, M. & RAVIV, A. 'Some results on incentive contracts with application to Education and Employment, Health Insurance and Law enforcement', American Economic Review, vol 68, 1, 1978.

HERZBERG, F. 'One more time: How do you motivate employees?', Harvard Business Review, 46:1, Jan.-Feb.1968.

HODGSON, G. M. 'Worker participation and Macroeconomic efficiency', Journal of Post-Keynesian Economics, vol 5, 2 1982-83 (Winter).

HOENACK, S. A. Economic Behaviour Within Organizations, Cambridge University Press, 1983.

HURWICZ, L. 'The design of mechanisms for resource allocation', American Economic Review, Papers and Proceedings, May 1973.

JONSSON, R. 'Recurrent Cost and Performance Comparisons within Public Administration: A method of generating efficiency-promoting competition', Public Choice, vol 42, no. 21, 1984.

LAFFONT, J. J. & MASKIN, E. 'The theory of incentives: An overview', Economic Theory Discussion Paper no. 41, Cambridge University, 1981.

LANG WELL, K. M. & WERNER, J. L. 'Economic Incentives in Health manpower Policy', Medical Care, vol XVIII, 11, 1980.

LAZEAR, E.P. 'Agency, earning profiles, productivity and hours restrictions, American Economic Review, 71, Sept. 1981.

LEIBENSTEIN, H. Beyond the Economic Man, Cambridge University Press, 1975.

LEVY-FERDINAND, K. 'Economic analysis of the non-profit-institution. The case of the private University', Public Choice, vol IV, Spring 1968.

MALCOMSON, J. M. 'Work incentives, Hierarchy and Internal labour markets', University of York, Discussion Paper no. 78, 1982.

MALCOMSON, J. M. 'Trade Unions and Economic Efficiency', The Economic Journal, Papers and Proceedings, 1983.

McCALL, J. J. 'The simple economics of incentive contracting', American Economic Review, Dec. 1970.

McCARTHY, C. 'Incentive Reimbursement as an impetus to cost-containment', Inquiry, vol XII, 4, 1975.

McLENNAN, A. 'A non cooperative definition of two person bargaining', Working Paper, University of Toronto, no. 8303, 1981.

MIRRLEES, J. A. 'The optimal structure of incentives and authority within an organization', Bell Journal of Economics, Spring 1976, 7 (1).

PAULY, M. V. 'Efficiency, Incentives and Reimbursement of Health Care', Inquiry, vol VIII, 1, 1970.

PONDY, L. R. 'Toward a Theory of internal resource-allocation', in M. N. ZALD (ed.) Power in organizations, Vanderbilt University Press, 1970.

PREMCHAND, A. 'Government Budgetary Reforms: An Overview', Occasional Papers, International Monetary Fund, Fiscal Affairs Department, April 1981.

PREMCHAND, A. 'Government and Public Enterprises: Budgetary Relationships', Occasional Papers, International Monetary Fund, Fiscal Affairs Department, IMF, August 1982.

RADNER, R. 'Monitoring cooperative agreements in a repeated principal-agent relationship', Econometrica, vol 49, 5, Sept. 1981.

RO, K. K. & AUSTER, R. 'An output approach to Incentive Reimbursement for Hospitals', Health Services Research, Fall, 1969.

ROSS, S. A. 'The Economic Theory of Agency: The Principal's Problem', American Economic Review, vol 63, 1973.

SHAVELL, S. 'Risk sharing and incentives in the principal and agent relationship', The Bell Journal of Economics, vol 10, 1979.

SCHULTZE, C. L. 'The role of incentives, penalties and rewards in attaining effective policy', in The analysis and evolution of Public Expenditures: the PPBS, Joint Committee, Congress of U.S.A., 1969.

VRACIU, R. A. 'Programming, Budgeting and Control in Health Care Organizations: The state of the art', Health Services Research, Summer 1979.

WEISBROD, B. A. 'Toward a theory of the voluntary non profit sector in a three sector economy', in E. PHELPS (ed.) Altruism, Morality and Economic Theory, New York, Russel Sage Foundation, 1975.

WEISBROD, B. A. The Voluntary Non Profit Sector, Lexington Books, Toronto, 1977.

WEITZMAN, M. L. 'The new Soviet incentive model', The Bell Journal of Economics, 1976.

WEITZMAN, M. L. 'Optimal Rewards for economic regulation', American Economic Review, Sept. 1980.

WELSH, W. A. 'Modeling Budgetary Strategies in Health Policy, East and West', Journal of Health Politics, Policy and Law, vol 8, no. 3, Fall 1983.

WILLIAMSON, O. E. 'The modern Corporation: origins, evolution, attributes', Journal of Economic Literature, vol XIX, Dec. 1981.

URI Ben ZION, & SPIEGEL, M. 'Efficiency and compensation of public officials', Public Choice, no. 35, 1980.

REFERENCES: CHAPTER IV

AKEHURST, R. L. & BLACKBURN,K. 'Geographic Cost Variations in the North Western Regional Health Authority: An examination of one of the assumptions of the DHSS Resource Allocation Working Party', Hospital and Health Services Review, November,1979.

ALLISON, R. F. 'Administrative Responses to Prospective Reimbursement', Topics in Health Care Financing, vol 3, 2, Winter, 1976.

BANTA, H. D. & RUSSELL, L. B. 'Policies toward medical technology: An international review', International Journal of Health Services, vol 11, no. 4, 1981.

BARER. M.L. 'Case-mix adjustment in Hospital Cost analysis: Information theory revisited', Journal of Health Economics, 1 1982 (pp. 53-80).

BARNARD, C. & ESMOND, P. 'DRG-Based Reimbursement: The use of concurrent and Retrospective Clinical Data', Medical Care, vol XIX, II Nov, 1981.

BECKER, E. R. & STEINWALD, B. 'Determinants of Hospital Case-mix Complexity', Health Services Research, 16:4, Winter 1981.

CULLIS, J. G. & WEST, P. A. The Economics of Health: An Introduction, Martin Robertson, 1979.

CULYER, A. J. 'Health and health indicators: Introduction', in A.J. CULYER (ed.) Health and Health Indicators: Proceedings of an European Workshop', Longman, 1981.

CULYER, A. J., LAVERS, R. J. & WILLIAMS, A. 'Social indicators: Health', Social Trends, 1971.

CULYER, A. J.,WISEMAN, J., DRUMMOND, M. F., & WEST, P. A. 'Joint costs and budgeting for English Teaching Hospitals', Institute of Social and Economic Research, University of York, School of Social Sciences, University of Sussex; mimeo, 1978.

BENNETT, J. E. & KRASNY, J. 'Functional Value Analysis: a technique for reducing hospital overhead costs', Topics in Health Care Financing , vol 3, 4, Summer, 1977.

BENTLEY, J. D. & BUTLER, P. W. 'Measurement of case mix', Topics in Health Care Financing, vol 8, 4, Summer, 1982.

BERKI, S. E. Hospital Economics, Lexington Books, 1972.

BERRY, R. E. 'Product Heterogeneity and Hospital Cost Analysis', Inquiry, vol VII, 1, 1970.

BERRY, R. E. 'On grouping hospitals for economic analysis', Inquiry, vol X, 5, Dec. 1973.

BERRY, R. E. 'Cost and efficiency in the production of Hospital Services', Milbank Memorial Fund Quarterly, Health and Society, Summer,1979.

CARR, W. J. & FELDSTEIN, P. J. 'The relationship of cost to hospital size', Inquiry, IV, 1967.

CES 'Unitary Grant: Bad arguments in a good cause', CES, (Centre for Enviroment Studies), no. 2, Dec. 1977.

COHEN, H. A. & ATKINSON, J. G. 'Case-mix regulation', Topics in Health Care Financing, vol 8, 4, Summer, 1982.

COOPER,R. A. & WEEKES, A. J. Data Models and Statistical Analysis , Philip Allan, Oxford, 1983.

CROMWELL, J. & KANAK, J. 'The effects of Prospective Reimbursement Programs on Hospital adoption and Service Sharing', Health Care Financing Review, vol 4, 2, Dec. 1982.

DANZON, P. M. 'Hospital 'Profits': The effects of Reimbursement Policies', Journal of Health Economics, 1, 1982, (pp. 29-52).

DAVIES, B. Social Needs and Resources in Local Services, London, 1968.

DAVIS, K., DEMPSTER, M. & WILDAVSKY, A. 'Towards a Predictive Theory of Government Expenditure', British Journal of Polictical Science, no. 4, 1974, (pag. 419-52).

DOE (Department of Enviroment), The New Block Grant 1981-82: Local Government, Planning and Land Act 1980.
Grant Related Expenditure: 'How the expenditure needs of Local Authorities are assessed in the new Block Grant', DOE, London, Dec. 1980.

DIEUDONNE, P. 'L'Etude economique de l'Hopital: Approache theorique et analyse empirique', Facultés Universitaires Notre Dame de la Paix, 1980, Memoire par Licencie et Maître en Sciences Economiques et Sociales.

DOWLING, W. L. 'Prospective Rate Setting: Concept and Practice', Topics in Health Care Financing, vol 3, 2, Winter 1976.

DUSANSKY, R. & KALMAN, P. J. 'Toward an economic model of the teaching hospital', Journal of Economic Theory, 7, 1974.

EASTWOOD, D. B. 'An adaptive linear expenditure system for state and local governments', Applied Economics, 1971.

ELNICKI, R. A. 'Hospital Productivity, Service Intensity and Costs', Health Services Research, Winter, 1974.

EVANS, R. G. 'Behavioural cost functions for hospitals', Canadian Journal of Economics, May, 1971, (pp. 198-215).

EVANS, R. G. & WALKER, H. D. 'Information Theory and the Analysis of Hospital Cost Structure', Canadian Journal of Economics, vol 5,3, 1972 (pp. 398-418).

FELDSTEIN, M. S. 'The effects of case-mix on Hospital Costs', Medical care, vol 3, 1965 (pp. 95-103).

FELDSTEIN, M. S. Economic analysis for Health Service Efficiency, Econometric studies of the British National Health Service, North-Holland Publishing Company, Amsterdam, 1967

FELDSTEIN, M. S. 'Econometric Studies of Health Economics', in M. D. INTRILIGATOR & D. A. HENDRICK (eds.) Frontiers of quantitative Economics (vol II), North Holland, 1974.

FELDSTEIN, P. L. 'The quality of hospital services', in PERLMAN (ed.) Economics of Health and Medical Care, (op. cit).

FETTER, R. B. et al. 'Case-mix definition by Diagnosis-Related Groups', (Center for Health Studies, Institution for Social and Policy Studies, Yale University), Medical Care, Supplement to vol 18, no. 2, Feb. 1980.

FOSTER, C. D. et al. 'The Reform of the Grant System', in Local Government Finance in a Unitary State, George Allen and Unwin, 1980.

GIBSON, J. B. 'The Block (and target) Grant System and Local Authority Expenditure: Theory and Evidence', Local Government Studies, vol 8, 3, May/June 1982.

GOLDFARB, M., HORNBROOK, M. H. & RAFFERTY, J. 'Behaviour of the Multiproduct firm: A model of the non-profit Hospital System', Medical Care, vol XVIII, 2, Feb. 1980.

GOODLEY, W. 'Grant Related Expenditure', Cambridge Economic Policy Review (C.E.P.R.) vol 6, no. 2, June 1980.

GREENBERG, J. & KROPF, R. 'A Case-mix method for Developing Health Planning Criteria for Hospital Services', Medical Care, vol XIX, 11, Nov. 1981.

GRIFFITH, J. R. et al. 'Measuring Community Hospital Services in Michigan', Health Services Research, 16:2, Summer, 1981.

HARRISON, T. F. 'The clinician as budget manager', Public Finance and Accountancy, Jan. 1982.

HEIGHAM, D. 'Grant related expenditure: an invisible push towards efficiency in local Government, Public Money, vol 2, 1, June 1982.

HENDERSON, J. M. & QUANDT,R. E. Microeconomic Theory: A Mathematical Approach , McGraw-Hill, Third Edition, 1980.

HERKIMER, A. G. 'Improving productivity', Topics in Health Care Financing, vol 4, 2, Winter 1977.

HJERPPE, R. T. 'The measurement of real output of public sector services', Review of Income and Wealth, 1980.

HORN, S. D. & SCHUMACHER, D. N. 'An analysis of Case-mix complexity using Information Theory and Diagnostic Related Grouping', Medical Care, vol XVII, no. 4, April 1979.

HORN, S. D. & SCHUMACHER, D. N. 'Comparing Classification Methods: Measurement of Variations in charges, LOS and Mortality', Medical Care, vol XX, 5, May 1982.

HORNBROOK, M. C. 'Hospital case-mix: its definition, measurement and use', Medical Care Review, vol 39, no. 1, 2, Spring 1982.

HORNBROOK, M. C. & RAFFERTY, J. 'The Economics of Hospital Reimbursement', in R. M. SCHEFFLER & L. F. ROSSITER (eds.) Advances in Health Economics and Health Services Research, J A S Press Inc, 1982.

INGBAR, M. L. & TAYLOR, L. D. Hospital Costs in Massachusetts, Harvard University Press, Cambridge, M.A. 1968.

JEFFERS, J. R. & SIEBERT, C. B. 'Measurement of Hospital Cost variations: Case-Mix,Service Intensity and Input Productivity Factors', Health Services Research, Winter 1979.

JOHNSON, K. W. 'The effect of the case-mix on LOS', North Western Regional Health Authority, mimeo, Research Report, no. 1 Regional Statistical Unit, August 1977. Manchester.

JOHNSTON, J. Econometric Methods, McGraw Hill, 1972.

JONES, R. & PENDLEBURY, M. 'Accountability and Budgets', Public Finance and Accountancy, 1983.

KIDDER, D. & SULLIVAN, D. 'Hospital Payroll Costs, Productivity and Employment under Prospective Reimbursement', Health Care Financing Review, vol 4, 2, Dec. 1982.

KLASTORIN, T. D. & WATTS, C. A. 'On the measurement of Hospital Case-mix', Medical Care, vol XVIII, 6, June 1980.

KLASTORIN, T. D., WATTS, C. A. & TRIVEDI, V.'A study of the Classification of Hospitals for prospective reimbursement', prepared by Health Care Financing Administration Report no. 10, DHEW, Washington University, Seattle U.S. Department of Commerce, National Technical Information Service, Jan. 1978, microfilm.

KMENTA, J. Elements of Econometrics, McMillan publish., 1971.

LAVE, J. R. & LAVE, L. B. 'Hospital Cost Functions', American Economic Review, 60, 1970.

LAVE, J. R. & LAVE, L. B. 'The extent of Role differentiation among Hospitals', Health Services Research, Spring 1971.

LAVE, J. R., LAVE, L. B. & SILVERMAN, L. P. 'Hospital Cost estimation controlling for case-mix', Applied Economics, 4, 1972.

LAVE, J. R., LAVE, L. B. & SILVERMAN, L. P. 'A proposal for incentive reimbursement for hospitals', Medical Care, vol XI, 2, March-April, 1973.

LAVERS R. J. & WHYNES, D. K. 'A Production Function Analysis of English Maternity Hospitals', ISER Reprint Series, 1977.

LEE, M. L. 'A Conspicuous Production Theory of Hospital Behaviour', Southern Economic Journal, no. 38, June 1971.

LEE, M. L. & WALLACE, R. L. 'Classification of Diseases for Hospital Analysis', Inquiry, vol 2, 1972, (pp. 69-72)

LEE, M. L. & WALLACE, R. L. 'Problems in estimating multiproduct functions: An application to hospitals', Western Economic Journal, 1973, (pag. 350-363).

LEE, M. L. & WALLACE, M. L.'Multiproduct cost functions', Western Economic Journal, vol XI, 1973.

LUKE, R. D. 'Dimensions in Hospital case-mix measurement', Inquiry, vol XVI, Spring 1979.

LYNCH, B. & PERLMAN, M. 'Predicting Local Authority Balances: can it be done?' Centre for Enviromental Studies, 1980.

MAGEE, CH. 'The potential for specialty costing in the NHS', Public Finance and Accountancy, March 1981.

MANN, J. K. & YETT, D. E. 'Hospital Costs', in M. H. COOPER & A. J. CULYER (eds.) Health Economics, Penguin Books, 1973. Australia.

McCLENDON, E. L. 'Health Care Cost Elements affecting Legislative and Planning Considerations', Topics in Health Care Financing, vol 4, 2, Winter, 1977.

McSTRAVIC,R.S. 'Hospital Occupancy Standards', Hospitals, Sept. 1982.

MILBANK MEMORIAL FUND, 'Organizational Determinants of Services, Quality and Cost of Care in Hospitals', Health and Society, vol 57, no. 2, 1979.

MILLS, R., RIEDEL, D. C., FETTER, R. & AVERILL, R. 'AUTOGRP: An interactive Computer System for the Analysis of Health Care Data', Medical Care, vol XIV, no. 7, July 1976.

MYERS, L. P. & SCHROEDER, S. A. 'Physician Use of Services for the Hospitalized Patient: A Review with Implications for Cost Containment', Health and Society, vol 59, no. 4, 1981.

MOORE, S. 'Cost Containment through size-sharing by primary care physicians', The New England Journal of Medicine, June 1979.

MUURINEN, J.-M. 'Economic Research into Health and Utilisation Behaviour: A theoretical retrospect and prospect', mimeo, Paper presented at the Scandanavian Health Economists' Meeting in Helsinki, Finland, 1982.

MUSKIN, S. J. 'Health Indexes for Health Assessments', in S. J. MUSHKIN & D. W. DUNLOP (eds.) Health: What is it Worth? Measures of Health Benefits, Pergamon Press, 1977.

PERLMAN, M. (ed.) The Economics of Health and Medical Care, London and Basingstoke: MacMillan Press, 1974.

PLESSAS, D. J. & CARPENTER, E. S. 'Empirical designation of Health Service Areas', Health Services Research, Winter 1975.

RAFFERTY, J. (ed.) Health Manpower and Productivity, Heath, 1974.

RECORD, J. C. et al. (Kaiser-Permanent Health Services Research Center) 'New Health Professions After a Decade and a Half: Delegation, Productivity and Costs in Primary Care', Journal of Health Politics, Policy and Law, Vol 5, no. 3, Fall, 1980.

REDISCH, M. A. 'Physician involvement in Hospital Decision-Making', in J. B. McKINLAY (ed.) Issues in Hospital Administration, no. 9.

REDDER, M. V. 'Productivity in the Medical Care Industry', in J. RAFFERTY, Health Manpower and Productivity, (op. cit).

REINHARDT, U. E. & SMITH, K. R. 'About manpower substitution (in Ambulatory Care)', in J. RAFFERTY, Health Manpower and Productivity, (op. cit).

RESTUCCIA, J. D. & MURPHY, D. V. 'Estimating Hospital Costs by Diagnosis Groups, Inquiry, vol XVII, Summer, 1980.

ROSS, J. P. & BURKHEAD, J. Productivity in the Local Government Sector, Lexington, Massachusetts, D. C., Heath., 1974.

RUCHLIN, H. S. 'Problems in measuring Institutional Productivity', Topics in Health Care Financing, vol 4, 2, Winter, 1977.

RUCHLIN, H. S. & LEVESON, I. 'Measuring Hospital Productivity', Health Services Research, Winter, 1974.

RUSSELL, L. B. Technology in Hospitals. Medical Advances and Their Diffusion, Brookings Studies in Social Economics, Massachusetts, 1979.

SHUMAN, L. J., WOLFE, H., WHETSELL, G. W. & HUBER, G. A., 'Prospective Reimbursement: Home Care', Inquiry, Sept. 1976.

SHUMAN, L. J., WOLFE, H. & HARDWICK, C. P. 'Predictive Hospital Reimbursement and Evaluation model', Inquiry, vol IX, 2, 1972.

SIMBORG, D. V. 'Sounding Board. DRG Creep. A new Hospital-acquired disease', The New England Journal of Medicine, June, 25, 1981.

SOULIER, M. Z. & DE COSTER, D. T. 'Productivity versus Cost control: Considerations for Health Care managers', Health Care Management Review, vol 7, no. 1,Winter 1982.

STACEY, M. 'Concepts of health and illness: A working paper on the concepts and their relevance for research in SSRC', Health and Health Policy, Priorities for Research, June 1977.

STERN, J. 'Social Mobility and the interpretation of Social Class Mortality differentials', Centre for Labour Economics, LSE, Discussion Paper, no. 23.

THOMPSON, J. D. 'One application of the DRG planning model', Topics in Health Care Financing, vol 8, 4, Summer, 1982.

THOMPSON, J. D., FETTER, R. B. & MROSS, C. D. 'Case-mix and Resource use', Inquiry, no. 12, 1975.

TRAVERS, T. 'Block Grant: origins, objects and use', Fiscal Studies, vol 3, 1, March, 1982.

TWADDLE, A. C. 'The concept of health status', Social Science and Medicine, vol 8 (pp. 29-39), Pergamon Press, U.K.

WEEDEN, R. Geographic Variations in the Cost of Health Service Inputs, mimeo, Government Economic Service, Working Paper, no. 35, October, 1980.

WEST, P. A. 'Allocation and equity in the public sector: The hospital revenue allocation formula', Applied Economics, vol 5, 1973.

WILLIAMS, A. 'Health Service Planning', in J. M. ARTIS & A. R. NOBAY, (eds.) Studies in Modern Economic Analysis, Basil Blackwell, 1977.

WILLIAMS, A. 'The nature, meaning and measurement of Health and illness: An economic viewpoint', mimeo, Health Economists Study Group and Medical Sociology Group, Joint Meeting, 1983.

WILLIAMS, S. U. et al. 'Improved cost allocation in Case-mix Accounting', Medical Care, vol XX, 5, May 1982.

WOLFSON, A. A Health index for Ontario, Toronto, Minister of Treasury, 1974.

WOOD, W. R., AMENT, R. P. & KOBRINSYI, E. J. 'A Foundation for Hospital Case-mix measurement', Inquiry, vol XVIII, Fall, 1981.

YATES, J. Hospital Beds, William Heinemann Medical Books, London, 1982.

YODER, J. L. & CONNOR, R. A. 'Diagnosis Related Groups and Management', Topics in Health Care Financing, Summer, 1982.

YOUNG, W. W., SWINKOLA, R. B. & HUTTON, M. A. 'Assessment of the AUTOGRP Patient Classification System', Medical Care, vol XVIII, no. 2, February 1980.

YOUNG, W. W., SWINKOLA, R. B. & ZORN, D. M. 'The Measurement of Hospital Case-mix', Medical Care, vol XX, 5, May, 1982.

ZARETSKY, H. W. 'The effects of patient mix and service mix on Hospital Costs and Productivity', Topics in Health Care Financing, vol 4, 2, Winter, 1977.

REFERENCES: CHAPTER V

AKEHURST, R. L. 'The effect of adjusting sub-regional target revenue allocation for case-mix within the NWRHA', The Hospital and Health Services Review, Oct. 1980.

AMENT, R. P. et al. 'Three Case-type Classifications: Suitability for use in Reimbursing Hospitals', Medical Care, vol XX, 5 May 1982.

ANDERSON, J. et al (eds.) Education in informatics of Health Personnel, JFIP medical informatics. Monograph Series vol 1, North Holland/American eslewier, 1974.

ASHBY, J. L. 'An Analysis of Hospital Costs by Cost Center 1971 through 1978', Health Care Financing Review, vol 4, 1, Sept. 1982.

ASHFORD, J., BAILEY, T. & BUTTS, M. 'Statistical alternatives to detailed clinical costing', Public Finance & Accountancy, March, 1981.

BARNETT, J. R., WARD, D. & TATCHELL, P. M. 'Hospital Resource Allocation in New Zealand', Reprint Series, 4, The Australian National University of Camberra.

BAUER, K. G. 'Information Needs for Hospital Rate Setting', Topics in Health Care Financing, vol 3, 2, Winter, 1976.

BERESFORD, J. C. & ELLIOT, G. C. 'Specialty costs at two general hospitals: Actual versus regresion approach - a comparative study', The Hospital and Health Services Review, Nov./Dec. 1982.

BERRY, R. E. 'Prospective Reimbursement: New York Formula', Inquiry, Sept. 1976.

BROADLEY, K., BROOK, S. & ROBERTS, J. 'District computer services in practice', The Hospital and Health Services Review, March, 1982.

COELEN, C. & SULLIVAN, D. 'An Analysis of the effects of Prospective Reimbursement Programs on Hospital Expenditures', Health Care Financing Review, Winter, 1981.

COLES, J., DAVISON, A. & WICKINGS, I. 'Allocating budgets to wards: An experiment', The Hospital and Health Services Review, Sept. 1976.

COLES, J., WICKINGS, I. & CARVELL, M. 'Improving variable cost budgeting in the NHS', Public Finance and Accountancy, April, 1982.

CONNOR, R. A. 'A management strategy for prospective case-based payment', Health Care Management Review, vol 6, 4, 1981.

COX, D. R. & SNELL, E. J. 'The Choice of Variables in Observational Studies', Applied Statistics, 23, no. 1 (p. 51-67) 1974.

CULLIS, J. G. & WEST, P. A. 'Comparing hospital costs. A reconsideration', The Hospital and Health Services Review, August, 1973.

CULYER, A. J., WISEMAN, J., DRUMMOND, M. F. & WEST, P. A. 'What Accounts for the Higher Costs of Teaching Hospitals', Social and Economic Administration, vol 12, no. 1, Spring 1978.

DAVISON, A. J. & GRACE, R. 'Management structures and computer policy', The Hospital and Health Services Review, Jan. 1982.

DRUMMOND, M. 'Welfare Economics and Cost Benefit Analysis in Health Care', ISER, University of York, Reprint Series, no. 331.

EDWARDS, M., STRUDWICK, B. & THOMPSON, M. A. 'Specialty budgeting for beginners', The Hospital and Health Services Review, Feb. 1982.

EDWARDS, M., MILLER, J. D. & SCHUMACHER, R. 'Classification of Community Hospitals by Scope of Services: Four indexes', Health Services Research, Winter, 1972.

ELAYAT, H. A., MURPHY, B. M. & PRABHAKAR, N. D. 'Entropy in the Hierarchical Cluster Analysis of Hospitals', Health Services Research, Winter, 1978.

ESPOSITO, A. et al. 'Abstracts of State Legislated Hospital Cost-Containment Programs', Health Care Financing Review, vol 4, 2, Dec. 1982.

EVANS, R. G. 'Health Care in Canada. Patterns of funding and regulation', in G. McLACHLAN, A. MAYNARD (eds.) The Public Private Mix for Health, The Nuffield Provinical Hospital Trust, 1982.(London).

GAUS, C. R. & HELLINGER, F. J. 'Results of prospective reimbursement', Topics in Health Care Financing, vol 3, 2, Winter, 1976.

GLASER, W. A. 'Paying the Hospital in England', mimeo, Health Care Financing Administration Project, U.S., D.H.H.S., June 1981.

GLASER, W. A. 'Paying the hospital: Foreign Lessons for the U.S.', Health Care Financing Review, vol 4, no. 4,Summer, 1983.

GRAY, A. 'Containing Health Care Costs in Western Europe: will present policies work?', Health Economics Research Unit, (H.E.R.U.) University of Aberdeen, Discussion Paper no. 03/1982.

HELLINGER, F. J. 'Prospective Reimbursement through Budget Review: New Jersey, Rhode Island and Western Pennsylvania', Inquiry, Sept. 1976.

HOOD, C. & DUNSIRE, A. Bureaumetrics, Gower Publishing, London, 1981.

HURST, J. W. 'Saving Hospital Expenditure by reducing in-patient stay', HMSO, 1977, Government Economic Service Occasional Papers, no. 14.

INSTITUTO NACIONAL DE ESTADISTICA Encuesta de Morbilidad Hospitalaria, Año 1978, Ministerio de Economía y Comercio, Madrid 1982.

JONES, T. & PROWLE, M. Health Service Finance. An Introduction, C.A.E.T., London, 1982.

KIND, P. & PROWLE, M. J. 'Information and Health Service Management', Technical Paper prepared for the Royal Comission on the NHS at the Centre for Industrial Economic and Business Research, University of Warwick, HMSO, April, 1978.

KLASTORIN, T. D. & WATTS, C. A. 'The determination of Alternative Hospital Classifications', Health Services Research 16:2, Summer, 1981.

MADDALA, C. S. Econometrics, McGraw-Hill, International Student Edition, 1977.

MAGEE, C. C. & OSMOLSKI, R. J. 'A Comprehensive System of Management Information for financial Planning decision making and Control in the Hospital Service', mimeo, Centre for Industrial Economics and Business Research, University of Warwick, Coventry, April, 1978.

NUNN, C. 'A computerised patient administration system in a new Scottish DGH', The Hospital and Health Services Review, Oct. 1982.

ONTARIO HOSPITAL ASSOCIATION RESEARCH COMMITTEE'S 'Outpatient Funding in Ontario Hospitals', mimeo, Background Paper no. 6, Ontario Hospital Association, 1980, Position Paper "Hospitals Inside Out".

PEARSON, D. A. & ABERNETHY, D. S. 'A Qualitative Assessment of previous efforts to contain Hospital Costs', Journal of Health Politics and Law, vol 5, no. 1, Spring, 1980.

PEEL, V. J., MALE, R. S. & GUNA WARDENA, A. 'Computers in in-patient information systems. An alternative local approach', Hospital and Health Services Research, Jan. 1981.

PHILIP P. J. & IER, R. 'Classification of Community Hospitals', Health Services Research, Winter, 1975.

PHILLIPS, D. F. 'Regulation and Data Systems: Questions of demand versus needs', Hospitals J. A. H. A., (51), Oct. 16, 1977.

PROWLE, M. 'Beyond specialty costing: patient costing?' Public Finance and Accountancy, March, 1981.

PINDYCK, R. S. & RUBINFIELD, D. L. Econometric Models and Economic Forecasts, McGraw-Hill, Kogakusha Ltd., Tokyo, 1976.

REISS, J. B. 'A Conceptual Model of the Case-Based Payment Scheme for New Jersey Hospitals', Health Services Research, Summer, 1980.

ROSSER, R. 'Issues of measurement in the design of health indicators: A Review', in A. CULYER (ed.) Health and Health Indicators: Proceedings of the European Workshop., 1981.

RUSSELL, E. M. 'Patient Costing Study', Scottish Health Service Studies, SHHD, no. 31, mimeo.

SCHUMACHER, D. N. et al. 'Hospital cost per case: Analysis using a state wide Data System', Medical Care, vol XVII, 10, Oct. 1979.

SLOAN, F. et al. Hospital Regulation, Lexington Books, 1982.

SMITH, M. I. 'Maximum Allowable Cost: Can the Government Control Drug Costs. (A Retrospective Political Analysis of Events and Issues)', Journal of Health Politics, Policy and Law, vol 4 no.2, Summer, 1979 (pp. 155-175).

SMITH, P. & STEWART, J. 'Cluster Analysis. A simple method for comparing performance', Public Finance and Accountancy, March and April, 1983.

STEELE, R. & GRAY, A. M. 'Statistical cost analysis: the Hospital care', Applied Economics, vol 14, 5, 1982.

STEINWALD, B. 'Commentaries on three Articles on Cost Containment', Health Services Research, 16:2, Summer, 1981.

STEINWALD, B. & SLOAN, F, 'Regulatory approaches to Hospital Cost Containment: A synthesis of the Empirical evidence', in M. OLSON (ed.): A New Approach to the Economics of Health Care, AEI, Washington, 1981.

STONE, D. A. 'Health Care Cost Containment in West Germany', Journal of Health Politics, Policy and Law, vol 4, no. 2, Summer, 1979.

SUGDEN, R. & WILLIAMS, A. The principles of practical cost-benefit analysis, Oxford University Press, 1978.

THOMPSON, J. D., FETTER, R. B. & MROSS, C. D. 'Case-mix and resource use', Inquiry, 12, 1975.

TREND, B. 'Policy and the Public Purse', Times Literary Supplement, July 16, 1982, (755-757).

WARNER, K. E. 'Effects of Hospital Cost Containment on the Developement and Use of Medical Technology', Health and Society, Milbank Memeorial Fund Quarterly, vol 56, 2, 1978.

WATTS, C. D. & KLASTORIN, T. D. 'The impact of case-mix on Hospital Cost: A comparative analysis', Inquiry, Vol XVII, 4, Winter, 1980.

WATTS, C. A. & KLASTORIN, T. D. 'The Cost of Regulation: A further note', Health Services Research, 16:4, Winter, 1981.

WEINER, S. M. 'On public values and private regulation: Some reflections on cost containment strategies', Health and Society Milbank Memorial Fund Quarterly, vol 59, 2, 1981.

WEST MIDLANDS HEALTH REGIONAL AUTHORITY, 'Financial Information Project 'NHS'. Research Program. First Research Project. General Statement', mimeo, West Midlands HRA, May, 1979.

STEERING GROUP on HEALTH SERVICES INFORMATION - WEST MIDLANDS H.R.A. 'A Report of the working Group A on Health Information' and 'A Report of the Working Group B on existing Costing Systems', mimeo, West Midlands HRA, May, 1979.

WHITE, K. L. et al. 'Public Health Services: Concepts and information for national planning and management', Public Health Papers, no. 67, World Health Organizatory Geneve, 1980.

WORTHINGTON, P. N. 'Prospective Reimbursement. New Jersey Formula', Inquiry, Sept. 1976.

WORTHINGTON, P. N. & PIRO, P. 'The effects of Hospital Rate Setting Programs on volumes of Hospital Services', Health Care Financing Review, vol 4, 2, Dec. 1982.

YOUNG, D. W. 'The use of computers in patient administration systems', The Hospital and Health Services Review, Jan. 1981.

ZUBKOFF, M., RASKIN, I. E. & HANET, R. S. (eds.) Hospital Cost Containment: Selected notes for future policy, Milbank Memorial Fund, PRODIST, New York, 1978.