

Factors Influencing Healthcare Spending in Singapore: A Regression Model

Seng Lee Huang

School of Mechanical and Manufacturing Engineering
Singapore Polytechnic, Singapore
Email:leehuang@sp.edu.sg

Abstract

In this paper, historical records of health care expenditure from 1960 to 2001 were gathered and a regression model was developed using ordinary least square technique. The regression model is developed at an aggregate level which helps to provide some insights on the factors affecting health care spending. The data analysed by using the PH-Stat2 Software, the program will suggest alternative models and the most suitable model is chosen, verified and evaluated for its aptness.

1. Introduction

Expenditures on health care are rising in many countries and Singapore is no exception. In absolute terms, the aggregate health care expenditure rose from \$0.1 billion in 1961 to \$5 billion (at current market price) in 2001, and government health care operating expenditure rose from \$0.05 billion to \$1.2 billion (at current market price) in 1961 and 2001 respectively, which accounts for 6.7% of the government operating budget in 2001 [1]. In terms of individuals' expenditure, the health care expenditure per capita rose from S\$177 in 1961 to S\$1 380 in 2001 (at 1990 price).

Health care service is manpower intensive, and the rising manpower cost is also a cause for concern. Furthermore, Singapore population is ageing rapidly, and the proportion of population who are 60

years old and older has risen from 4.1% in 1961 to 11% in 2001 and it is expected to reach 27% in 2030 [2].

Even though the Singapore economy has enjoyed rapid growth; the gross domestic product per capita average annual growth rate is 6.3 % from 1960 to 2001. But the government is concerned as health care costs are rising steadily while the population is ageing and the economy is maturing.

In October 1993, the government has spelt out its philosophy, policy and approach to controlling health care costs in a White Paper on "Affordable Health Care", thus setting the direction for the health care system.

The Singapore health care philosophy emphasises on building a healthy population through preventive health care programs and the promotion of healthy living. Emphasis is placed on health education, immunization and health screening for early detection of diseases. The government is committed to provide good and affordable basic medical services to all Singaporeans through the provision of heavily subsidised medical services at the public hospital. But individuals are encouraged to take responsibility for their own health by saving for medical expenses and avoid reliance on state welfare or medical insurance and even in the most heavily subsidised wards they are expected to co-pay. And if the market fails to keep the health care costs down, the government will step in and intervene.

2. Determinants of Health Care Expenditure

Spending on health care is a result of a multitude of factors interacting with one another. In line with the current literature on determinants of health care expenditure and taking into the consideration of health care system and the way it is financed, five key determinants were chosen: gross domestic product, government health operating expenditures, supply of doctors, aging population and the use of Medisave withdrawal for payment of health care expenditures.

Many empirical studies have shown that gross domestic product has a strong influence on health care spending. For example Newhouse in 1977[3], Gerdtham and Jonsson in 1992 [4] and Hitiris and Posnett in 1992 [5], using cross-sectional OECD (Organisation of Economic Cooperation and Development) data, reported that gross domestic product was one of the most important determinants of aggregate health care expenditure. In Getzen's 1990 paper [6], the model that he proposed uses only gross domestic product and inflation factors. In 2000, Getzen's paper [7] suggested that health spending is more a function of income over previous five years than of current. That is gross domestic product growth effectively has impact on private health spending after three to five years.

Another influencing factor is the percentage of elderly people in a country. The old spend more on medical care than the young, and it is, therefore, intuitively correct to assume that the aggregate health care rises with higher percentage of elderly population in the country. Studies made in Japan by Fujino published in 1987 [8] showed that the elderly consume about 3.2 times more medical services than the average person. In addition, studies done in the United States by Murthy and Ukpolo published in 1994[9] showed that it is one of the most significant

determinants of health expenditures.

Government subsidy also influence health care expenditure, in Matteo and Matteos' paper [10], 1998, they showed that the real per capita provincial government expenditures is positively correlated to provincial per capita federal transfer revenues.

In the health care industry, there is asymmetrical information between the doctor and patient; therefore, there is this potential for exploitation when the supplier who also acts as the agent can bring about a level of consumption different from that which would have occurred if the fully informed consumer were to choose freely. Hence the "inducement hypothesis" which states that health care is supply driven that is supply creates its own demand since. Therefore, as the number of doctors grows, the health care consumption increases as more physicians would induce more demand.

The aggregate health care expenditure which is a function of economic wealth of a nation, the amount of government subsidy, the proportion of aging population and , the ratio of doctor to per thousand population and lastly the effect of MediSave scheme. These variables were chosen because a) they reflect the changes in health care expenditure and not changes in inflation, or population or recession and b) good historical data is available for the study.

$$AHE = F (GDP_{(lag 0 - 5)}, GHR, DOC, AGE, MED) \dots \dots \dots (1)$$

AHE = aggregate health care expenditure per capita

GDP_(lag 0 -5) = gross domestic product per capita with a lag of 0 to 5 years

GHR = percentage of government health expenditure to gross domestic product

DOC = doctor per thousand population

AGE = proportion of population ≥ 60 years old

MED = Medisave scheme

The gross domestic product per capita plotted in figure 2, and the percentage of population ≥ 60 years old is in figure 3. The data for doctor per thousand is

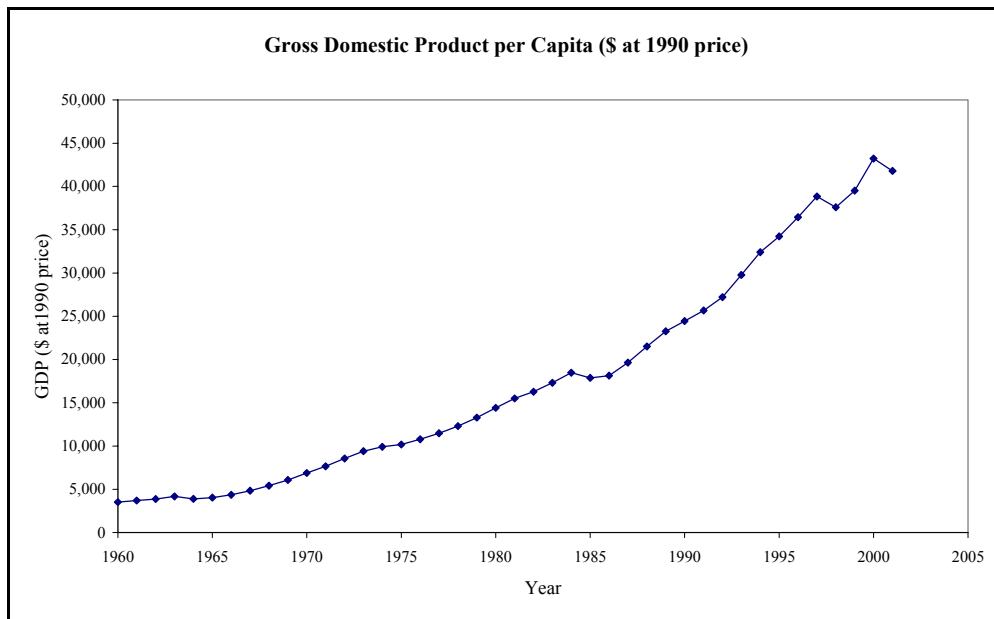


Figure 1: Time Series of Gross Domestic Product per Capita

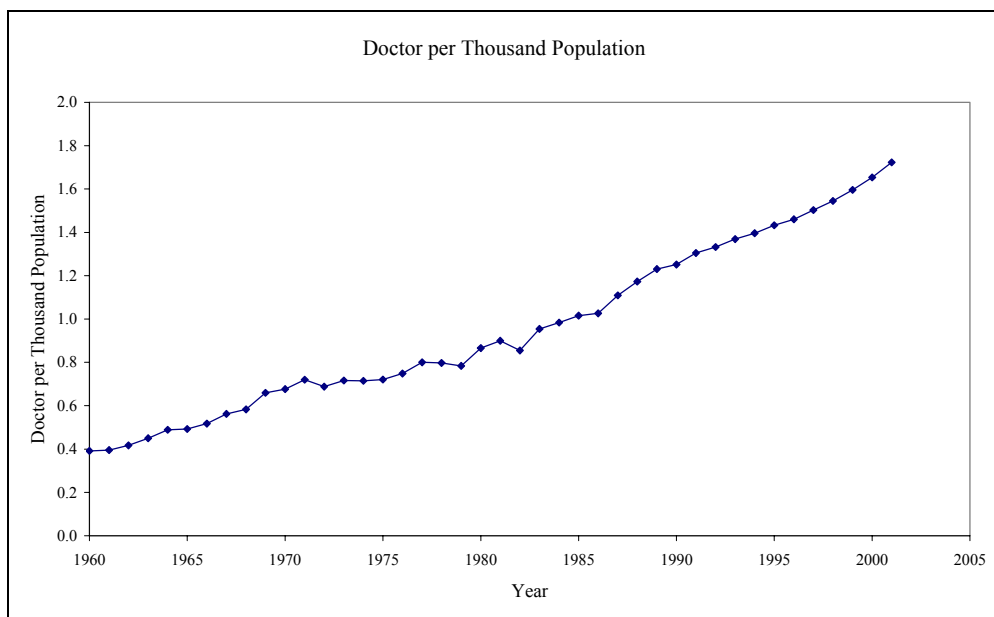


Figure 2: Doctor per Thousand Populations from 1960 to 2001

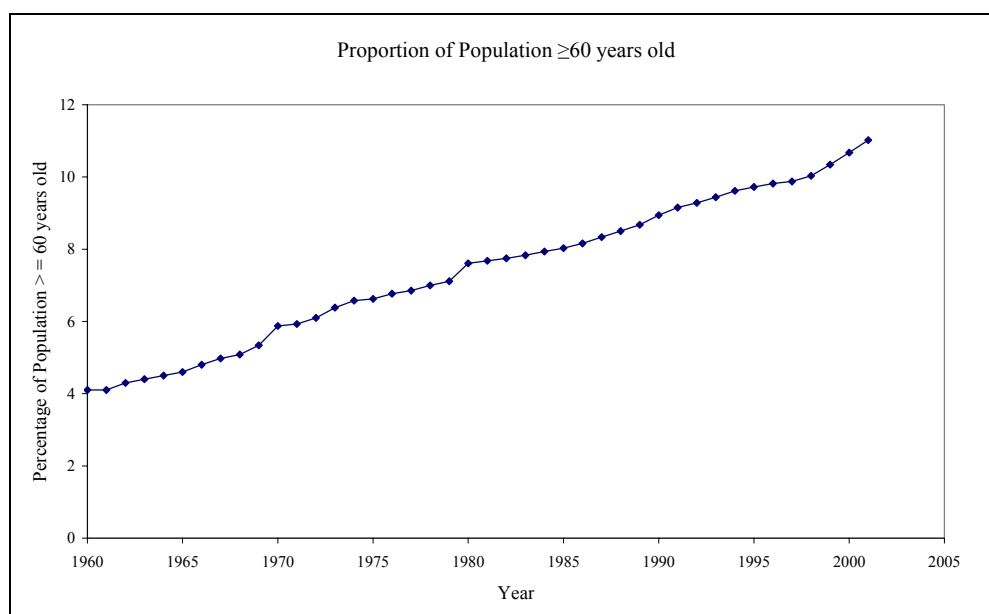


Figure 3: Proportion of Population ≥ 60 Years Old from 1960 to 2001

3. Methodology

The health care expenditure is analysed at an aggregate level, and records of health care expenditures were obtained from Yearbook of Statistics Singapore from 1960 to 2001. The expenditures were adjusted for inflation by using the gross domestic product deflator based on 1990 prices, and then divided by the population to obtain the real expenditures per capita.

Aggregate Health Expenditure = Private Health Expenditure (PHE) + Government Health Expenditure (GHE)

The aggregate health expenditure is defined as the sum of private health expenditure and government health expenditure. Government health expenditure only includes direct and related health expenses by the Ministry of Health and excludes development expenditures and other expenditures related to environmental health. Private health expenditure includes expenditures on medical and pharmaceutical

products, and expenditures on services of physicians and nurses. The aggregate health care expenditure over time is consistent as this definition did not have significant changes over times. The time series of distribution of government health and private health expenditure are shown in figure 4.

The government health expenditure is funded mainly from tax revenue and a small amount from interest of Medifund. Therefore the government subsidy is also constraint by the performance of the country economy and the usage of her resources. The times series for percentage of aggregate health expenditure and government health expenditure to gross domestic product are shown in figure 5.

The private health care expenditure is the amount spent by individuals and organisations. Therefore, it is made up of payments from individuals Medisave account, insurance, organisation group health insurance and individual out of pocket payment. The proportion of Medisave claims to private health care expenditure is shown in figure 6.

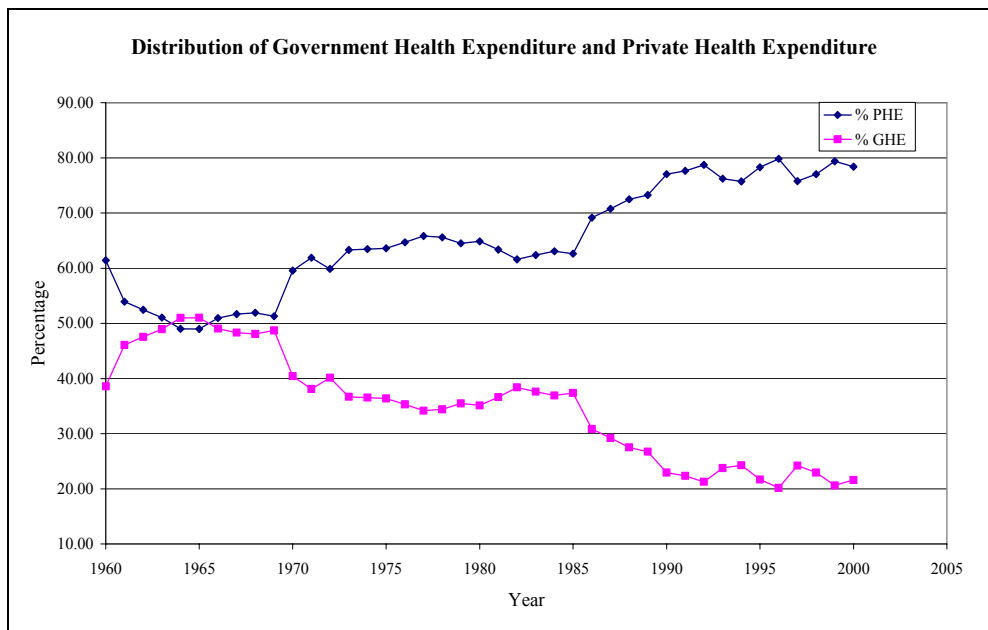


Figure 4: Time Series of Distribution of Government Health Expenditure and Private Health Expenditure

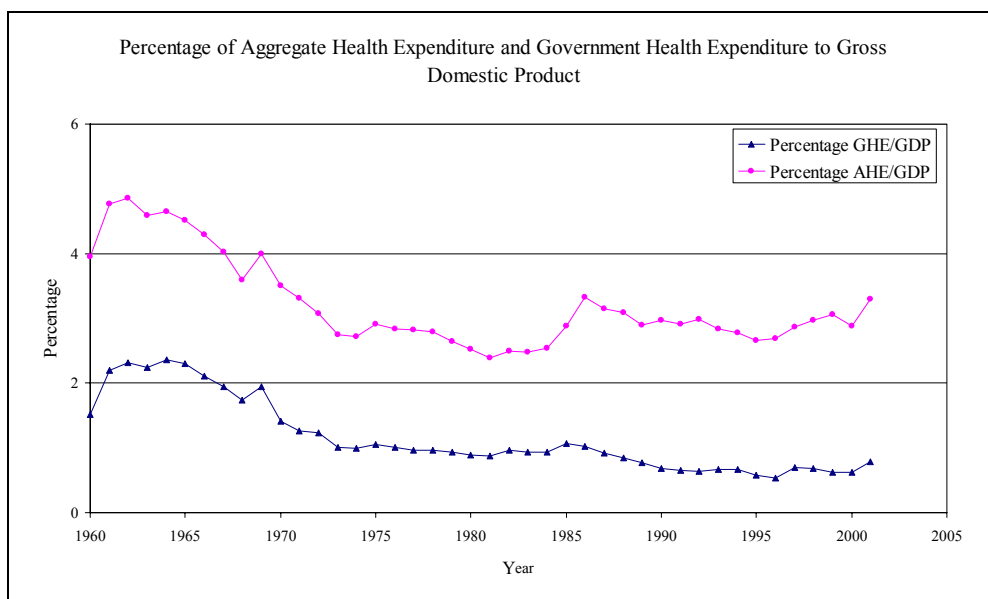


Figure 5: Percentage of Aggregate Health Expenditure and Government Health Expenditure to Gross Domestic Product from 1960 to 2001

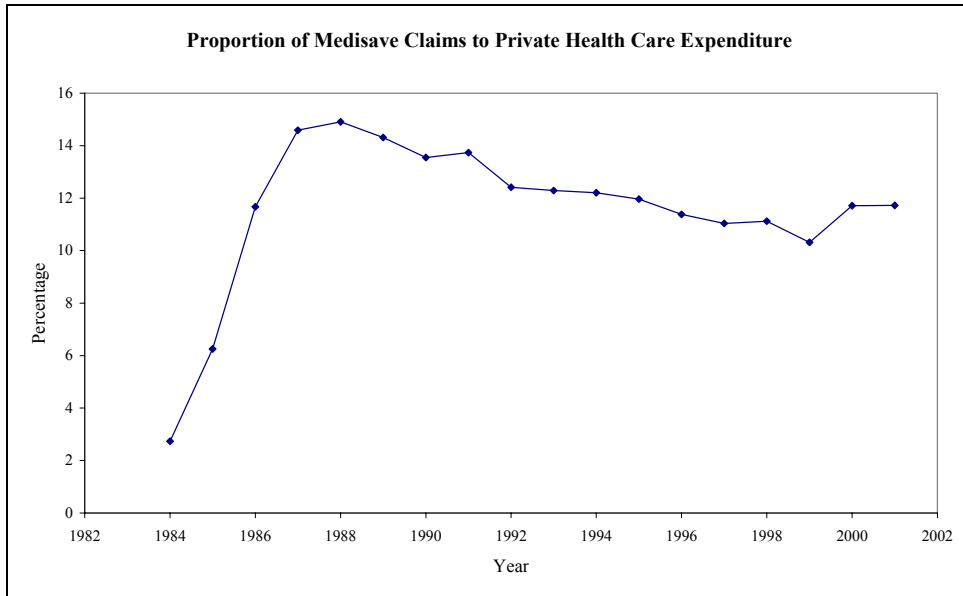


Figure 6: Proportion of Medisave Claims to Private Health Care Expenditure

Now, the production of the healthcare industry in terms of aggregate expenditure is assumed to be in Cobb-Douglas form (see appendix A1), hence equation (1) can be expressed as :

$$AHE = b_0 * GDP^{b1} * GDP_{(lag1)}^{b2} * GDP_{(lag2)}^{b3} * GDP_{(lag3)}^{b4} * GDP_{(lag4)}^{b5} * GDP_{(lag5)}^{b6} * GHR^{b7} * DOC^{b8} * AGE^{b9} * MED^{b10} \dots\dots\dots(2)$$

- AHE* = aggregate health care expenditure per capita
- GDP* = gross domestic product per capita
- GDP_(lag1)* = gross domestic product per capita with 1 year lag
- GDP_(lag2)* = gross domestic product per capita with 2 years lag
- GDP_(lag3)* = gross domestic product per capita with 3 years lag
- GDP_(lag4)* = gross domestic product per capita with 4 years lag
- GDP_(lag5)* = gross domestic product per capita with 5 years lag
- GHR* = percentage of government health expenditure to gross domestic product
- DOC* = doctor per thousand population
- AGE* = proportion of population ≥60 years old

MED = Medisave scheme (replaced by dummy variable in the actual implementation)
bi (*i*=0,1, 2...10) are the elasticities terms

Taking logarithm on both sides, the above equation becomes:

$$\log(AHE) = \log b_0 + b_1 \log(GDP) + b_2 \log(GDP)_{(lag1)} + b_3 \log(GDP)_{(lag2)} + b_4 \log(GDP)_{(lag3)} + b_5 \log(GDP)_{(lag4)} + b_6 \log(GDP)_{(lag5)} + b_7 \log(GHR) + b_8 \log(DR) + b_9 \log(AGE) + b_{10} \log(MED) \dots\dots\dots(3)$$

Since the variables above are trended, therefore to make the model stationary, the difference must be taken [11], [12] hence the equation becomes:

$$\Delta AHE_t = b_1 \Delta GDP + b_2 \Delta GDP_{(lag1)} + b_3 \Delta GDP_{(lag2)} + b_4 \Delta GDP_{(lag3)} + b_5 \Delta GDP_{(lag4)} + b_6 \Delta GDP_{(lag5)} + b_7 \Delta GHR + b_8 \Delta DR + b_9 \Delta AGE + b_{10} \Delta MED \dots\dots\dots(4)$$

where
 $\Delta AHE_t = \log AHE_t - \log AHE_{(t-1)}$
 and

$t = 1, 2, 3 \dots 41$ (1961, 1962, ..., 2001)

ΔMED is replaced by a dummy variable, hence the equation becomes

$$ahe = b_1 * gdp + b_2 * gdp_{(lag1)} + b_3 * gdp_{(lag2)} + b_4 * gdp_{(lag3)} + b_5 * gdp_{(lag4)} + b_6 * gdp_{(lag5)} + b_7 * ghr + b_8 * dr + b_9 * age + b_{10} * dum \dots (5)$$

$dum = 0$ (Medisave scheme not present)

$dum = 1$ (Medisave scheme present)

small letters to represent (log-log)

4. Implementation

There are 10 possible dependent variables: gdp , $gdp_{(-1)}$, $gdp_{(-2)}$, $gdp_{(-3)}$, $gdp_{(-4)}$, $gdp_{(-5)}$, ghr , doc , age , dum . The development of the model is done on the PH2-Stat software which uses the approach of best-subsets regression analysis. This program has a limitation of running a maximum of seven dependent variables at one time. Therefore, the analysis was split into several runs.

For each run, the variables ghr , doc , age and dum remained in the all the runs. For variables gdp , $gdp_{(-1)}$, $gdp_{(-2)}$, $gdp_{(-3)}$, $gdp_{(-4)}$, and $gdp_{(-5)}$ were first grouped into two groups: gdp , $gdp_{(-1)}$, $gdp_{(-2)}$, as one group and $gdp_{(-3)}$, $gdp_{(-4)}$, and $gdp_{(-5)}$ as another group. To determine which of these variables should be chosen for the next run, the expected signs of the coefficients were compared with the results. If it does not agree with the expected signs that variable is dropped. Then, the variables with the right signs were chosen from the two groups for the final run

The best-subsets approach evaluates all the possible regression models for a given set of independent variables. The software will suggest alternative models based on the statistic C_p . This evaluation criterion for competing models is developed by Mallows

[13]. The statistic C_p measures the differences between fitted regression model and the true model. When a regression model with k independent variables contains only random differences from a true model, the average value of C_p is $k+1$, the number of parameters. Thus in evaluating many alternative models, the goal is to find models whose C_p is close to or below $(k+1)$.

In general, the steps involved in building the model are:

1. Choose a set of independent variables to be considered for the regression model.
2. Fit a full regression model that contains all the independent variables so that the variance inflationary factor, VIF (see appendix A2) can be determined for each independent variable.
3. Determine if any independent variables have a $VIF > 5$.
4. Three possible results can occur:
 - a. None of the independent variables has $VIF > 5$, proceed to step 5.
 - b. One of the independent variables has a $VIF > 5$, eliminate the variable and proceed to step 5.
 - c. More than one independent variable has $VIF > 5$, eliminate the independent variable that has the highest VIF and go back to step 2.
5. Perform the best-subsets regression with the remaining independent variables to obtain the best models (in terms of C_p) for a given number of independent variables.
6. List all the models that have $C_p \leq (k+1)$.
7. Choose the best model based on C_p , interpretability and parsimony (perform partial F test).

8. Perform a complete analysis of this model including residual analysis. Depending on the results of the residual analysis, remodel the regression equation. The regression analysis is based on ordinary least square method; therefore residual analyses are performed to check for

any violations of assumptions: constant variance, independence and normality.

5. Results

The results from the various sets of dependent variables are summarized below.

Model	1.1	2.1	4.7	5.1
<i>gdp</i>	0.4435#	0.4506#	-	0.4532#
<i>gdp</i> ₍₋₁₎	-	-	-	-
<i>gdp</i> ₍₋₂₎	-	0.2428	-	0.2346
<i>gdp</i> ₍₋₃₎	-	-	0.0103	-
<i>gdp</i> ₍₋₄₎	-	-	-	-
<i>gdp</i> ₍₋₅₎	-	-	-	-
<i>ghr</i>	0.3194**	0.3187**	0.3236**	0.3201**
<i>doc</i>	-	-	-	-
<i>age</i>	-	-	0.7040*	-
<i>dum</i>	0.0107*	0.0128#	0.0069	0.0123#
<i>R Square</i>	0.4179	0.4541	0.4457	0.4535
<i>Adj R Square</i>	0.3694	0.3899	0.3764	0.3872

(Notes: significant at $\alpha = 0.01^{**}$, $\alpha = 0.05^{\#}$, $\alpha = 0.10^*$)

Table [1]: Summary of Regression Results

Comparing the regression results of models 1.1, 2.1, 4.7 and 5.1, the model 2.1 is chosen as it has the highest adjusted R square value. The variance inflationary factor is checked again for model 2.1, and all the values are below 1.4 which is not very much greater than one, and the Durbin Watson statistic test = 2.15 which is greater than $d_{u,0.01} = 1.72$ ($k=4$, $n=38$), which means that there is no evidence of positive autocorrelation among the residuals.

The assumptions of constant variance, independence and normality were checked and residual plots were plotted. It is observed that the plots were random, and therefore do not violate the above assumptions.

Hence, the model is :

$$ahe = 0.4506 \text{ gdp} + 0.2428 \text{ gdp}_{(-2)} + 0.3187 \text{ ghr} + 0.0128 \text{ dum}$$

Since the equation is in log-log form, this would make the analysis easy. The elasticity of gross domestic product per capita is 0.69 and the elasticity of percentage of government health care expenditure to gross domestic product is 0.32 and Medisave scheme elasticity is 0.01. Assuming that all the dependents grow at a rate of 1% annually, this would result in a 1.0249% growth in aggregate health expenditure.

Based on the equation obtained from the least square technique, a simulation run was done for 1960 to 2001; the result of the run is shown in the figure below. The predicted result was very close to the actual.

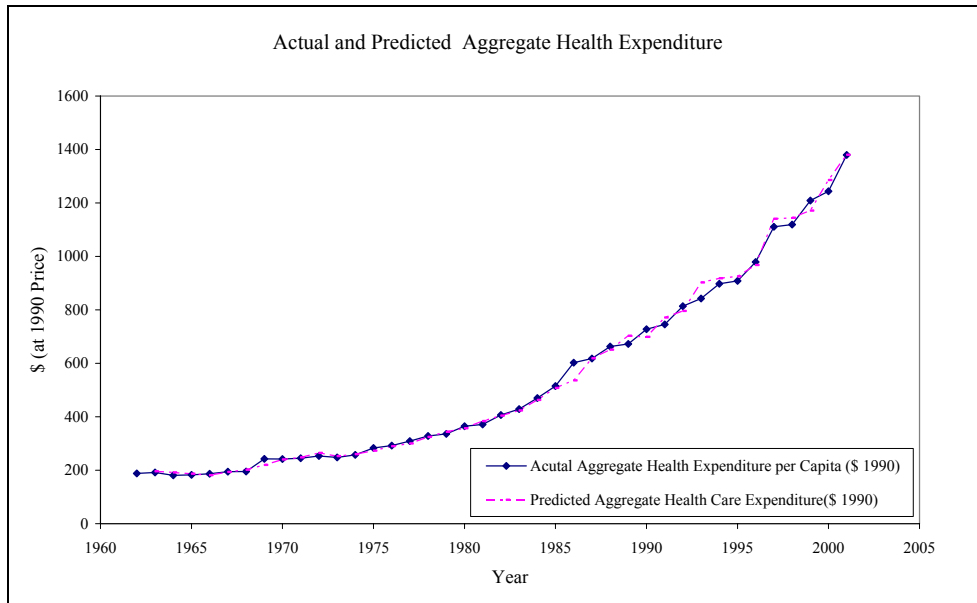


Figure 7: Actual and Predicted Aggregate Health Care Expenditures

6. Conclusion

The key determinants are growth in gross domestic product and percentage of government health expenditure to gross domestic product which contribute 68% and 31% in growth of aggregate health care expenditure respectively. The other factors like the ageing population and the number of doctor per thousand populations do not significantly affect the growth of health care expenditure.

Medisave scheme contributes about 1 % of growth in expenditure, therefore it not an effective tool for cost containment, however it is an importance means of alternative financing resource.

Using the least square ordinary technique coupled with the PH2-Stat software above to analyse health care expenditure has indeed provided some insight on what are the factors that will impact the growth in health care expenditure, the government control on resources spent on health care as well as the performance of the economy are important determinant that cannot be ignored.

7. References

- [1] *Yearbook of Statistics Singapore*, 1960 to 2001, Department of Statistics, Singapore
- [2] Ministry of Health Singapore website, <http://www.gov.sg/moh> [Accessed June 2003].
- [3] Newhouse, J.P., 1977. "Medical Care Expenditure: A Cross National Survey". *Journal of Human Resources*, 12(1), 115-125.
- [4] Gerdtham, U. G. and Jonsson, B., 1992. "An Econometric Analysis of Health Care Expenditure, A Cross Section Study of OECD Countries". *Journal of Health Economics*, 11(1), 63-84.
- [5] Hitiris, T. and Posnett, J., 1992. "The Determinants and the Effects of Health Expenditure in Developed Countries". *Journal of Health Economics*, 11(2), 173-181.
- [6] Getzen, T.E. , 1990. "Macro Forecasting of National Health Expenditures. *Advances in Health Economics and Health Services Research*", 11, 27-48
Neter, J., M.Kuter, C. Natchtsheim, and

- W. Wasserman, 1996. *Applied Linear Statistical Models*. 4th ed. Homewood, IL: Irwin.
- [7] Getzen, T.E., 2000. "Forecasting Health Expenditures: Short, Medium and Lon(Long) Term". *Journal of Health care Finance*, 26(3), 56-72.
- [8] Fujino, S., 1987. "Health Economics in Japan: Prospects for the Future". In: Smith G.T., ed., *Health Economics: Prospects for the Future*. London; New York: Croom Helm.
- [9] Murthy, N.R.V. and Ukpolo, V., 1994. "Aggregate Health Care Expenditure in the United States: Evidence From Cointegration Tests". *Applied Economics*, 26(8), 797 – 802.
- [10] Matteo, L.D., Matteo, R.D., 1998. "Evidence on the Determinants of Canadian Provincial Government Health Expenditures:1965-1991". *Journal of Health Economics*, 17(2), 211-228.
- [11] Getzen, T.E., 2000. "Forecasting Health Expenditures: Short, Medium and Lon(Long) Term". *Journal of Healthcare Finance*, 26(3), 56-72.
- [12] Getzen, T.E., 1992. "Population Aging and the Growth of Health Expenditures". *Journal of Gerontology: Social Sciences*, 47(3), S98-104.
- [13] Thomas, R.L., 1985. *Introductory Econometrics: Theory and Applications*. London; New York: Longman.
- [14] Levine, D.M., Stephan, D., Krehbie, T.C., Berenson, M.L., 2002. *Statistics for Managers using MicroSoft Excel*. 3rd ed. Prentice Hall.
- [15] Snee, R. D., 1973. "Some Aspects of Nonorthogonal Data Analysis, Part I. Developing Prediction Equations". *Journal of Quality Technology*, 5(2), 67-69.

Appendix

Appendix 1: Cobb-Douglas Production Function¹

The production function that has been most frequently used in empirical work is the Cobb-Douglas production function, $Q = AK^\alpha L^\beta$

where parameters α and β measure the elasticities (assumed to be constant, values between 0 and 1) of output with respect to capital K and labour L respectively. Q is the aggregate output and parameter A is a constant term.

The above equation is non-linear, however by taking logarithms on both sides of the equation, we obtain a linear equation:

$$\log(Q) = \log A + \alpha \log K + \beta \log L$$

Therefore, the least square technique can be used to estimate the parameters. The non-linear original variables are transformed and the new variables and parameters become linear.

Now, taking the partial differentiation of the Cobb-Douglas production function yields:

$$\frac{\partial Q}{\partial K} = \alpha AK^{\alpha-1} L^\beta = \alpha Q/K$$

$$\alpha = (\partial Q/Q) * (K / \partial K),$$

which corresponds to the partial elasticity of Q with respect to K.

$$\text{Similarly, } \beta = (\partial Q/Q) * (L / \partial L)$$

which corresponds to the partial elasticity of with respect to L.

¹ Source: Introductory Econometrics Theory and Applications by R L Thomas [13]

Appendix 2: Variance Inflationary Factor [14]

One of the important problems in the application of multiple regression analysis involves the possible collinearity of the independent variables. This condition refers to situations in which some of the independent variables are highly correlated with each other. In such situations, collinear variables do not provide new information, and it becomes difficult to separate the effect of such variables on the dependent response variable.

One method of measuring collinearity uses the variance inflationary factor (VIF) for each independent variable. VIF is defined as:

$$VIF_j = \frac{1}{1 - R_j^2}$$
where R_j^2 is the coefficient of multiple determination of dependent variable X_j with all other X variables

If the set of dependent variables are uncorrelated, then VIF_j is equal to 1. If the set of dependent variables is highly intercorrelated, then VIF_j might even exceed 10. However, some statisticians [15] have suggested a more conservative criterion that would employ to least squares regression, that is VIF_j should not exceed 5.