UNIVERSITY
OF YORK

CENTRE FOR HEALTH ECONOMICS
HEALTH ECONOMICS CONSORTIUM

The Market for Private Health Care and the Demand for Private Insurance in Britain

by

Carol Propper
and
Alan Maynard

January 1989

DISCUSSION PAPER 53

UNIVERSITY OF YORK

CENTRE FOR HEALTH ECONOMICS

THE MARKET FOR PRIVATE HEALTH CARE AND THE DEMAND FOR PRIVATE INSURANCE IN BRITAIN

by

Carol Propper

and

Alan Maynard

THE AUTHORS

Carol Propper is a Lecturer in Quantitative Methods in the School of Advanced Urban Studies at the University of Bristol and a Visiting Research Fellow at the Centre for Health Economics, University of York.

Alan Maynard is Professor of Economics and the Director of the Centre for Health Economics at the University of York.

ACKNOWLEDGEMENT

The authors would like to acknowledge financial support for this work from the Economic and Social Research Council and the Health Promotion Research Trust. They would like to thank the secretaries, in particular Sal Cuthbert, for typing the manuscript.

FURTHER COPIES

Further copies are available (at price £6.00 to cover the costs of publication, postage and packing) from:

The Secretary
Centre for Health Economics
University of York
Heslington
York YO1 5DD

Please make cheques payable to the University of York. Details of other Discussion Paper can be obtained from the same address, or telephone (0904) 433648 or 433646.

ABSTRACT

There has been little economic analysis of the private health care industry in Britain and as a result much of the policy debate about its role has been ill-informed and based on rhetoric. The purpose of this paper is to analyse the market for private health care and the demand for private insurance in Britain.

It is shown that the market is highly specialised, providing care for non-emergency (mostly surgical) cases. Care is provided in NHS pay beds and just over 10,000 private beds, the majority of which are owned by for-profit organisations. The sector is very specialised, providing predominantly cold elective surgery and whilst its turn-over is small relative to the NHS, in relation to the particular sub-market in which it operates its role is very significant, with over 1 in 6 cold elective procedures in England being provided in the private sector. Geographical inequalities in this provision mean that some London Regions now have nearly one third of such procedures carried out privately.

The finance for such care comes from predominately non-profit making (provident) insurers. Over 5 million people are covered by private health care insurance and in 1987 over £580 million was paid out in benefits. If private payments in cash are added to this figure, total private health expenditure in 1987 was of the order of £750 million.

The insurers have sought to control costs, arguing that market growth is dependent on premia stability and that increases in contributions will reduce coverage. However, the implied price elasticity (ie the implication that coverage is reduced significantly by price increases) has not been estimated until now. In this paper the authors use available date to

provide estimates of the short (impact) and long run price elasticity of -0.6 and -2.55 respectively.

The low value for the estimate of the short run elasticity implies that if providers raised their charges aggressively and these costs were passed on in high premia, market demand would decline by a relatively small amount. Elsewhere Propper and Eastwood (1988) have argued that private buyers of insurance tend to be conservative in adjusting to market changes. This behaviour is consistent with these elasticity estimates.

The tentative conclusion from this paper, based on the short run elasticity estimate of -0.6, is that providers could be aggressive price makers with little risk of losing market share. This result merits further investigation as the price elasticity value may vary in different sub-markets (for example, it may be higher in the corporate sector). If data were made available such questions could be examined. Further, without better data the discussion of the efficiency of the private sector will remain vague and policy will be based on the images rather than the reality of the market's operations.

INTRODUCTION

Every five or six years familiar problems are rediscovered about the efficiency and distribution of the services provided by the National Health Service and its funding. This process is accompanied by advocacy of the extension of the private sector (e.g. Green (1988)) which usually has little regard to the efficiency of its functioning.

The purposes of this paper are first, to examine the characteristics of the market for private health care in Britain (section 1) and second, to analyse using multivariate techniques the aggregate demand for private insurance (section 2). Estimates of aggregate responses to price and income changes are necessary to predict changes in the market. Throughout the paper our concern will be with the finance and provision of private health care for acute, elective surgery. The private provision of residential and nursing homes for the elderly, the mentally ill and mentally handicapped, which provides 52,000 places at a social security cost to the Government of £1 billion, is dealt with elsewhere (Maynard and Smith (1983), Griffiths (1988)).

1. THE MARKET FOR PRIVATE HEALTH CARE IN BRITAIN.

1.1 Introduction

The market for private elective (mostly surgical) health care in Britain periodically experiences increased growth. "Crises" in the National Health Service (NHS), for instance the industrial disputes of the early years of the Thatcher Government and the recent (e.g. the Governments NHS Review) of parsimonious funding of the NHS, have produced more rapid growth in private insurance. However the sector remains basically very

conservative with, until recently, little market innovation in the design of insurance products or in the control of provider costs.

Indeed the private sector fits in many ways the caricature of the NHS, i.e. constrained by the substantial market powers of financers (such as BUPA, the dominant insurer) and providers (especially hospital owners and doctors) and non-responsive to consumer preferences. Potentially inefficient behaviour of actors in the private health care sector is given impetus by poor data collection and little evaluation of the behaviours of suppliers and financiers. The information that is available appears to substantiate the hypothesis that, as in so many other markets, capitalists are the enemies of capitalism i.e. the effective defence of market position ensures that resources are not used efficiently and that consumers' interests are not enhanced. Insurers, rather than controlling provider costs, have been able to pass on the increasing costs of health care in the form of increased premia, thus ensuring a stable market, rather than appropriate reimbursement (identified using, for instance, methods such as those proposed by Hsino, Braun, Yntema and Becker (1988)) for providers, especially the medical profession. Whilst there has been competition for market shares both in provision and finance the private health care market in Britain remains relatively small, highly specialised and very conservative in evolving new products and controlling old markets.

1.2 The supply of private health care.

The market for private health care provides cold, elective surgery for patients willing and able to pay, either directly (out of pocket) or indirectly from insurance. The market is highly specialised with less than two dozen procedures (e.g. impacted teeth, small benign lesions, hernia

repairs, hysterectomies and haemorrhoid procedures) accounting for over 70 per cent of expenditure. These procedures are generally not life threatening but reduce considerably the quality of life of the potential patients who are often on NHS waiting lists.

Such patients can be referred directly to the private sector by their general practitioners or may choose this care route after a NHS referral to a hospital consultant. The care will be provided in either a private hospital or NHS pay beds and generally the medical care is provided by NHS doctors working part time in the private sector, and by so doing increasing their income if they are surgeons by, on average, about £17,000 (Laing (1987)).

The distribution of private acute care facilities (i.e. in hospitals with an operating theatre) is set out in Table 1. Between 1979 and 1988 the number of private acute hospitals in England increased from 134 to 188 with no growth in the rest of the UK. The growth in the hospital stock over the period was very uneven : there was an increase of 50 hospitals in the period 1979-84 but only 4 have been opened since then, with 22 closures in the years 1985-87 barely being matched by 24 openings (Independent Hospitals Association (1988, p4).

The 40 per cent increase in the number of private hospitals in the 1979-88 period resulted in a 58 per cent increase in the private acute bed stock in England. However, the time trend is uneven over the period with the bed stock declining in 1985 (by 112 beds) before resuming its steady but modest and uneven growth path subsequently as can be seen from the data in Table 1.

Table 1 : Private Acute Hospital Beds

NHS Region	Number of Hospitals			Number of hospital beds			Beds per 100,000 population in 1988
	1979	1985	1988	1979	1985	1988	
Northern	1	3	5	30	144	154	5
Yorkshire	9	14	15	341	535	580	16
Trent	9	14	11	286	524	457	10
East Anglia	5	9	9	123	332	344	17
NW Thames	13	16	17	837	1225	1271	36
NE Thames	24	25	23	1383	1704	1661	44
SE Thames	15	18	19	609	893	1100	30
SW Thames	12	16	14	762	1001	887	30
Wessex	7	15	14	191	630	593	20
Oxford	7	12	12	232	470	478	19
South Western	9	10	11	345	425	475	15
West Midlands	13	17	18	419	606	709	14
Mersey	5	6	7	273	312	354	15
North Western	5	10	13	283	616	605	15
England (total)	134	185	188	6114	9416	9668	20
Wales	4	5	4	202	239	215	8
Scotland	9	9	9	265	414	402	8
Northern Ireland	2	2	2	82	86	86	5
Rest of UK (total)	15	16	15	549	739	703	7
UK (total)	149	201	203	6663	10155	10371	18

Source: Independent Hospitals Association (1986 and 1988) page 2.

These time trends reflect unfounded optimism about the future of the private health care sector in the first half of the 1980s which resulted in low utilisation rates, sometimes barely over 50 per cent, and very low rates of return on investments (Fitzhugh (1987)). The more modest growth rates in the second half of the decade reflect consolidation resulting from hospital competition and more vigorous cost containment policies.

In part these policies are the result of a change in ownership patterns. In 1979 less than 30 per cent of beds were owned by for-profit hospitals. By 1988 57 per cent of hospitals and 55 per cent of beds were in the for-profit sector. The pressures for greater economy induced by this changed pattern of ownership are being enhanced by the actions of the insurers, some of whom (e.g. BUPA) own hospitals, in an effort to reduce cost pressure on the rate of growth of their premia.

An alternative supply of private beds is to be found in the NHS, in pay beds. After restriction by the Labour Government in 1976, which contributed substantially to the growth of investment in private sector beds, the level of the NHS pay bed stock has grown since 1979 by over 15 per cent to about 3000 beds, but throughout this period utilisation has been been uneven and sometimes low. The growth in the pay bed stock and the efforts to increase utilisation reflect the drive in the NHS for "income generation" by NHS District managers and the preference for pay beds of insurers, a preference arising from the shorter lengths of stay (and hence costs) of patients in these beds (Williams, Nicholl, Thomas and Knowelden (1985)).

In summary, in comparison to the total expenditure on the NHS, the size of the private acute health care sector is modest. However, in

comparison to NHS cold (elective) surgery it is substantial with over 16 per cent of all such procedures being carried out in the private sector in England. Reflecting the unequal distribution of private facilities shown in Table 1, private sector elective facilities carry out over 30 per cent of all cold procedures in some Thames (London) regions (Nicholl, Beeby and Williams 1988).

The supply of private health care is small (in relation to the NHS) but highly specialised, significant in its particular sub-market (elective surgery) and growing modestly. Expenditure on these services, mostly financed from insurance, is about £750 million per annum.

1.3 The finance of private health care.

The principal insurance carriers in the UK health care market are the provident associations which are insurance companies registered under various Insurance Companies Acts as "pecuniary loss" insurance companies, limited by guarantee and non-profit distributing. The role of the individual guarantors, whose support is nominal, is limited to access to information and a role in advising the company. All these companies' surpluses of income over expenditure becomes part of the company's reserves and are not subject to corporation tax.

These companies, the largest of which are British United Provident Association (BUPA), Private Patients Plan (PPP) and Western Provident Association (WPA), once they have met minimal legal requirements (Maynard (1983)), can design and market policies. At present they control nearly 90 per cent of the private health care insurance market and although BUPA's market share has fallen significantly during the 1980s, it remains the market leader.

Table 2 Private Health Care Insurance Market 1979, 1985 and 1987

	1979	1985	1987
Subscribers ('000s)	1292	2107 (2364)	2234 (2573)
Persons insured ('000s)	2765	4506 (5030)	4526 (5254)
Subscriptions paid (£mn)	122	470 (519.6)	628 (711.0)
Benefits paid (£mn)	84	414 (455.9)	510 (581)
Persons insured (% UK pop)	5	8 (8.9)	8 (9.3)

Notes 1. The data (outside brackets) are for BUPA, PP and WPA only.

2. The 1985 and 1987 data in brackets are for all insurers. There are no 1979 data for all insurers.

Source: Laing (1988) Table 3.1, Page 42

It can be seen from table 2 that the number of subscribers and the estimated number of persons insured has increased significantly since 1979 although the growth rates between 1985 and 1987 were modest. The margin between subscriptions and benefits paid is considerable with the latter (the level of benefits paid) being in excess of £580 million for all insurers in 1987. In 1987 it was estimated that over 9 per cent of the UK population was covered by private insurance.

For both coverage and benefits paid there has been a slow erosion of the market share of the non-profit providents. However the share of the for-profit companies remains modest and less than 10 per cent of the total market. With insurance finance of just over £580 million and self-finance of about 30 per cent, total expenditure on private health care in 1987 was of the order of £750 million.

The providents offer three basic policies : these are individual, employee and company purchase. The growth in these sectors over the last decade has been very uneven with the most significant growth being in the company sector and the most significant decline being in employee sector after its peak early in the 1980s.

Table 3 : Purchase Categories for BUPA, WPA and PPP, 1977, 1982 and 1987

(000s %)

Year	Individual purchase	Employee purchase	Company purchase	Total
1972	459 (43.4)	175 (16.5)	424 (40.1)	1058
1982	546 (28.5)	438 (22.8)	933 (48.6)	1917
1987	610 (27.3)	378 (16.9)	1246 (55.6)	2234

The benefits paid are divided largely between room charges and surgeons' fees. The share of surgeons' fees over the period 1977-1987 has been relatively stable at 23 per cent of the total benefits paid. The share of room charges over the same period has fallen from 50 to 44 per cent. The other large item (15.6 per cent) in 1987 is miscellaneous inpatient and outpatient fees (Laing (1988) Table 3.2).

As Laing shows, the growth in real subscription income (deflated by the Retail Price Index) has been highest during the period when inflation of private insurance prices has been greatly in excess of general price movements. Thus whilst enrolments seem to be affected by premia changes, subscription income appears to be less price elastic.

The consideration of the price elasticity of demand for private health care insurance is important not only in relation to enrolments and subscription income, but also in regard to supply side inflation. The financers, both provident and for-profit, have emphasised that hospital and other fees must be controlled tightly if the market is to grow. If, however, the price elasticity of demand for insurance is low, insurers can pass on cost increases in higher premiums. To determine whether cost containment or cost pass-ons in higher premiums is the appropriate behaviour it is necessary to have an estimate of the price elasticity of demand for private health insurance.

1.4 Overview.

To summarise, the market for private health care is small in relation to the NHS and highly specialised. There is some competition between insurers for market share and the innovators have succeeded in reducing the share of the market leader, BUPA. The major changes in the 1980s have been

a cost explosion induced by fee increases (initiated by BUPA), changes in the contracts of NHS doctors which have enabled them to do more private work, and an extension of the market to higher risk (blue collar) occupations. The insurers initially responded to the cost increases with rhetoric and moral suasion: recently have they sought to control costs with the use of fixed price contracts with providers and financer-provider links on the preferred provider organisation (PPO) model of the USA This latter innovation in the UK market was made by Crusader Insurance and Nuffield Hospitals in 1988.

The insurers face difficulty in expanding their activities because aggregate market demand for their product is determined by the capacity of the NHS to care for patients. The tighter public expenditure controls are, the greater the illusion and perhaps reality of NHS "collapse" appears, ceteris paribus, the greater is the scope for private insurance. In the near future the private market is unlikely to expand rapidly given Government assurances of their continued finance of the NHS. However, the division of the market between individual insurers and providers may further change. Such developments will be influenced by factors such as the price elasticity of demand.

2. ECONOMETRIC ESTIMATION OF THE AGGREGATE DEMAND FOR PRIVATE HEALTH INSURANCE IN BRITAIN.

The objective of this section is to identify the effect of changes in price and income on the aggregate purchase of private health insurance in Britain. We seek to estimate an aggregate demand function of the form

$$q = f(y, p, z)$$

where the dependent variable, q, is total health insurance policies sold in any one calendar year by the three largest suppliers of health insurance, y is income, p is price and z denotes other factors hypothesised to affect the level of demand. The sales of the three largest companies (all Provident Associations) account for over 90 per cent of all sales for most of the time period under analysis. The total number of policies sold (referred to in the health insurance market as subscriptions) is the sum of the sales of the three types of subscription, individual, employee and corporate. The benefits offered by the policies sold to each group are very similar. The first two types of policy are paid for by the individual who is covered, the latter generally being purchased at a lower price; the last type of policy is purchased by the employer. While the motivations for non-corporate and corporate purchase may be different, data on the total number of policies sold to each type of subscriber is available as a continuous series only from 1977, and so there is insufficient information to permit econometric analysis of trends for each subscriber group separately.

In sections 2.1 and 2.2 we describe the data used in the analysis. In section 2.3 we describe the econometric methodology used for the analysis. Readers familiar with, or willing to take on trust the econometric methodology can proceed directly to sections 2.4 and 2.5, in which the results of the analysis are presented.

2.1 Data and trends.

The data set comprises annual observations for the period 1955 to 1986 on the following variables:

SUB : total subscriptionds of BUPA, WPA, PPP

REV : total revenue from subscriptions for BUPA, PPP, WPA

CEXP : nominal consumers expenditure/population over 15

RPI : retail price index (1975=100)

PDI : personal disposable income/population over 15.

The subscription data were obtained from published sources (Lee Donaldson 1982; Grant 1985; Laing 1987), the other information from published CSO tables. From these basic variables the average subscription price was calculated as

PRICE = REV/SUB,

and the real price, RPRICE, derived by deflating the nominal price by the rpi. Consumer expenditure and personal disposable income were also deflated by the rpi to obtain real (1975 prices) values. All variables were transformed to logs and these are denoted by lower case letters. Differences are represented by the prefix d, for example, $dprice_t = price_t - price_{t-1}$.

The scope of the private sector is to a large extent defined by the existence of the NHS (Propper 1987), and private health insurance provides most of the finance for private sector treatment. Therefore we would expect the demand for private health insurance to be affected by changes in the NHS. During the sample period, there were two periods of intense debate over the nature and funding of the NHS and the role that a private sector should play in the medical care market of the UK. The first period was the mid 1970s, marked by considerable industrial unrest such in the health sector and the abolition of pay beds by the Labour administration in 1976, and the second was the late 1970s and early 1980s, marked by the

return of a Conservative administration, new contracts for consultants in 1980, which gave greater incentives for private practice (Maynard 1982) and limited tax deductibility on private health insurance premiums. Dummy variables were use to capture the effect of these shocks. The constructed dummies were:

D75 : dummy for beginning of period of uncertainty over future of private health care in mid 1970s. Has value 1 for 1975, 0 otherwise.

D7577 : dummy for period of uncertainty in mid 1970s. Has value 1 for 1975 - 1977, 0 otherwise.

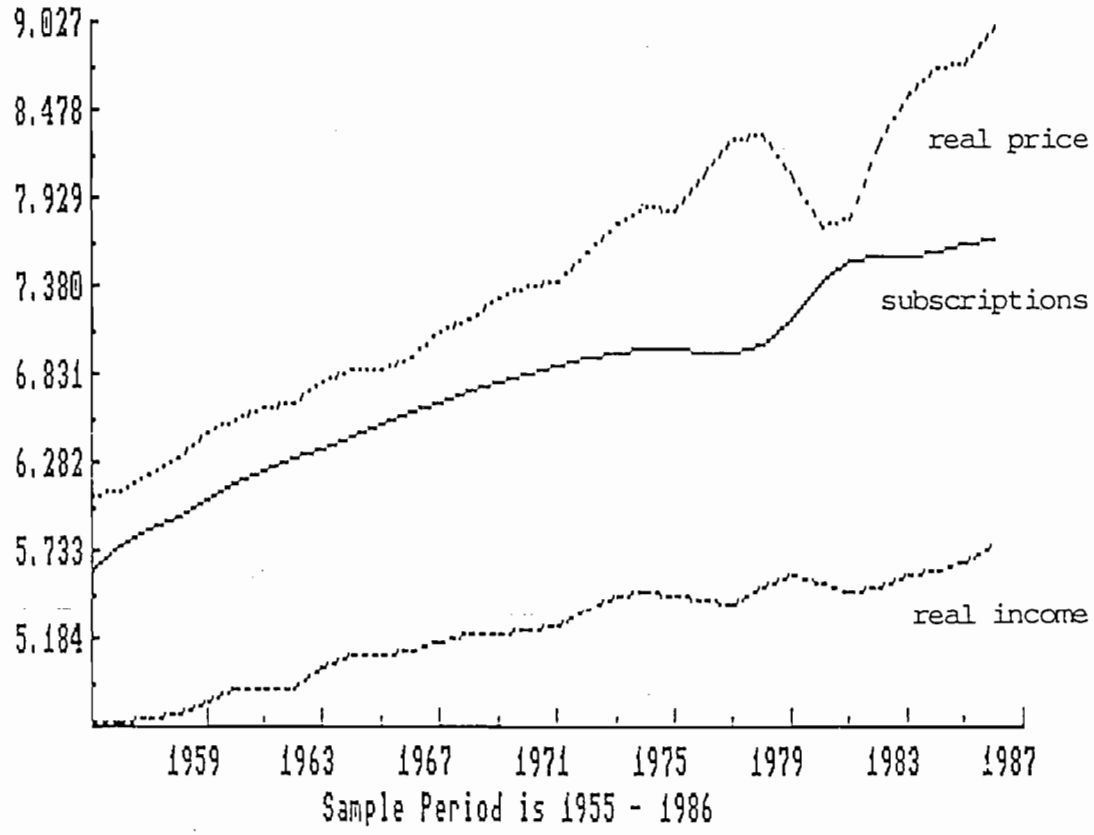
D8086 : dummy to mark period covered by new contracts for consulants. Has value 1 for 1980-1986, 0 otherwise.

2.2 Analysis of trends

The trend of the (log of) total subscriptions, real price and real income is plotted in Figure 1. The number of subscriptions grew steadily but at a relatively slow rate until the mid 1970s, plateaued, then rose sharply in the late 1970s/early 1980s, plateuing off again from 1982 onwards. The overall trend in real consumer expenditure has been upwards, although there was a fall in the late 1970s, followed by a rise 1982 onwards. In annual data, the pattern of real consumers expenditure and real personal disposable income are very similar for the sample period and in the subsequent analyses we present results using only real consumer expenditure. The trend in the real price of subscriptions has been upwards, but with some fluctuations.

Figures 2 and 3 show the graphs of subscriptions against real income and real price respectively. The second graph shows the relationship between subscriptions and income is as predicted from both theoretical

Fiugre 1: (Logarithm of) total subscriptions, real price of subscriptions and real income, 1955-1986.



Notes: price and income are scale up by 2 to permit presentation on the same graph.

total income proxied by total real consumer expenditure

Figure 2: (logarithm of) total subscriptions plotted against (logarithm of) average real price

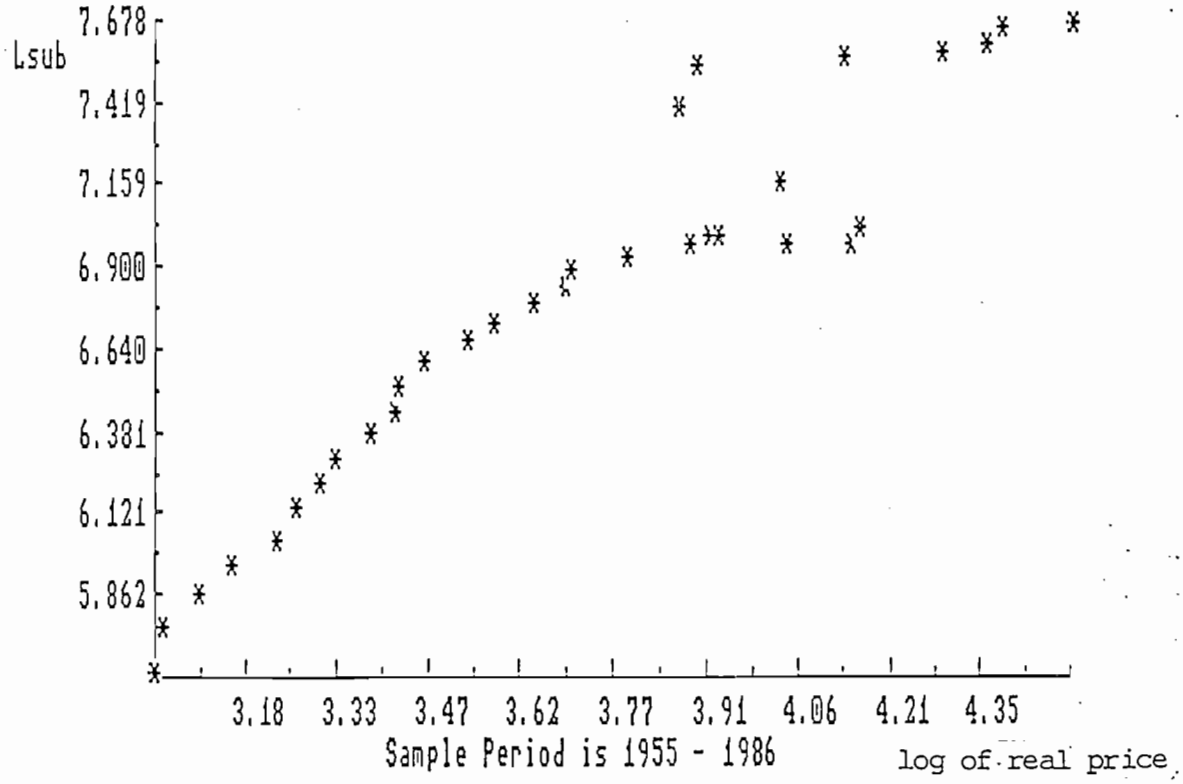
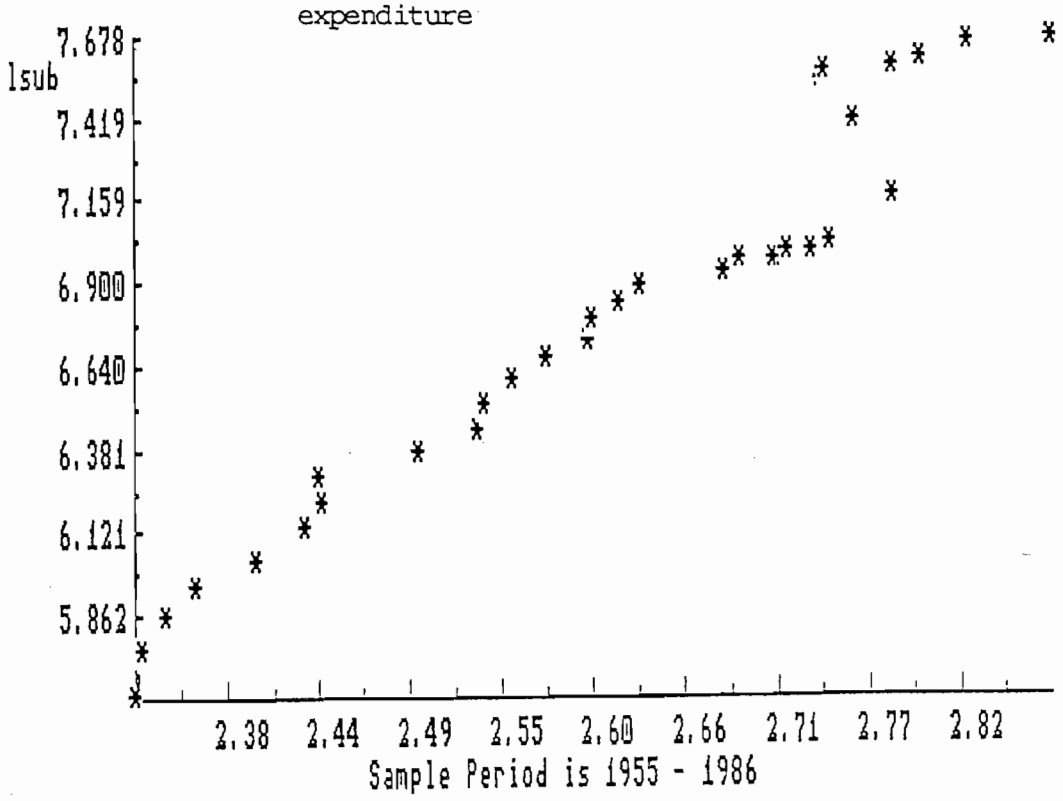


Figure 3: (logarithm of) total subscriptions plotted against (logarithm of) total real consumer expenditure



Note: real income is proxied by total real consumer expenditure.

analysis and prior cross-sectional empirical analysis in the UK (Propper 1987; Smith 1988) and USA (Manning et al 1987). The positive relationship between price and subscriptions of Figure 2 is somewhat unexpected. However, the figure does not preclude a negative, demand function relationship between price and quantity, holding other variables constant.

Generally, the correlations between the levels of subscriptions, income and price in one year, and the autocorrelations between the levels of each variable in one year and its own lagged level is very high. These correlations are presented in Table 4. The high values for the correlation coefficients suggest that it may be difficult to identify the separate effect of each variable in analyses using levels of the series.

Table 4: Correlation Matrix of Levels of Variables

	Sub	Price	Income	drpi
Sub	1.00			
Price	.9451	1.00		
Income	.9691	.9656	1.00	
d rpi	.4461	.4604	.5351	1.00

All variables in logs

Our aim is to estimate a demand function. It is therefore necessary to consider the issue of causality and the question of the exogeneity of the variables defined as exogenous and the endogeneity of the variable defined as the dependent variable. It seems reasonable to assume that the level of national income (or real consumer income) and the change in the general price level in the economy are exogenous and that the direction of causality runs from changes in these variables to changes in quantity of health insurance purchased. However, the exogeneity of the price of insurance contracts may be more of an issue.

Estimation of a demand function presumes that changes in price determine changes in quantity; however, under certain circumstances the causality in the health insurance market may run in the opposite direction. If insurance companies respond to a change in the number of subscribers by altering the price of insurance, then price may be endogenous. Relatively little is known about the price setting mechanism in the UK health insurance market. The price of contracts is the outcome of discussion

between the leading health insurance supplier (BUPA) and the medical care providers (the BMA and the private sector hospitals, some of which are owned by BUPA). Premiums generally appear to rise in line with increases in medical care costs. An increase in the demand for private insurance is likely to be but one determinant of the overall level medical care costs. Other factors include the capital costs of building hospitals, the ratio of capital to staff, the target income of doctors (if doctors seek a target income), and the quantity of uninsured patients. Thus an increase in quantity of medical insurance sold will not necessarily lead to an increase in the prices charged by private hospitals. Generally, therefore, the price of insurance would appear to be exogenous. However, there is one situation in which price may not be exogeneous - the case of adverse selection. If the insurance companies were to attract new (additional) subscribers who had a higher claims rate than previous buyers, and the insurance companies did not ex ante adjust their premiums, then the suppliers would make a loss (or lower profits than expected). If the insurance companies respond to these lower than expected profits by raising the premia, then the direction of causality would run from the change in quantity to a change in price and not vice versa. This occurrence of this event depends on an unexpected change in the average risk of subscribers and subsequent action by suppliers to recoup the unexpected shortfall. This may have occurred in the UK market in the early 1980s. There was, as noted above, a sharp rise in the number of subscribers in 1979 and 1980. This was accompanied by an increase in the ratio of benefits paid to subscriptions earned in 1980 and 1981 and this followed by a sharp increase in price in 1982. However, with the exception of this period, it seems reasonable to assume that price is exogeneous.

2.3 Econometric methodology

We adopted a general to simple modelling strategy in which the starting point was a dynamic relationship between the level of subscriptions, price, income, changes in the general price level and shocks to the health insurance market, in the form of the dummies described above. The division of relationships into dynamic and static can be made on the basis of periodicity of the data; dynamic models being used for quarterly or monthly data, static models for annual data (Harvey 1981: Ch.1). Health insurance contracts run for a year, and can be purchased at any time, so in aggregate, the level of demand may respond only slowly to changes in price and income. Therefore it was felt that although only annual data was available, a dynamic model would be a useful starting point.

A general specification of a dynamic model is

$$Y_t = \alpha + \sum_{j=1}^{n} \beta_j Y_{t-j} + \sum_{j=0}^{n} \gamma_j X_{t-j} + \varepsilon_t \qquad (2)$$

where y_t denotes the variable selected as the endogenous variable, x_t is a vector of explanatory variables, n is chosen by the researcher to be appropriate to the underlying process and the periodicity of the data.

Within a time series framework it is necessary not only to estimate a model, but also to test that the model represents a long-run relationships between the series of interest. Recent econometric work has argued that for a long run relationship betwen two or more variables to exist, the variables must be cointegrated (see Hall (1987)) for an exposition of the theory underlying this statement and the Oxford Bulletin of Economics and Statistics (1986) for further discussion). Whether a set of variables is cointegrated depends first, on the properties of each of the variables in

the set and second, on the properties of the set taken as a whole.

Many economic variables are non-stationary; they do not have a mean which is independent of the sample period (for example, GNP appears to grow over time). However, they may be stationary in first or second differences. If so, the variable is termed integrated, of order 1 or 2 respectively. Cointegration theory states that a linear combination two or more non-stationary series may be a stationary series. In other words, even if two series are trended and so non-stationary, the difference between them may be constant. In order for two variables to be cointegrated, both must be integrated at the same level of integration. A starting point for econometric analysis is therefore to check the order of integration of the series under consideration to see whether cointegration is possible. If it is possible, it is then worth proceeding to the estimation of a model.

Given the trended nature of many economic series it is quite often the case that there is high correlation between the levels of different variables at any time t and between the value of each variable at time t and the own variable lagged terms. This can mean that estimation of a model in the form of equation (2) results in poorly defined parameter estimates. However, such a model can be re-expressed as an equation in differences and lagged levels. For example, letting n=2, (2) can be rewritten, without imposing any restrictions, as

$$dY_t = \alpha_o + \alpha_1 \ dY_{t-1} + \alpha_2 dX_t + \alpha_3 d^2X_t + \alpha_4(Y_{t-2} - X_{t-2}) + \alpha_5 X_{t-2} + \varepsilon_t \quad (3)$$

$$\alpha_1 = (\beta_1 - 1),\ \alpha_2 = (2\gamma_o + \gamma_1),\ \alpha_3 = -(\gamma_o + \gamma_1),$$
$$\alpha_4 = (1 - \beta_1 - \beta_2),\ \alpha_5 = (\beta_1 + \beta_2 + \gamma_o + \gamma_1 + \gamma_2 - 1)$$

Estimation of a model in the form of equation (3) rather than (2) is useful if the levels of the series are highly correlated (Gilbert 1986).

On one level (3) is simply a reparamaterization of (2), which may allow the researcher to overcome the problem of highly correlated series. However, in addition, if certain restrictions are imposed on equation (3), it is not only a useful reparamaterization, but also falls into a class of models known as Error Correction Models (ECM) (Gilbert 1986). These models, first used by Sargan (1964), posit that the dynamic relationship between a set of variables can be characterised by changes in the dependent variable being a function of changes in explanatory variables and lagged differences between the levels of dewpendent and explanatory variables. The long run equilibrium is defined by the relationship between the levels of the variables given in the ECM. Estimtion of an ECM model amounts to estimation of a model of the form of equation (3) subject to certain restrictions on the parameters. These restrictions are given by the long run relationship between the levels variables. (For further discussion of ECM models see Gilbert (1986)).

Dolado and Jenkinson (1987) have pointed out that there are two approaches to estimation of an ECM model. The first is to estimate the long run relationship between the current levels of the series from a regression of the dependent variable on the current levels of the explanatory variables (termed the 'cointegrating regression'), and then to estimate an equation of the form of (3), subject to the restrictions given by the cointegrating regression. This is the approach suggested by Engle and Granger (1987). (For an application see Hall (1986)). The second approach is to start by estimating a model of the form of equation (3). Once such a model is estimated, the researcher then tests to see whether

(3) satisfies a long run stable relationship by estimating the cointegrating regression derived from the levels terms in equation (3) and checking to see that this vector cointegrates.

In this paper, although we hypothesised that the effect of an increase in price would be to decrease quantity and an increase in income to increase in income to increase quantity, economic theory provided no hypotheses about the magnitude of these coefficients or of the other factors that might determine purchase. Nor had we previous estimations to take as a starting point. Therefore, the ECM approach was adopted basically as a reparamaterization of a relationship of the form of equation (2), in order to deal with series which are highly correlated.

The approach taken in this paper was first to test the order of integration of each of the variable under analysis. Conditional on the possible existence of a cointegrating vector a model of the general form of equation (3) was estimated. Finally, a cointegrating regression between the significant lagged levels variables from the estimated difference equation was estimated. This was then tested to check for the existence of a stable long run relationship.

2.4 Analysis

The analysis began with a check on the order of integration of the series under consideration. If the series are not integrated at the same order (or a subset of the series integrated at the same order as all other series under consideration) then no long relationship will exist between these variables. If this is the case, then the researcher may simply be estimating a spurious regression - a regression of two or more trended

variables which have similar trends but no long run stable relationship. If the series are integrated at the same order, then it is possible to move on to estimate this relationship.

2.4.1 Order of integration of the variables

To test the order of integration of the date used in the analysis, three tests proposed in the literature on cointegration were used. These were the Dickey-Fuller (DF) test, the Augmented Dickey Fuller (ADF) test and the Cointegrating Regression Durbin Watson (CRDW). The null hypothesis for all these tests is that the series under consideration follows a random walk: the alternatives differ between the tests. For a discussion of the properties of these tests see Engle and Granger (1987). Table 5 reports these statistics and critical values for the series and their first differences. It is clear that none of the levels of the variables are stationary processes. Both the DF and ADF tests are below 2 and the CRDW close to 0. In first differences, subscriptions, real consumer expenditure and real price (with the rpi as deflator) appear to be stationary. Thus these series appear to be integrated of order 1. However, the remaining series rpi is not stationary, but the tests indicate that differencing d rpi twice produces a stationary series. These results imply that a linear combination of all variables except rpi in levels, plus drpi (i.e. the inflation rate), could be a cointegrating vector. In other words, there may exist a stable long run relationship between these variables.

Table 5: Tests for the order of integration of the variables

Variable	DF	ADF	CRDW
Sub	-1.84	-1.94	0.03
rprice	-0.537	-0.54	0.064
rcexp	-1.539	-1.12	0.042
rpi	0.993	-0.43	0.0134
dsub	-2.57	-3.78	0.65
drprice	-2.91	-5.56	1.08
drcexp	-4.06	-4.810	1.60
drpi	-1.784	-1.55	0.4023
d^2rpi	-5.204	-4.86	2.00

Critical value for CRDW at 5% = .386; critical value for DF at 5% = -3.37; critical value for ADF at 5% = -3.17.

Engle and Granger (1987) suggest in situations where test statistics conflict that an ADF test should be used.

2.4.2 Model estimation

We began by estimating an unrestricted dynamic equation linking total subscriptions to real price, income (as measured by real consumer expenditure), the inflation rate (drpi) and the various dummies, with 2 lags on each of the non-dummy variables. The final specification indicated subscriptions lagged one period, price in period t and with one lag and the change in rpi were significantly associated with the level of subscriptions, but income, in the current period or lagged, did not appear

to be significant. However while the R^2 of this model was high, the DW close to 2 and the RESET test statistic well below the 1 per cent critical value, the variables in the equation were highly correlated. This may make individual parameter estimates unreliable guides to the effect of variables on purchase.

To obtain more orthogonal, but still interpretable variables, the model was re-expressed in the form of equation (3) above. The dependent variable was therefore sub_t, the independent variable lags and differences in price, income and the inflation rate. The most general model estimated allowed for up to 2 period lags. A fairly simple search procedure led to the following model

$$\begin{aligned} dsub_t = &- \underset{(.036)}{.127}\, sub_{t-1} - \underset{(.039)}{.342}\, price_{t-1} - \underset{(.046)}{.604}\, dprice_t \\ &- \underset{(.016)}{.083} D75 + \underset{(.003)}{0.21} t + \underset{(.280)}{1.86} \end{aligned} \qquad (4)$$

$T = 1957 = 1986$

$R^2 = .922$ $\partial = .0153$ DW = 1.83

AR F(2,23) = .11 RESET F(2,23) = 1.04

ARCH F(2,21) = .37 Normality $\chi^2(2) = 7.95$ Chow (10) = 1.47.

standard errors in parentheses

The effect of lagged subscriptions in this equation is negative, as would be expected if habit were a determinant of purchase, and the effect of a price change and the level of price are negative. The trend term, intended to proxy the growth in the private health care sector, so reflecting greater choice and perhaps quality of private facilities, is positive. Of the dummies for change and debate in the health sector, only D75 was consistently significant. The effect of lagged and first

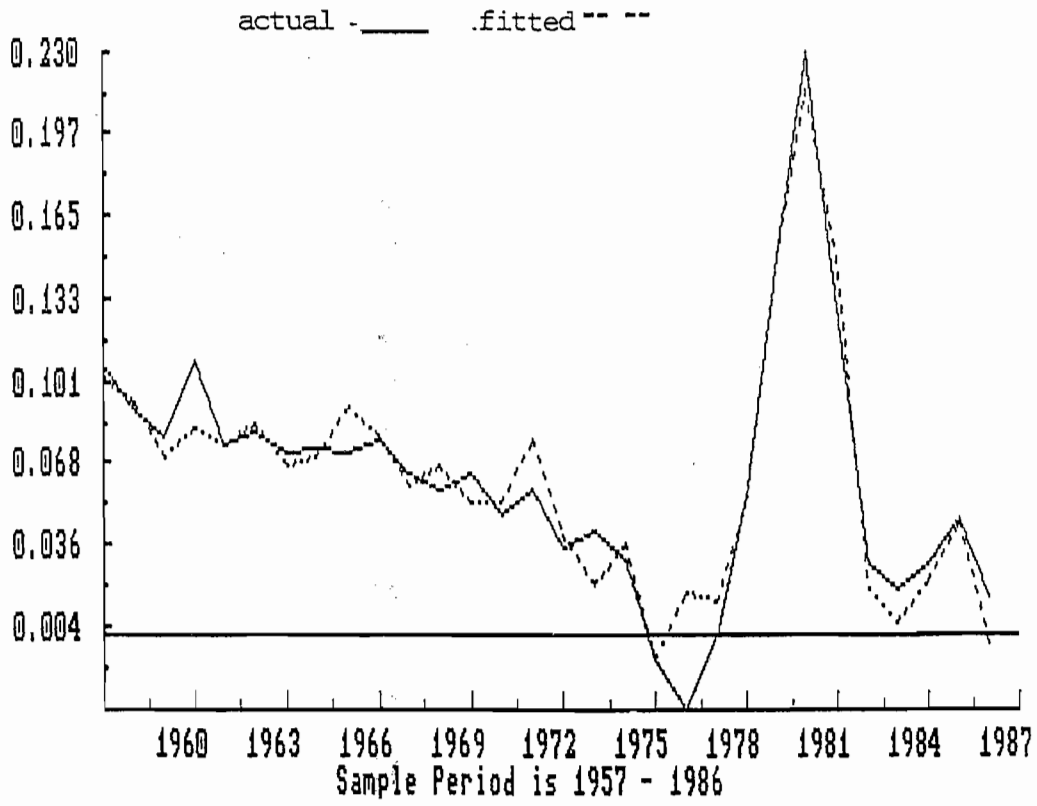
difference terms in income were small in magnitude and not significantly different from zero in all specifications of a differences model, so no income terms are included in equation (4).

The functional form of equation (4) appears acceptable. The RESET test and the test for residual autocorrelation are below the 1% critical value and the test statistic for residual heteroscedasticity (ARCH F[..], Engle 1982) is similarly insignificant. The tests statistics for normality (Jarque and Bera 1982) indicates some misspecification, perhaps due to omitted variables. The normality requirement is less essential than that of homoscedasticity, but suggests that the statistics which assume a normal distribution should be interpreted with caution. However, the parameter estimates are all well defined and no t statistics are close to their 5 per cent critical values, so the departure from normality may not be a large problem. Estimation of equation (4) using only the first 20 observation and then carrying out a Chow test to detect differences between this and the full sample model indicated no significant differences between the full and sub-period estimates. Finally, the goodness of fit, as measured by the standard error of the regression, σ , is close. This is also shown in the plot of actual and fitted values in Figure 4. The graph shows the sharp drop in subscriptions in 1976 and the sharp rise in 1980, both of which are well picked up by the model.

A further test of the model is to estimate equation (4) by recursive least squares (RLS) and to examine the constancy of key parameters as the sample size is increased. In order to use recursive procedures, the parameter of the dummy variable for the debate over the future of the private sector within the NHS, D75, was first estimated freely and then imposed at the full sample value to construct an adjusted level of

subscriptions, denoted dsub*. Equation (4) was re-estimated using RLS for this dependent variable, using 16 observations to 'start up' the estimation. Plots of the one step residuals (one step forecast errors) indicated satisfactory consistency of the model. To examine the stability of particular parameters, RLS can be used to show how the parameters change as the sample size increases. This exercise indicated that while the parameters of lagged price and difference in price both show some variation they are both fairly constant from 1975 onwards. The variation in early years of the sample might have been expected given the deviation of the fitted from the actual values in the early part of the sample shown in Figure 4 and the small size of the sample.

Figure 4: Plot of fitted values from model 4 against actual values



From the levels terms in (4) a cointegrating regression for this model was estimated and tested for cointegration using the CRDW. The only levels variables in (4) are price and subscriptions. The cointegrating regression had a DW statistic of .513 (5% critical value .386), indicating that the hypothesis of a random walk can be rejected. We therefore conclude that (4) is not a spurious regression i.e. net the result of a regression of trended variables which are actually unrelated.

The estimated model appears to provide a close and reasonable consistent fit to the data. The sample is the outcome of an (unobservable) data generation process and it might be expecteds that there are other models which provide a similar fit. The model was re-estimated including the lagged change in rpi (i.e. treating drpi as a levels variable). The parameter estimates remained fairly constant, and drpi had a negative coefficient but the RESET test statistic of 6.4 indicated some functional form misspecification.

It is possible that the general first order model of equation (4) could be simpliied to an AR(1) process. An AR(1) process is a special case of the general dynamic model of equation (1); moreover, when only annual data is available ARMA models may be used to capture the effects of aggregation in what is basically a dynamic relationship. An AR(1) model with a lagged dependent variable was estimated using an algorithm proposed by Hatanka given in the statistical package LIMDEP (Greene 1985). The resulting model is given below.

$$sub_t = \underset{(.221)}{.813}\ income_t \underset{(.073)}{-.323}\ price_t + \underset{(.066)}{.927}\ sub_{t-1}$$

$$T = 1957 - 1986 \qquad (5)$$

$$RESET\ F\ (2,24) = 7.83$$

The RESET test indicates functional form misspecification, so the AR(1) specification can be rejected. In addition, equation (5) gives a very high long run price elasticity.

2.5 Elasticity estimates

We present, the elasticities for equation (4). The effect of income is insignificant in this equation and therefore income elasticity estimates cannot be derived from the current research.

	Impact	Long Run
Equation (4)	-.604 (.046)	-2.55 (.523)

Standard errors (derived from a Taylor series approximation for the long run elasticity) are given in brackets.

The difference between the long and short run elasticity is in the expected direction. The 95% confidence interval for the impact elasticity for model (4) lies between -.5 and -.7, so demand appears to be price inelastic, as would be expected in a market in which habit plays a large part or in which individuals make their decision over a longer time period that the contract period of a year (see Propper (1987) for further discussion). Both the point estimate and the interval estimate for the

long run elasticity is considerbly larger than the impact elasticity. In the short run therefore, the price elasticity appears to be inelastic; in other words the aggregate response to a unit increase in price is less than a unit fall in aggregate subscriptions.

3. LAPSES AND NEW ENROLMENTS

The level of subscriptions in any one year is the sum of the number of new subscribers plus subscribers in the previous year who have renewed their subscriptions minus those who have not renewed their subscriptions. In the health insurance market the first group are termed new enrolments, the last lapses. The above analysis of aggregate subscriptions is therefore complemented by an analyss of lapses and new enrolments. Unfortunately, published data on lapses and enrolments is only available for the years 1960-1983. The short length of the series precludes extensive econometric analysis of possible dynamic relationships, so we have used graphical analysis together with simple econometric estimation to check of the hypotheses suggested by the two dimensional graphs.

From Figure 5 it is clear that the overall trend in both lapses and new enrolments during the sample period has been positive, but both series show some fluctuations, particularly in the second half of the sample period. In this latter half, the changes in enrolments appear to be followed, with a one or two year lag, by similar changes in the level of lapses. To examine this in more detail, changes in lapses were plotted on the same graph as changes in enrolments (Figure 6). In the first half of the sample, there appears to be no clear pattern, but from about 1976 onwards, the behaviour of the change in lapses appears to follow the changes in new enrolments fairly closely with a two year lag. This was

Figure 5: Plot of (logarithms of) enrolments and lapses, 1960-1983.

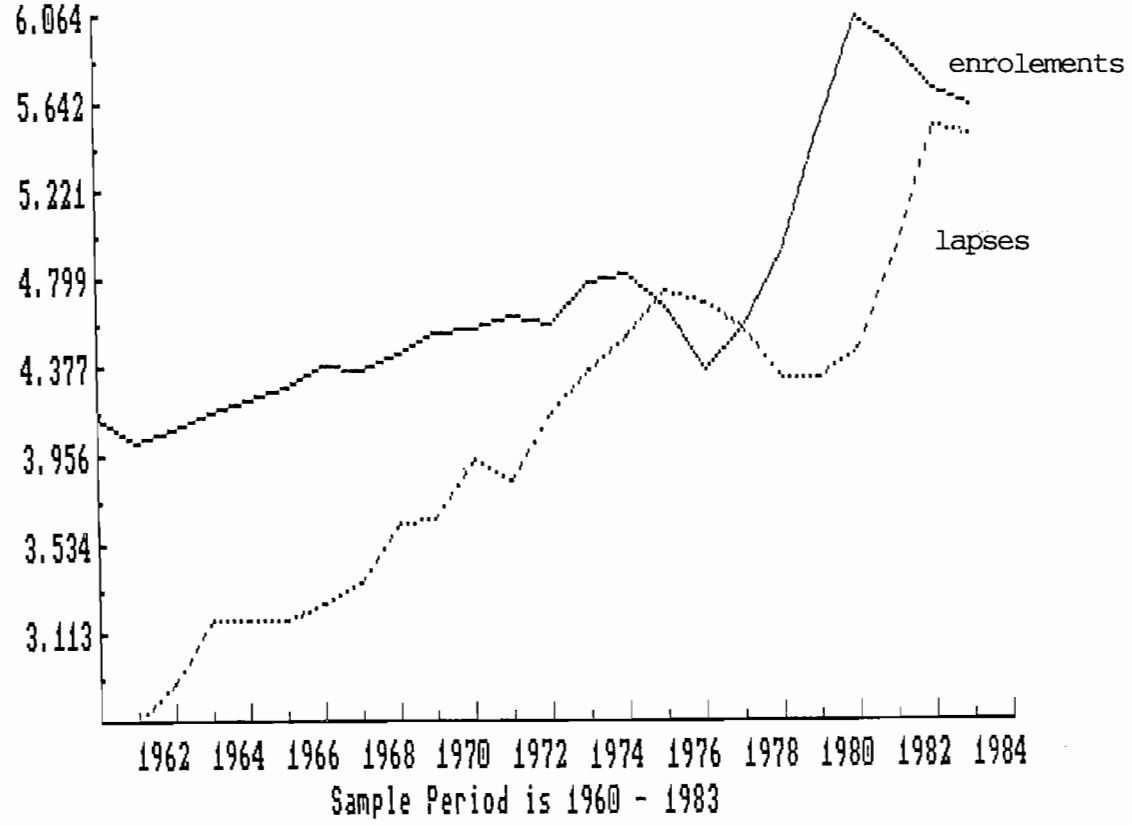
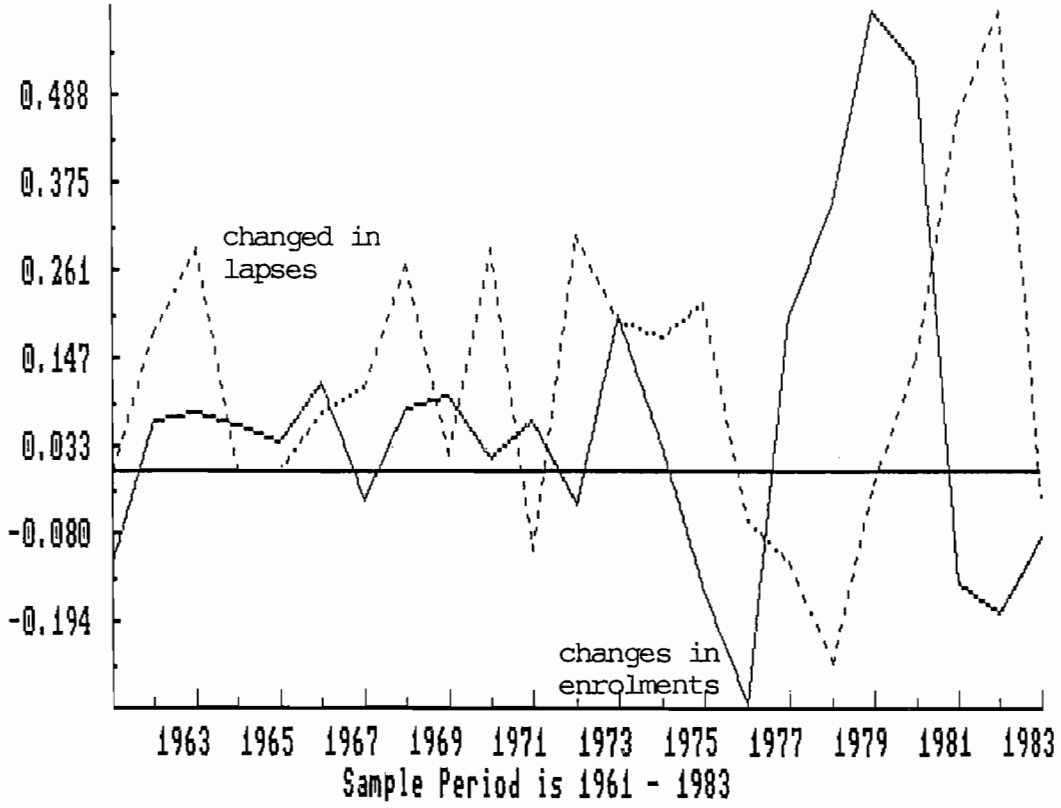


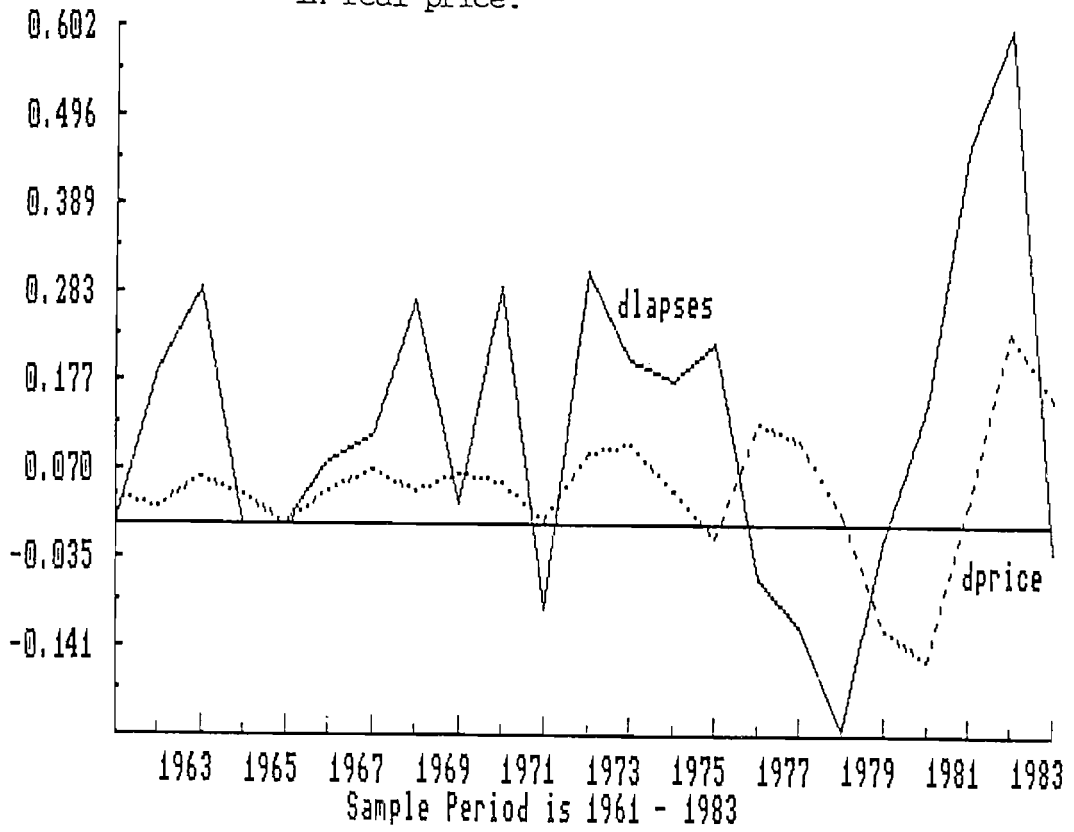
Figure 6: Change in lapses and enrolments over sample period



confirmed by the plot of lapses against two year lagged enrolments. From this it appears, at least for the second part of that period, that one possible explanation for the level of lapses is the extent of market activity, as measured by the number of enrolments. However, the pattern in both lapses and enrolments may be due to the relationship between other variables and each of the two series, rather than to any relationship between lapses and enrolments. Obvious candidate 'other' variables, given our analysis of aggregate subscriptions, are aggregate income, the average price of subscriptions and the level of inflation.

Taking lapses first, it was expected that the level of lapses would be a positive function of the real price of subscriptions and the rate of inflation (a proxy for a fall in the level of wealth) and a negative function of the level of income. A plot of the level of prices and the level of lapses indicated a very close association in the expected direction. However, the plot of the level of lapses against the level of income lagged one year also indicated a close positive relationship. This latter association is of the opposite direction from that expected. These two graphs taken together suggested that the apparent association between either or both price and income and lapses may be spurious. Crossplots of the changes in lapses against changes in price and changes in income respectively indicates that this is probably the case for income, although there may be some association, in the expected direction, in the latter part of the period between price lagged one or two periods and lapses (Figure 7).

Figure 7: Plot of changes in lapses against changes in real price.



Taking this analysis of two dimensional associations as the starting point, a small number of multivariate analyses were carried out. Estimating the relationship between lapses, income, price and enrolments (the latter lagged by two periods) by OLS gave the following equation

$$\underset{}{lapse_t} = \underset{(.707)}{.907\ lcexp_t} + \underset{(.235)}{1.006\ lprice_t} + \underset{(.614)}{1.77\ drpi_t} \qquad (9)$$

$$+ \underset{(.067)}{.684\ lenrol_{t-2}} - \underset{(1.15)}{5.45}$$

$T = 1962\text{-}1983 \quad R^2 = .982 \quad \sigma = .1097 \quad DW = 1.49$

Reset F (2,15) = .13 AR F (1.16) = .87

standard errors in parentheses

The specification test indicates no misspecification of functional form and the goodness of fit, as measured by σ, is reasonable for a short series. Although the coefficient of income is insignificant, the sign of the parameter estimates for the inflation rate and price are in the expected

direction However, high correlation between the price, aggregate income and drpi render interpretation of the parameter estimates difficult. When the levels equation is re-expressed in terms of differences, income remains insignificant and that of rpi decreases (the parameter estimates are small and poorly defined). After dropping terms in income and drpi, the estimated equation was

$$\underset{}{dlapse_t} = \underset{(.089)}{.0605}\ lprice_{t-1} + \underset{(.436)}{.961}\ dprice_t - \underset{(.079)}{.077}\ lapse_{t-1} \qquad (10)$$

$$+ \underset{(.046)}{.78}\ denrol_{t-2} + \underset{(.118)}{.019}\ enrol_{t-3}$$

T = 1964-1983 σ = .1214 DW = 2.15

Reset F(2,13) = .03 AR F(1,14) = .20

Arch F(1,13) = 1.37 Normality $X^2(2)$ = .44

standard errors in parentheses

The functional form of the model does not appear to be mis-specified and the goodness of fit reasonable close (though the fit is better for the second half of the period than the first). The coefficient on change in price is positive, and significantly different from zero, though it has a large standard error. The coefficient on the change in enrolments (lagged two years) is positive and well defined. In the short run therefore, lapses are associated positively with changes in price and recent changes in enrolments. However, as no levels terms are significant, despite the CRDW statistics in equation (9), we are unable to determine whether any long run relationship exists between income, price and lapses.

This analysis does not show that lapses are unaffected by economic variables. What it does show is that aggregate income, and perhaps also the inflation rate and average price, do not appear to be long run

determinants of the aggregate level of lapses, despite the apparent close two dimensional relationships between the levels of price and lapses and income and lapses. (It also indicates the danger of drawing conclusions from two dimensional graphs used in isolation from other analyses.) However, the lack of relationship between aggregate income and average price and lapses does not preclude a long run relationship between lapses and more disaggregated economic factors, such as changes in transitory income, changes in the relative price of insurance and changes in the relative attractiveness of private health care. (It is also interesting to note that analysis of lapses suggests, as the analysis for total subscriptions, that the impact of price is more important than the impact of aggregate income.)

The analysis of new enrolments found some indication in the latter part of the period of a negative relationship between change in price and the change in new enrolments, but the whole period plot of the two variables against each other was fairly scattered. The specification tests for estimates of the relationship between enrolments, price, aggregate income, inflation expressed in either levels of first differences and lagged level terms indicated functional form misspecification for all models estimated. In the light of these tests and the results of the lapses analysis, no further analysis was undertaken.

While it has not been possible to determine whether lapses and new enrolments are functions of price and income, or are responsive to changes in other factors, this section of the analysis has indicated that the nature of the health insurance market may have altered in the late 1970s. Prior to the late 1970s, the market grew at a fairly steady, if unspectacular, rate. The fluctutations in total number of subscriptions,

lapses and new enrolemnts are relatively small. After this date, the movements in these series are larger and show a greater association with chages in average price. It may be that with the entry of more aggressive for profit insurers and the increase in the size of the private health sector, economic factors play a larger role in the determination of purchasing behaviour. However, in the absence of a longer time series, we are not able to take the investigation further.

CONCLUSION

The nature of the private health care insurance industry in Britain is poorly understood in much of the policy debate and is inadequately analysed from an economic perspective. This paper has sought to mitigate these problems by examining the supply and finance of this sector and by presenting estimates of the aggregate price elements of demand for private health care insurance.

It has been shown that the role of the private health care sector is very specialised (predominantly cold elective surgery) and whilst its turnover is small relative to the NHS, in relation to the particular sub-market in which it operates its role is very significant with over 1 in 6 cold elective procedures in England being provided in the private sector. The finance for this activity is largely from insurance (over £580 million in 1987). Perhaps as many as 30 percent of private patients pay cash and thus total expenditure was about £750m in 1987. The insurers have, during the 1980s, experienced considerable cost containment problems as hospitals and doctors have sought to enhance their incomes. These actions have forced insurers to raise their premia, but have also induced insurers to develop various policies, ranging from moral suasion to preferred provider

provider organisations (Crusader and Nuffield Hospitals), to control costs.

The estimate of the price elacticity of demand for private insurance presented in this paper indicates that providers could be quite bold in increasing their prices as demand responses are low in the short run. Thus whilst the insurers may be anxious to control costs to enhance or defend market share, such policies may have limited effects on aggregate market size. Indeed providers, such as doctors, will be able to use their monopoly powers to raise their fees with little concern for reduced demand for their services. The financers appear, in the short run and in aggregate, to be price takers, rather than price makers of the services they purchase on behalf of the customers. However these conclusions must be tentative for a number of reasons. Firstly it is essential to note that whilst the estimate point of the price elasticity is statistically significant and less than one, the interval estimate is quite large and the long run relationship between prices and lapses would appear to be greater than one. Secondly the magnitude of the price elasticity of demand in the component parts of the health insurance industry may differ from that of the aggregate market. It is not possible to estimate the price elasticities for each of the corporate, employer and individual markets separately because of the absence of a publicly available data set of adequate length.

However, despite the absence of such data it is possible to speculate about the price elasticity characteristics of these components of the total market. From published data it is evident that the growth in the individual market is very limited. From Propper and Eastwood's data (Propper and Eastwood (1989) it seems that individual buyers have a long time horizon, reconsider their decision infrequently and thus their price elasticity may be low. It is likely that the price elasticity for the

corporate sector of the market will be higher than that for the individual because managers may manipulate employee reward packages in line with short term changes in the labour market. Ideally these hypotheses should be tested empirically and hopefully this will be possible in due course. The data are available within companies but there is no evidence that detailed econometric methods are being used to explore them.

The NHS Review is likely to lead to increased competition both within the private sector and between the public and private sectors. The effects of such competition can only be appreciated if the behaviour of the private health care sector is more fully understood. Hopefully this paper will encourage others to do more detailed work not only on price elasticities at a disaggregated level but also on the behaviour of the suppliers of hospital and medical services.

REFERENCES

Dolado J and Jenkinson T (1987) 'Cointegration: A Survey of Recent Developments', University of Oxford Applied Economics Discussion Paper 39.

Engle R F and Granger C W J (1987) 'Cointergration and Error Correction: Representation, Estimation and Testing' Econometrica 55, 2, 251-276.

Fitzhugh Directory (1987) Directory of Independent Hospitals and Provident Associations, Health Care Information Services, London.

Gilbert C (1986) 'Professor Hendry's Econometric Methodology' Oxford Bulletin of Economics and Statistics 48, 3, 283-307.

Grant C (1985) 'Private Health Care in the UK; A Review' Economist Intelligence Unit Special Report, 207.

Green D (1988), Everyone a Private Patient, Institute of Economic Affairs Hobart Paper 27, London.

Griffiths R (1988) Community Care: an agenda for action, HMSO, London.

Hall S G (1988) 'An Application of the Granger and Engle two-step estimation procedure to UK aggregate wage data' Oxford Bulletin of Economics and Statistics 48, 3, 229-40.

Hall S G (1987) 'Dynamic Modelling and Cointegration', mimeo, Bank of England.

Harvey A C (1981) The Econometric Analysis of Time Series, Oxford: Philip Allan.

Hsio, W.C., Braun, P., Yntema, D., and Becker, E.R. (1988), 'Estimating physicians' work for a resource based relative-value scale,' New England Journal of Medicine, 319, 13, 835-41.

Laing W (1987) Laings Review of Private Health Care 1987, Laing and Buisson Publications Ltd, London.

Laing W (1988) A Review of Private Health Care 1988, Laing Publications, London.

Laing, W. (1988). Laing's Review of Private Health Care 1987. Laing Publications, London.

Lee Donaldson Associates (1981) UK Private Medical Care.

Maynard A (1982) 'The Private Health Care Sector in Britain' in G McLachlan and A Maynard (eds) The Public/Private Mix for Health, Nuffield Provincial Hospitals Trust.

Maynard A and Smith J A A (1983), The Elderly: who cares? who pays?, Nuffield York Portfolio 1, Nuffield Provincial Hospitals Trust, 1983.

Manning et al (1987), 'Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment' American Economic Review 77, 251-277.

Nicholl, J.P., Beeby, N.R. and Williams, B.T. (1989), 'The role of the private sector in elective surgery in England and Wales in 1986', British Medical Journal.

Oxford Bulletin of Economics and Statistics 48, 3 (1986) [Special issue on cointegration].

Propper C (1987) 'An Econometric Estimation of the Demand for Private Health Insurance in the UK', Centre for Health Economics, York University Discussion Paper 24 (forthcoming in Applied Economics).

Propper, C. and Eastwood, A. (1989) 'The reasons for non-corporate private health insurance purchase in the UK: the results of a new survey and an econometric analysis of the determinants of purchase'. Discussion Paper 52, Centre for Health Economics, University of York.

Sargan, J.D. (1964) 'Wages and Prices in the United Kingdom: A Study in econometric Methodology' in P E Hart, G Mills and J N Whittaker (eds) Econometric Analysis for National Economic Planning, London: Butterworths.

Smith, S (1988) 'Modelling the Demand for Private Health Insurance' Paper presented at the Health Economics Study Group meeting, Newcastle Univeristy, Jan 6-8 1988.

illiams, B.T., Nicholl, J.P., Thomas, K.J. and Knowelden, J. (1985). difference in length of stay for surgery in the NHS and the private sector in England and Wales, British Medical Journal, 290, 978-80.