New Evidence on an Old Issue: Is Health Care a Luxury?

By

Gang Zhao

2773911

Major Paper presented to the

Department of Economics of the University of Ottawa

in partial fulfillment of the requirements of the M.A. Degree

Supervisor: Professor Kathleen Day

ECO 7997

Ottawa, Ontario
January 2004

Abstract

In recent years, health care expenditures have grown very fast in most developed countries. A number of researchers have tried to explore the forces that drive the health care expenditure and how to control it. Many researchers have studied the issue of whether health care is a luxury because if health care is a luxury, then a government cannot control health care expenditure effectively by reforming its health care policy. Working with different methods on different data sets, these authors obtain different results. In this paper, I re-examine this issue with cross-national data. I include more countries in my data set and I add some new independent variables to the model to explain health care expenditure. My results confirm the conclusion that health care is a luxury.

Keywords: Health care expenditure; Income elasticity; Cross-national analysis.

New evidence on an old issue: Is health care a luxury?

1. Introduction

There are two striking trends that have appeared in health care expenditure data in the last half century. First, for wealthier countries, as their Gross Domestic Product (GDP) grew, health care expenditures (HE) grew even faster. In the member countries of the Organisation for Economic Co-operation and Development (OECD), the ratio of health care expenditures to GDP became increasingly large (see Table 1). For instance, in 1960 Canada's total expenditure on health care represented 5.4% of its GDP, while by 2000, its total expenditure on health care had increased to 9.2% of its GDP. These huge health care expenditures became a heavy burden on the governments of OECD countries. Second, in the poor countries of Africa, during the 1980s national income decreased and health care expenditures dropped dramatically, worsening the already terrible health situation there.

These two phenomena interested a great number of researchers. Since these two trends are seemingly consistent with the hypothesis that health care is a luxury instead of necessity, some researchers tried to explain these phenomena with elasticity analyses; that is, they tried to examine the income elasticity of health care (ζ)[1] to see if it is larger than one. If the income elasticity of health care is larger than one, then health care is a luxury.[2] In this case, the phenomena mentioned above merely reflect the nature of health care demand. What is more, if health care is a luxury, the ratio of health care expenditures to GDP will become increasingly large with the growth of

[1] Here we use GDP as the indicator of aggregate income and then define the income elasticity of health care as: $\varsigma = \dfrac{\Delta HE / HE}{\Delta GDP / GDP}$.

[2] Luxuries are goods with an income elasticity larger than one.

GDP. In this case, a government should focus on how to allocate health care expenditure in a more efficient way rather than trying to control the increase in health care expenditure. Therefore, whether or not health care is a luxury became a very interesting question not only for academic researchers but also for policy making.

Although many researchers have done a lot of studies on this issue, almost all of them have focused on OECD countries. Among these researchers, some authors used cross-sectional data. However, the small sample size made their results statistically unreliable. Others researchers used panel data which included time series of health care expenditures and GDPs. Although using panel data can increase sample sizes, regressing one time series on others may result in spurious regressions: in other words, the strong correlation between health expenditure and GDP may be due to the common trend of the two variables instead of a real economic relationship. On the other hand, with different data sets, their results are different. Consequently, there is no clear conclusion on this issue yet.

In my paper, I carry out a purely empirical analysis of the relationship between health care expenditure and GDP using data from 171 countries, including 30 OECD member countries and 141 developing countries. To the best of my knowledge, my sample size is the largest among the cross-national studies on this issue. This larger sample size increases the reliability of my results. In an attempt to deal with heterogeneity among different countries, I include some dummy variables related to the geographic location of each country. I also include an independent variable that proxies the population health status, which was not discussed by former researchers. My conclusions provide new evidence that health care is a luxury.

The remainder of my paper is organized as follows. Section 2 is a literature review. Section 3 is a theoretical analysis of the demand for health care. Section 4 describes the data and the possible model forms and provides an outline of my empirical analysis, including the tests I will use, the purpose of these tests and the methods of a few of these tests. Section 5 reports the test results and regression results. Section 6 presents the conclusions of my paper, the limitations of my analysis and suggestions for further studies.

2. Literature Review

2.1 Cross Section Studies

In Newhouse (1977), the author tries to unravel the relationship between per capita health spending and per capita GDP using 1971 data for thirteen developed countries. Estimating a single linear equation with Ordinary Least Squares (OLS), the author obtains the following two results. First, the variation in per capita GDP can explain more than 90 percent of the variance in per capita health spending. Based on this finding, he concludes that GDP is the most important determinant of health spending and that other factors such as price factors and the organization of the delivery system are not important in order to explain health spending. Second, taking GDP as the proxy for income, the income elasticity of health spending is larger than one; therefore, Newhouse concludes that health care is a luxury good.

Leu (1986) argues that some institutional factors may affect health expenditure. In his paper, he includes the share of public financing of health in total health spending (SPF) to measure the extent of public sector provision of health services. He justifies the inclusion of this variable on the grounds that government financing of health care raises moral hazard problems; that is, people will want to spend more, the greater the share of health spending covered by government. Meanwhile, the author also uses a dummy variable for the National Health Service (NHS).[3] Specifically, the author lets NHS be 1 for those countries having centralized budgetary control and 0 for other countries. Assuming a log-linear functional form and using the OLS method of estimation, the author analyzes the data covering 19 OECD countries in 1974 and confirms that per capita GDP is the major determinant of health expenditure and the income elasticity of health expenditure is above unity. What is more, the author finds that the effects of SPF and NHS on health expenditure are also statistically significant.

Although these two authors use data from different countries and different years, they obtain similar results. They find that the relationship between per capita health expenditures and per capita GDP is quite stable and the conclusion that health care is luxury good seems very robust.

Should the case be closed? Not at all. Parkin et al. (1987) raises two questions: whether the choice of conversion factors used to convert health expenditures and GDP denominated in national currencies to a common currency might affect the result of the analyses significantly, and whether the estimate of income elasticity varies across different functional forms. To answer these two questions, the authors use data from 18 OECD countries in 1980 and convert the health expenditure per capita of each country with both exchange rates and GDP Purchasing Power Parities (GDP PPPs).[4] What is more, the authors do OLS estimation of different functional forms representing the relationship between per capita health expenditures and per capita

[3] Leu does not specify why he included NHS in his paper.
[4] GDP PPPs are Purchasing Power Parities based on the prices of a basket of all goods in different economies. One can get GDP PPPs for OECD member countries from OECD databases.

GDP. Using a linear functional form and converting health expenditures per capita and GDP per capita with exchange rates, Parkin et al. reach results similar to those of Newhouse (1977). However, when the authors use PPP factors to convert the same data set and regressed per capita health care expenditure on per capita GDP, they obtain a different result. Furthermore, they conclude that the income elasticity of health care is not stable across different functional forms: linear, semi-log, exponential and log-linear (see Table 2).

This seemingly simple matter of conversion factors stimulated a serious discussion thereafter. Gerdtham and Joensson (1991) used data for 22 OECD countries in 1985 to examine the issue of conversion factor instability again. To do so, the authors convert health expenditure per capita using exchange rates, GDP PPPs and health PPPs[5] respectively. Meanwhile, the authors convert GDP using exchange rates when they use exchange rates to convert health expenditures and using GDP PPPs when using either GDP PPPs or health PPPs to convert health expenditures. Using OLS and a log-linear functional form, the authors regress the three sets of values of health spending on GDP per capita. The authors report that the income elasticity of health expenditure is 1.24 with exchange rates as conversion factors and is 1.43 with both GDP PPPs and health PPPs as conversion factors. No matter what conversion factors the authors used, their estimates of the income elasticity are significantly larger than one. Therefore, they conclude that no matter which conversion factors are used, health care is a luxury good.

[5] Health PPPs are Purchasing Power Parities based on the prices of a basket of health related goods and services. One can get health PPPs for OECD member countries from OECD health databases.

Then another researcher, Murthy, in his paper published in 1992, argues that there might be heteroscedasticity[6] and outliers within the data set Gerdtham and Joensson (1991) have used. The author first regresses per capita health expenditures on per capita GDP using the same data, same conversion factors, same functional form (Log-linear) and same estimation method. However, in addition to the OLS *t*-values for the estimates, the author also reports the White-corrected *t*-values for them. Observing a large difference between the OLS *t*-values and the White-corrected *t*-values, he concludes that there must be heteroscedasticity in the error terms. In addition, after carrying out the DFBETAS test, DFFITS test and RSTUDENT test for outliers, Murthy concludes that some observations, such as Turkey, Portugal and Sweden, are outliers.[7] To decrease the effect of outliers, he uses the Least Absolute Error (LAE)[8] estimator instead of OLS to re-estimate the income elasticity of health care. His conclusions include: (1) the income elasticity is not the same when using GDP PPPs and health PPPs as conversion factors, and therefore there exists conversion factor instability in international comparisons of health care expenditure; (2) converting per capita health expenditure by health PPPs, the income elasticity is not significantly greater than unity; therefore, "health care is a necessity instead of a luxury" (Murthy 1992, 187).

---

[6] Murthy did not specify the nature of the heteroscedasticity and did not do any tests for the existence of heteroscedasticity. However, since the existence of heteroscedasticity is common in cross-country analysis, his concern is reasonable.

[7] I describe the details of these tests in section 4.

[8] In his paper, the author did not mentioned why he decided to use the LAE instead of the OLS to do estimation. However, Gerdtham and Joensson (1992) pointed out that Murthy' s idea may be: "since LS is based on squared deviations from the regression line $\sum(y_i - \alpha - \beta x_i)^2$, there will be a tendency for the coefficients to be influenced greatly by extreme deviation." (Gerdtham and Joensson 1992, 189). In contrast, the LAE method Murthy used is based on absolute deviations, so outliers may have smaller influence on estimated coefficients.

Responding to Murthy (1992), Gerdtham and Joensson (1992) argue that based on the Breusch-Pagan and White tests for heteroskedasticity, the hypothesis of homoscedasticity can not be rejected. At the same time, the authors confirm that Murthy's concern about outliers is a good point; it is very important for a cross-sectional analysis of only about 20 countries (small sample size). From this standpoint, Gerdtham and Joensson re-estimate their model with some adjustments. First, the authors introduce a dummy variable to distinguish outliers from the remaining countries; second, the authors exclude outliers from the sample completely. They report that the two adjustments don't change the results significantly and therefore, they insist that there is no conversion factor instability and health care is a luxury.

In summary, a lot of researchers have carried out cross-sectional analyses of the relationship between health care expenditure and GDP, using data from different countries and different years. The common result is that GDP is the most important determinant of health care expenditure. The main arguments focus on: (1) whether the income elasticity of health care expenditure is larger than one or not, that is, whether health care is a necessity or a luxury; (2) whether or not there exists conversion factor instability in cross-sectional analysis on this issue; and (3) what is the best functional form to represent the relationship between per capita health expenditures and per capita GDP. The main limitations of cross-sectional analyses on this issue are that the sample sizes of these studies are so small (less than 30) that the statistical results are not very reliable, and that they assume that all parameters of the relationship are the same for all countries.

2.2 Time-Series and Panel Studies

In recent years, with the development of OECD annual health care statistics for its member countries, some researchers have begun to use panel data to examine the relationship between GDP and health expenditure. For example, Hitiris and Posnett (1992) use a pooled cross-section time-series data set including observations covering twenty OECD countries for 28 years from 1960 to 1987. In their paper, the authors examine the relationship between per capita health expenditure and per capita GDP (taking per capita GDP as the only explanatory variable) by estimating both linear and log-linear models.[9] The estimation results provide evidence that health care is a luxury again. The authors also examined two possible non-income determinants of health care expenditure: the public sector's share of financing and the proportion of the population over 65. Adding these two variables into their original equations, the authors report that the effect of the second one is significant while that of the first one is not.

In OECD (1995), the authors use data from 32 OECD countries for a twenty-year-period and employ a fixed effects model to analyze the relationship between health care expenditure, GDP, and other variables reflecting demographic characteristics, health behaviour and supply-side factors. These variables include the proportion of the population 75 years and over (POP75), the proportion of the population 4 years and under (POP04), the female labour force participation ratio (FPR), alcohol intake (ALCC), tobacco consumption (TOBC), the stock of practising physicians (DOCTCA), the stock of practising physicians with fee-for-service

[9] In their paper, authors did not clarify what estimation method they used. However, with their statement that "the method used to estimate the regression equation takes account of the fact that the

payments (DOCTCA*FFSA), the proportion of in-patient expenditure of total in-patient care beds (TEXMC), the number of public in-patient care beds as a proportion of total in-patient care beds (PUSH), the public insurance coverage of the population (COVERO), the proportion of coverage for in-patient care of state and social security schemes (COVERI), the proportion of coverage for pharmaceuticals of state and social security schemes, the average share of medical care billing paid for by public insurers in ambulatory care sectors (COPAYA), the average share of medical care billing paid for by public insurers in in-patient care (COPAYI) and the average share of medical care billing paid for by public insurers in pharmaceuticals (COPAYP). Estimating a log-linear model with Generalized Least Squares (GLS), they find that besides that of GDP, only the coefficients of TOBC, TEXMC, DOCTCA, DOCTCA*FFSA and PUSH are statistically significant. The income elasticity reported in the paper is 0.74, which is significantly less than one. This finding challenges the conclusion of most former studies that health care is a luxury.

In Hitiris (1997), the author takes the following variables as the determinants of health care expenditure: per capita GDP (GDP), the "population aged 0-19 plus 65 and over as percent of population aged 20-64" (p. 4) (P), the share of health care expenditure in total public spending (G), the rate of inflation (I), and a dummy variable for "differences in countries' health service systems" (p. 4) (D) in his model.[10] The data he uses in this paper consist of observations for the 32 years from 1960 to 1991 for ten EC member states. Based on his estimation results, the author concludes that: (1) the coefficients of all the main explanatory variables are

disturbance terms are found to be cross-sectionally heteroscedastic and time-wise autoregressive," (p. 174), I guess they used GLS or FGLS to do the estimation.

[10] D=1 for countries have a National Health Service (NHS) plan and D=0 otherwise.

significant; (2) GDP is still the most important determinant of health spending; and (3) the income elasticity of health spending is almost equal to one (Hitiris 1997, 5).

There is no doubt that panel studies may increase the sample size dramatically, but are they really superior to cross-section analysis? It is difficult to say. Using time series data to study this issue may cause some econometric problems. It is well known that the stationarity of the data set is a very important assumption underlying OLS. Regressing a non-stationary dependent variable on non-stationary regressors may cause a spurious regression problem, namely, the strong correlation between health care expenditure and GDP may be due to the common trend in the two variables instead of a real economic relationship.

Hansen and King (1996) raise the question of the stationarity of the data series used by Hitiris (1992). Using the augmented Dickey-Fuller (ADF) test, the authors point out that "two-thirds of the variables tested were found to be non-stationary in levels and no one country possessed a data set that is entirely stationary in levels" (p. 130). Therefore, the authors are sceptical of the results of those articles that use panel data. Furthermore, with the Engle-Granger cointegration test, they state that the hypothesis of no-cointegrating relationship among health expenditure and GDP cannot be rejected.

Blomqvist and Carter (1997), who use data consisting of "32 annual observations for 18 countries" (p. 210), employ time series techniques to try to find a long-run relationship between health care expenditure and income. To do so, the authors test for the presence of unit roots in the time series of health care expenditure per capita and GDP per capita. Using the Phillips and Perron (1988) test, the authors accept the unit root hypothesis in all series except the health care expenditure of

Finland. This finding may imply that per capita health expenditure and GDP are non-stationary and thus "some of the conventional methods used in earlier literature to estimate the long-run relationship between them may have yielded misleading results" (p. 213). Then the authors test the cointegrating relationship between health spending and GDP using the Shin (1994) test and point out that the null of cointegration cannot be rejected for any sample country. Furthermore, estimating their dynamic log-linear model of health spending with the Phillips-Loretan (1991) method, the authors report an income elasticity of 0.976, which casts doubt on the conclusion that health care is a luxury.

McCoskey and Selden (1998), however, argue that the ADF test has low power and proposed the use of the "z-bar" test (see Im et al., 1997). What is more, the authors also argue that "the inclusion of a time trend can cause a large loss of power with little improvement in fit" (p. 373). Therefore, the authors omit the time trend in their model and re-analyzed Hansen and King (1996)' s data again. They reject the existence of unit roots in both health care expenditure series and GDP series in the OECD panel data set.

With the same starting point as Hansen and King (1996), Roberts (2000) re-examined Hitiris' (1997) data set and argue that the high statistical value of the coefficient of determination (R^2) and the low value of the Durbin-Watson (DW) statistic imply that there might be a spurious regression problem. In his paper, the author not only employs augmented Dickey-Fuller (ADF) tests and Philips-Perron (1988) tests to examine the existence of unit roots in each country' s series, but also uses the "t-bar test" (see Im et al., 1997) to examine the presence of unit roots in the panel data set. The author concludes that for each country, health care expenditure and GDP are non-stationary. Consequently, the author employs both the Engle-Granger test and the Johansen (1991) test to test for the presence of cointegration between health expenditure and GPD, and the two tests lead to different results: the Engle-Granger test cannot reject no-cointegration while the Johansen (1991) test can.

In summary, we can see that even though panel data can increase the number of observations, it introduces more technical problems. Using time series, especially long time series, may cause a problem of spurious regression, so it is necessary to test the stationarity of each series included. Unfortunately, the methods developed to test for unit roots and cointegration are far from satisfactory so far: all of them have their own problems, such as having low test power or being sensitive to the selection of the number of lags. But as the above review of former research shows, it seems that cross sectional analysis and panel data analysis both have their own limitations.

3. Theoretical analyses of the demand for health care

From the viewpoint of traditional theories of demand, the factors that affect the demand for health care, as well as other common commodities such as appliances and cars, should include price factors, income factors and the tastes of consumers (see Varian 1993). The way that income affects the demand for certain commodities is obvious. When people have more income, they have more purchasing power. Therefore, they will generally have more demand for all normal goods.[11] The only exceptions are inferior goods. Higher incomes make people more concerned about the qualities of commodities. Therefore, consumers will demand less of inferior goods

[11] Normal goods are those goods for which demand increases when income increases.

when their incomes increase. The relationship between income and the demand for certain goods can be shown by an Engle curve.

The effect of price on demand is much more complicated. These effects could come in several ways. First, when other factors remain unchanged, that is, when the demand curve is fixed, the higher the price of a certain commodity is, the lower the quantity of the commodity that is demanded.[12] People will shift their demand for the commodity to its substitutes. This phenomenon is called the substitution effect. Second, an increase in the price of a certain commodity lessens the real income of consumers, holding other factors fixed. Consequently, the decrease in real income will cut people' s demand. This is called the income effect. Third, the price of other goods may affect the demand for certain goods as well. There are two situations in this case: if the price of its substitute increases, the demand for a certain good will increase likewise; if the price of its complement increases, the demand for the good will decrease.

The tastes of consumers are hard to describe and even harder to measure. They may be influenced by countless factors themselves. In addition, these factors vary from one commodity to another and change over time. For example, the medical discovery that drinking tea will help prevent cancer boosted the demand for tea in the Japanese market dramatically. As another illustration, an action movie entitled "Motor Boy" incited a large number of young Chinese men to buy large powerful motorcycles.

[12] Theoretically, there exists a kind of good called a Giffen good, the demand for which falls when its price falls, but it is very hard to find an example in the real world.

Considering health care as a commodity, the factors that affect the demand for it can be catalogued into the three previously mentioned categories. However, due to the characteristics of health care, the factors that affect the demand for health care are very complicated. The way that these factors affect particularly demand may be very different from that of other commodities. A number of authors have discussed the characteristics of health care and the determinants of the demand of health care in many papers or books [Culyer (1971), Sorkin (1975), Anderson (1995), Zweifel and Breyer (1997) and McPake et al. (2002)]. These authors demonstrate many ways in which the market for health care differs from those of traditional goods and services.

Characteristic 1: Externalities.

Externalities in health care mainly stem from the issue of communicable diseases. If one suffers from a communicable disease, he or she will put others around him or her into danger. In this case, health care to the patient benefits not only the patient (the health care consumer), but also other people that the patient may come into contact with. Because of these externalities, health care is largely financed by the government. According to the data set provided in *The World Health Report 2002* by the World Health Organization (WHO), the mean value of the share of government expenditure on health care in total expenditure on health care (SGT) is 62%. One hundred and twenty-two out of 190 member countries have SGTs larger than 50%. This large share of government expenditure, along with different private insurance coverage, reduces the out-of-pocket expenditure of consumers significantly. According to the same data set, the average ratio of out-of-pocket expenditure to total health care expenditure of 190 member countries is only 32%. Since out-of-pocket expenditure is the payment individual consumers make when "buying" health care, a decrease in out-of-pocket expenditure could affect the demand for individual consumers.

Characteristic 2: Derived demand

No one really wants to undergo surgery because surgery always causes temporary pain, or at least temporary inconvenience. People undergo surgery just because they want to recover from illnesses. What they really want is health. The demand for health incites the demand for health care. For this reason, in contrast to most other commodities, for which improvements in quality will generally increase demand, improvements in the quality of health care may decrease the potential demand for further health care.

Characteristic 3: Health care is an investment good and a consumption good

Grossman (1972) pointed out that health care is an investment as well as a consumption good. In his research, the author "assumed that individuals inherit an initial stock of health that depreciates over time" (p. XV) and health care will improve consumers' stock of health. In the light of utility analysis, the author states that "health is demanded by consumers for two reasons" (p. XV). First, consuming health care (treatment and cure specifically) reduces patients' sick days, which are a source of disutility. Second, investing in health care may improve consumers' stock of health and therefore increase consumers' total time for "market and non-market activities" (p. XV). What is more, the higher the wage rate is, the more utility one will obtain when spend time on market activities, leading to the higher demand for health care. Therefore, Grossman includes the wage rate in his empirical model. In addition, he suggests that the consumer's health status, which reflects the "stock" of health, will affect the demand for health care. However, since " the stock of health, like the stock of knowledge, is a theoretical concept, one that is difficult to quantify empirically" (p. 45),[13] Grossman employed another variable TL (which stands for "time lost from market and non-market activities due to illness and injury" (p. 45) and is inversely correlated with health status) in his empirical model.

Characteristic 4: Information and Uncertainty

When people plan to buy a common commodity such as a TV set or computer, they generally have a clear idea of their tastes. For instance, they may prefer a colour TV to a black and white one or they like computers with a Pentium IV processor much more than those with a Pentium II. However, when people "consume" health care, they generally have no idea which kinds of treatments are better and what are the side-effects of the treatments. They also have little knowledge of the price of the health care they receive.

In recent years, due to this characteristic of health care, a lot of researchers [OECD (1995), Culyer (1971), Sorkin (1975), Zweifel and Breyer (1997) and McPake et al. (2002)] have discussed an effect called "supplier-induced demand." That is, the suppliers of health care have the power to induce consumers to consume more health care than necessary. This phenomenon stems from the asymmetrical information available to health care consumers and health care providers. Generally, health care consumers do not have a lot of knowledge about health care. They also do not know what their health problems are and what is the best way to solve their problems. Their consumption of health care, to a great extent, is based on the

[13] Indicators such as health-adjusted life expectancy (HALE) had not been developed at the time Grossman wrote this paper.

suggestions of health care suppliers such as physicians or surgeons. Therefore, if these health care suppliers desire, they could possibly lead the patients to consume more health care than they need. Of course, the magnitude of "supplier-induced demand" depends on the incentive of the supplier to let consumers pay more. For example, when the number of doctors increases while the number of patients remains unchanged, in order to maintain their own incomes, those doctors might have a tendency to let each patient receive more treatment.

Even though the theoretical analysis gives us some hints on searching for determinants of health expenditure, it does not give us many guidelines on finding possible explanatory variables and the causal mechanisms (see Gerdtham and Joensson, 2000). In my opinion, the possible determinants of health expenditure may include income, the relative price of health care and other variables affecting consumers' tastes. First, the theory of demand for conventional goods suggests that my model should include (1) GDP per capita, which represents aggregate income level; (2) a relative price index for medicines and wage rates of health workers, which reflect price factors; (3) the public share of health care expenditure in total health care expenditure (SPT), which not only affects consumers' out-of-pocket spending, but also reflects public intervention. Second, the research of Grossman (1972) suggests that life expectancy (which reflects population health status) and the average wage rate should also be included in my empirical model. Thirdly, the fact that the demand for health care is a derived demand also motivate me to include demographic variables which reflect the age structure of the population in my model. Fourthly, variables reflecting the level of supplier-induced demand (these variables could include the stock of physicians, the number of in-patient beds etc) can also be explanatory variables due to "supplier-induced demand". Lastly, dummy variables to account for differences between countries are also necessary.

Due to the inavailability of data, all the researchers I mentioned in section 2 only include some of the variables among above-mentioned variables. For example, Newhouse (1977), Parkin et al (1987), Gerdtham and Joensson (1991) and Murthy (1992) used only GDP per capita as the single explanatory variable in their papers. Meanwhile, in those papers [Leu (1986), Hitiris and Posnett (1992), OECD (1995) and Hitiris (1997)] attempting to find the missing explanatory variables for health expenditure, the results of their empirical analyses are sometimes controversial. For example, Leu (1986) reported a positive effect of SPT on PHE; in Gerdtham and Joensson (1992), however, the authors reported a negative impact of SPT on PHE. What is more, the underlying theory does not provide any ideas on the functional form of the relationship between health expenditure and its determinants. Therefore, empirical analyses become very important for us to understand this issue well.

4. Data description, Model Specification and Methodology

4.1 Data Description

The sample I use in this paper covers 171 countries that are member countries of both the WHO and the United Nations (UN). My data set includes data for the year 2000 for these 171 countries. (The Appendix lists the names of my sample countries.)

According to the theoretical analysis in section 3, I should try to collect data reflecting the effects of income, price and consumers' tastes on the demand for health care. Since it is extremely difficult to collect data for such a large sample, I just picked one or two variables for each aspect respectively. (The Appendix lists all the variables I included in my models and their data sources.) Most of the variables I selected in my paper were commonly used by other authors who already did research on this issue.

The dependent variable in my model is real health care expenditure per capita (PHE) in level or log form, already converted at source from national currency units to US dollars using GDP PPPs. It reflects the level of consumption of health care. This variable was used by almost all previous researchers [Newhouse (1977) and Leu (1986),[14] Parkin (1987), Gertham and Joensson (1991, 1992), Hitiris and Posnett (1992), Hitiris (1997), Hanson and King (1996), Blomqvist and Carter (1997)]. The data for this variable are from the *World Health Report 2002* by the World Health Organization (WHO).

Nine explanatory variables were included in the estimation. The first one is real GDP per capita (PGDP), also converted at source from national currency units to US dollars by GDP PPPs. This variable represents the income of health care consumers. It too was used by almost all previous researchers. Theoretically, this variable will have a positive effect on PHE: people with more income tend to increase their consumption of all normal commodities. All the research mentioned in Section 2 found that the income elasticity of health care is larger than zero; that is, that health care is a normal good. The main dispute focuses on whether health care is a luxury or a necessity, the question that is the main concern of my paper. The data for this variable are collected from the *Human Development Reports 2002*.

The second explanatory variable in my estimation is the public share of health care expenditure in total health care expenditure (SPT), in percentage terms. This variable has an effect similar to price. The increase in the public share of health care expenditure will reduce the out-of-pocket expenditure of people and therefore reduce the price of health care to consumers correspondingly. However, since this variable is not a direct indicator of price and it includes some other information such as the government' s power to control the health system, the effect of this factor on health care expenditure is hard to forecast. Leu (1986) reported a positive effect of SPT on PHE. In Gerdtham and Joensson (1992), however, the authors reported a negative impact of SPT on PHE. The data for SPT is collected from the *World Health Report 2002* by the WHO.

The third explanatory variable is the percentage of aged people (60 and older) in the total population (SOP). This is a demographic variable used by Leu (1986), Gerdtham and Joensson (1992), Hitiris and Posnett (1992), Hitiris (1997) and Roberts (2000).[15] Since older people generally have worse health than younger people, this variable may have a positive effect on PHE. The data for this variable is collected from the website of the WHO. The fourth explanatory variable is health-adjusted life expectancy (HALE). This variable reflects the health status of the whole population of a country.[16] It is surprising that it seems that no researchers had included this

---

[14] Newhouse (1977) and Leu (1986) used nominal health care expenditure converted by exchange rate.
[15] The demographic variables are different from one paper to another. For example, Leu (1997) used the population over 60, 65 and 75 years of age while Gerdtham and Joensson (1992) chose the population over 65 years old. However, the idea is the same: older people may be in worse health. Therefore, they need more health care.
[16] According the definition given by the World Health Organization, "healthy life expectancy (HALE) is based on life expectancy (LEX), but includes an adjustment for time spent in poor health. This indicator measures the equivalent number of years in full health that a newborn child can expect to live based on the current mortality rates and prevalence distribution of health states in the population." (WHO official website: http://www3.who.int/whosis/hale/hale.cfm?language=en) With this definition, we can see that HALE not only includes information on current mortality rates but also information about health status in the population. On the other hand, life expectancy (LEX) only reflects mortality rates. Therefore, HALE may reflect population health status better than LEX.

variable in their estimations so far. In my opinion, health status may affect the tastes of health care consumers in a significant way. Ordinarily, healthy people may have less demand for health care. Therefore, this variable should have a negative impact on health care expenditure. I collected the data for HALE from the website of the WHO.

Table 3 lists some descriptive statistics of the five variables I mentioned above. Looking at Table 3, we can see that the means of the five variables, especially those of PHE and PGDP are quite different across different regions. For example, the mean value of PHE for African countries is 135, which is less than one-tenth of that of European countries. This fact motivates my decision to include dummy variables to account for geographic differences between the sample countries. At the same time, we can see that regions with a higher mean value of PGDP generally have a higher mean value of PHE. For example, European countries as a whole have highest mean value of both PGDP and PHE. North American countries have the second highest mean value of both. Meanwhile, African countries, which on average have the lowest PGDP, also have the lowest PHE. This feature reveals the strong relationship between PGDP and PHE.

To account for these regional differences, I include five dummy variables to account for geographic differences between the sample countries. Being close to each other, having similar cultures and environments and frequent human flows among neighbouring countries, may cause countries within the same geographic region to have a similar level of health expenditure. For example, this summer when Severe Acute Respiratory Syndrome (SARS) broke out, almost all the countries that suffered local mass transmission were Asian countries. Canada was the only exception. To control the transmission of SARS, these countries increased their health expenditure greatly. These five dummy variables are defined as follows:

Afr = 1 for African countries, 0 otherwise

Nam = 1 for North American countries, 0 otherwise

Oce = 1 for the Oceania countries, 0 otherwise

Lat = 1 for South American (Latin American) countries, 0 otherwise

Eur =1 for European countries, 0 otherwise

To avoid a "dummy variable trap" problem, I do not define a dummy variable for Asian countries.

4.2 Model specification

Since there is no theoretical guideline for the form of the equation explaining health care expenditures, I will list three possible models with different forms here. With several specification tests, I will select the best one among them.

Linear Form Model:

$$\begin{aligned} PHE_i &= c_1 + c_2 PGDP_i + c_3 SOP_i + c_4 SPT_i + c_5 HALE_i + c_6 EUR_i + c_7 NAM \\ &+ c_8 OCE_i + c_9 LAT_i + c_{10} AFR_i + \varepsilon_i . \end{aligned} \quad (1)$$

For this form, the way to calculate the income elasticity of health care is:

$$\zeta = C_2 \frac{\overline{PGDP}}{\overline{PHE}},$$

where $\overline{PGDP}$ is the sample mean value of the per capita GDP and $\overline{PHE}$ is the sample mean value of per capita health expenditure.

Semi-log Form Model:

$$LNPHE_i = d_1 + d_2 PGDP_i + d_3 SOP_i + d_4 SPT_i + d_5 HALE_i + d_6 EUR_i + d_7 NAM + d_8 OCE_i + d_9 LAT_i + d_{10} AFR_i + \varepsilon_i . \quad (2)$$

For this form, the way to calculate the income elasticity of health care is:

$\zeta = d_2 \overline{PGDP}$,

where $\overline{PGDP}$ is defined above.

Log-linear Form Model:

$$LNPHE_i = f_1 + f_2 LNPGDP_i + f_3 LNSOP_i + f_4 LNSPT_i + f_5 LNHALE_i + f_6 EUR_i + f_7 NAM + f_8 OCE_i + f_9 LAT_i + f_{10} AFR_i + \varepsilon_i . \quad (3)$$

For this form, the income elasticity of health care ζ is equal to f_2 .

4.3 Methodology

My empirical analysis consists of several steps. First, I will test all three models for the presence of outliers using three tests proposed by Belsley, Kuh and Welsch (1980).[17] The first one is the studentized residuals (RSTUDENT) test. The critical value of the statistic is 2. If |RSTUDENT| for any observation is larger than 2, the related observation could be considered an outlier. The second test is DFFITS. The critical value of this statistics is $2\sqrt{K/n}$, where K is the number of regressors and n is the number of observations. Here K equals 10 and n equals 171, so the critical value equals 0.4836. If |DFFITS| is larger than 0.4836 for some observation, the related observation would be considered an outlier. The third test is based on the statistic DFBETAS, the value of which is computed for each coefficient in each model respectively. The critical value of this test equals $2/\sqrt{n}$, where n is the number of observations. Here n is 171, so the critical value is 0.1529. If |DFBETAS| is larger than 0.1529 for some observation, the related observation would be an outlier.

Second, I will exclude common outliers for all three models and then do OLS estimation of model (1), model (2) and model (3). For each model, I will do the following diagnostic tests: the Jarque-Bera test for normality of the error terms; White' s test for the heteroscedasticity of the error terms; and Ramsey' s RESET test for mis-specification. I use White' s test for heteroscedasticity since it does not rely on the normality of error terms. The results of Ramsey' s RESET test in particular may provide evidence on which model fits the data best.

Then I will carry out further tests to see which model fits the data best. Since models (2) and (3) have the same dependent variable (LNPHE), I will compare the Akaike infomation criterion values (AIC), the Schwarz criterion values (SC), the R^2 values and the log-likelihood values of both models. The model with the larger R^2 and log-likelihood values, or smaller AIC and SC values is the one that fits the data better. If model (2) fits the data better than (3), I will use Davidson and MacKinnon' s J-test to see which of models (1) and (2) is better. If model (3) is better than (2), I will use Davidson and Mackinnon' s two-way P_E[18] test to see which of models (1) and (2) is better. In order to carry out the P_E test to test the linear functional form against the log-linear one, I will first form the compound model as follows:

$$PHE_i = X_i'\gamma + \alpha_1 [\overline{LNPHE_i} - LN(X_i'c)] + \varepsilon_i, \qquad (4)$$

where $X_i = [1, PGDP_i, SOP_i, SPT_i, HALE_i, EUR_i, NAM_i, OCE_i, LAT_i, AFR_i]'$,

$\gamma = [c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8, c_9, c_{10}]'$,

[17] The tests were carried out in SHAZAM using the command OLS/influence and OLS/dfbetas.

$\overline{LNPHE_i}$ is the fitted value of *i*th observation in Model (3), and

c is the vector of OLS estimates of γ in model (1).

If the linear functional form is correct, then α_1 will be zero. Letting $DFF_{1i} = \overline{LNPHE_i} - LN(X_i^{'}c)$, I just need to add the value of DFF_{1i} to model (1), and then estimate the new equation using OLS and do a *t*-test to see whether its coefficient is significant. Similarly, in order to carry out the P_E test to test the log-linear functional form against the linear one, I need to form the following compound model:

$$LNPHE_i = Z_i^{'}\beta + \alpha_2[\overline{PHE_i} - EXP(Z_i^{'}f)] + \varepsilon_i, \tag{5}$$

where

$Z_i = [1, LNPGDP_i, LNSOP_i, LNSPT_i, LNHALE_i, EUR_i, NAM_i, OCE_i, LAT_i, AFR_i]^{'}$,

$\beta = [f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8, f_9, f_{10}]^{'}$,

$\overline{PHE_i}$ is the fitted value of *i*th observation in Model (1), and

f is the vector of OLS estimates of β in model (3).

If the log-linear functional form is correct, then α_2 will be zero. Letting $DFF_{2i} = \overline{PHE_i} - EXP(Z_i^{'}f)$, a *t*-test of significance of its coefficient will indicate whether or not the null hypothesis that log-linear model is correct can be rejected.

Once the "best" model has been identified, I will examine the significance of the coefficients and the goodness of fit of the best model. I will also re-estimate this model after deleting all outliers for it, to see how goodness of fit and the significance and magnitude of the coefficient estimates are affected.

---

[18] See Greene (2000) for the details of this test.

Then, I will calculate the income elasticity of health care expenditure with and without the outliers. If the income elasticity depends on PGDP and/or PHE, I will use the mean values of these variables in the calculation. Next, I will use a *t*-test to test if the income elasticity is greater than one, with and without outliers.

5. Estimation Results

5.1 Outliers

The results of the outlier tests for models (1), (2) and (3) are presented in Tables 4, 5 and 6 respectively. Looking at Table 4, we can see that the United States, Singapore, Switzerland and Kuwait are identified as outliers by all six tests; Qatar is identified as an outlier by five tests, Germany by four tests, Israel and Iceland by three tests and Norway and Georgia by two tests. Taking those that were identified by more than three tests as outliers, the United States, Singapore, Switzerland, Kuwait, Qatar and Germany are considered to be outliers for the linear functional form.

Turning now to Table 5, we can see that Botswana, Liberia, Luxembourg and Moldova were identified as outliers by four tests; South Africa, Lebanon, Iran and Japan were identified as outliers by three tests; Tajikistan, Ecuador, Yugoslavia, Georgia and Turkey were identified as outliers by two. If those that were selected more than three times are considered to be the outliers, then Botswana, Liberia, Luxembourg and Moldova are the outliers for the semi-log function form.

Studying Table 6, we can see that Congo, Ecuador, Equatorial Guinea, Iran, Lebanon and Liberia were identified as outliers by four tests; Bhutan, Kiribati and Cote D' lvoire were identified as outliers by three; Azerbaijan, Romania, Singapore, Zimbabwe and Jordan were identified as outliers by two tests. Once again taking those that were identified by more than three tests as outliers, then Congo, Ecuador, Equatorial Guinea, Iran, Lebanon and Liberia are the outliers for the log-linear functional form. What is interesting is that four of these six countries are quite unique. Iran and Liberia are famous for not cooperating with international organizations. Therefore, the data from these two countries are not very reliable. Congo and Lebanon suffered civil war in recent years; this fact may make the data from these two countries unreliable too.

Overall, the analysis of outliers indicates that the outlier tests are quite sensitive to the functional form and there are seemingly no common outliers for the three functional forms. However, we notice that Romania appears in the PGDP or LNPGDP items for all three models. Since the coefficients of PGDP or LNPGDP are the most important to my paper, I will delete Romania from the sample.

The OLS estimates for the three models, after Romania has been deleted, are shown in Table 7.[19] Looking at this table, we can see that the functional form affects the signs of the coefficients of some variables. For example, in the log-linear model and the linear model, the signs of the coefficients of LNHALE and HALE are both positive, while in the semi-log form, the sign of the coefficient of HALE is negative. We can also see that the functional form affects the estimate of income elasticity of health care dramatically. In the linear and log-linear models, the income elasticity is around one; however, in the semi-log model, the income elasticity is only about 0.5.

5.2 Diagnostic Tests

[19] I also estimated the three models with the data of Romania. However, the results are not very different from those obtained when Romania is excluded.

Table 8 summarizes the results of the diagnostic tests for all three models. It indicates that the normality of the residuals is rejected at the 1% significance level for the linear and log-linear models. Therefore, some tests that rely on the normality of the residuals, such as the Goldfeld-Quandt test for heteroscedasticity, are no longer valid. What is more, the *t*-tests and *F*-tests of hypotheses about the coefficients are not valid either. However, due to my relatively large sample size, we can rely on the asymptotic properties of the OLS estimates of my models. In fact, since the critical values of the *t*-tests and *F*-tests are generally more rigorous than the critical values of their asymptotic counterparts, we can still use the *t*-tests and *F*-tests to test the significance of the coefficients if no further problems are detected. Looking at the results of White' s test, we can reject the null hypothesis of homoscedasticity at the 1% significance level for the linear model. For the semi-log model, we cannot reject the null hypothesis at the 1% and 5% significance levels, but we can reject it at the 10% significance level. Finally, we cannot reject the null hypothesis for the Log-linear model at the 1%, 5% and 10% significance levels. Due to the presence of heteroscedasticity in the linear model, I should use feasible generalized least squares (FGLS) to estimate it thereafter.

Based on the results of Ramsey' s RESET tests, I can reject the null hypothesis of no specification error at 1% significance level for linear model and semi-log model. In contrast, I cannot reject the null hypothesis at the 1%, 5% and 10% significance levels for log-linear model. These results provide evidence that model (3) is the best model form among the three. The only problem is that since the error terms of the linear model are heteroscedastic, the RESET test for linear model is not reliable (the results for the semi-log and log-linear models are reliable, though). Therefore, I will do some more specification tests to select the best model.

Since the semi-log and the log-linear models have the same dependent variable (LnPHE), here I just compare the AIC values, the SC values, the R^2 values and log-likelihood values of both models (see table 7). Table 7 shows that log-linear model has larger R^2 and log-likelihood values and smaller AIC and SC values. Therefore, we can conclude that the log-linear model does a better job of explaining variations in LNPHE than the semi-log model.

Carrying out the two-way P_E test following the procedure described in section 4, I obtained the following results: the estimate of α_1 in Model (4) is 97.65293 with a *p*-value of 0.003, and the estimate of α_2 in Model (5) is -.0000306 with a *p*-value is 0.2605. Since $\hat{\alpha}_1$ is significant while $\hat{\alpha}_2$ is not, we can see that the two-way P_E test strongly rejects the linear model form in favour of the log-linear model form.

Since the log-linear model fits the data better than the semi-log model and it is accepted by two-way P_E tests against linear model, we can conclude that the Log-linear model is the "best" model among the three. What is more, there is no evidence of mis-specification from the Ramsey' s RESET tests which had been done. Therefore, I will focus on the analysis of log-linear model in the remaining part of my paper.

In the log-linear form, the income elasticity of health care equals the coefficient of the variable LNPGDP. In order to see if the outliers affect the estimate significantly, I report the OLS estimates for all the variables in the log-linear model with and without outliers in Table 9.

As Table 8 shows, deleting outliers or not does not affect the results of OLS estimation in any obvious way. Meanwhile, we can see that the R^2 values reported with or without outliers are very high (both are larger than 0.95), so we can conclude that the log-linear model fits the data set very well. The income elasticity of health care expenditure (the coefficient of LNPGDP) is statistically significant, with or without outliers. The positive sign of this coefficient implies that health care is a normal good; in other words, as income increases, the demand for health care increases as well. More specifically, a one percent increase in per capita GDP will increase per capita health care expenditure by about 1.1 percent, holding all else constant. In order to see if this value is significantly larger than one, I carried out a one-sided *t*-test of the null hypothesis f_2=1 against the alternative that f_2>1. Without outliers, the value of the *t*-statistic is 11.1178 and its *p*-value is 0.001, implying that we can reject the hypothesis that f_2=1 in favour of f_2>1. Similarly, with outliers, the value of the *t*-statistic is 33.7849 and its *p*-value is 0.0000, so we can reject the hypothesis that f_2=1 in favour of f_2>1 once again. The income elasticity is significantly larger than one. Health care appears to be a luxury instead of a necessity.

With or without outliers, the coefficient of LNHALE, the log value of health-adjusted life expectancy, is significant at the 5% level. Its negative sign implies that health status has a negative impact on health care expenditure. This result is consistent with theoretical analyses of the demand for health care: because the demand for health care is a derived demand, healthy people may have less demand for health care.

The coefficient of LNSOP, the log of the percentage of aged people in the total population, is also significant at the 1% level, with or without outliers. Its positive sign is consistent with the belief that older people may demand more health care. In addition, the negative sign of the coefficient of LNSPT, the log of the public share of health care expenditure in total health care expenditure, is consistent with the result reported in Gerdtham and Joensson (1992). This coefficient is also statistically significant at the1% level, whether or not the outliers included.

The coefficients of all the dummy variables except AFR, which accounts for the African countries, are statistically significant at least the 5% level, both with and without outliers. Therefore, we can conclude that if the values of all the other variables are the same, a country in Africa and a country in Asia will have the same health care expenditure level. However, holding other variables the same, a country in Europe tends to spend more on health care than Asian countries, as do the countries in Latin America, North America and Oceania. What is more, holding the other variables the same, with outliers, countries in Oceania spend the most on health care; when outliers are deleted, Latin American countries spend the most on health care.

After the above steps, I delete the dummy variable AFR to get the following simplified model (6):

$$\begin{aligned} LNPHE_i =& g_1 + g_2 LNPGDP_i + g_3 LNSOP_i + g_4 LNSPT_i + g_5 LNHALE_i \\ &+ g_6 EUR_i + g_7 NAM + g_8 OCE_i + g_9 LAT_i + \varepsilon_i \end{aligned} \qquad (6)$$

In this model, I group African and Asian countries as a reference group because due to the historical and environmental reasons (long history of colonization, long periods civil war and severe environmental problems), most countries within these two continents have a similarly low level of development, except Japan and a few emerging market countries in Asia.

The OLS estimates of the coefficients of equation (6), with or without outliers, are listed in Table 9. The high values of R^2 and the adjusted R^2 indicate that the model fits the data very well. Furthermore, deleting outliers or not does not affect the sign or the magnitude of the estimates much. The signs of all the estimates seem to be consistent with the predictions of health economics. The income elasticity of health care is seemingly larger than one, implying that health care is a luxury.

However, readers may raise the question of whether countries with different wealth levels have same income elasticity. In other words, is the income elasticity stable among different countries with different wealth levels? This concern is important since my sample includes so many countries with different income levels. In order to test the stability of the income elasticity, I divide my sample countries into four groups following World Bank practice. The World Bank classifies countries all over the world into four groups based on the level of per capita GDP. Group 1 are the poorest countries- their per capita GDPs are less than 735 US dollars. Group 2 are the relatively poor countries; their per capita GDPs range from 735 to 2936 US dollars. Group 3 are the relatively rich countries- their per capita GDPs range from 2936 to 9076 US dollars. Finally, Group 4 consists of the richest countries with per capita GDPs in excess of 9076 US dollars. In order to account for the wealth level of each country in my sample set, I define three additional dummy variables as follows:

W1=1 for these relatively poor countries (in Group 2), 0 otherwise;

W2=1 for these relatively rich countries in (Group 3), 0 otherwise;

W3=1 for these richest countries in (Group 4), 0 otherwise.

To avoid a "dummy variable trap" , I do not define a dummy variable for the poorest countries. The alternative to model (6) is:

$$LNPHE_i = h_1 + h_2 LNPGDP_i + h_3 W_1 LNPGDP_i + h_4 W_2 LNPGDP_i + h_5 W_3 LNPGDP_i$$
$$+ h_6 LNSOP_i + h_7 LNSPT_i + h_8 LNHALE_i + h_9 EUR_i + h_{10} NAM + h_{11} OCE_i \quad (7)$$
$$+ h_{12} LAT_i + \varepsilon_i .$$

In model (7), I assume that the constant terms and the parameters of all variables except LNGDP are the same across groups with different income levels. Of course, it would be more general to relax this assumption to let all parameters differ accross different groups. However, since the number of rich countries and relatively poor countries are small (about 30), the regression results for these two groups will not be as statistically reliable.

The results of a Wald test of the null hypothesis that $h_3=h_4=h_5=0$ are presented in Table 10. We cannot reject the null hypothesis that $h_3=h_4=h_5=0$. Therefore, we can conclude that the income elasticity is quite stable among countries with different income level. [20]

6. Conclusions

The issue of whether health care is a luxury or a necessity has been discussed by a great number of researchers over the past three decades. By analyzing different data sets with different methods, those researchers obtained different results.

With more recent data from a much wider selection of countries, I examined this issue again. Three main conclusions can be drawn from my research. First, the relationship between per capita health expenditure and its determinants can be expressed in a log-linear model. Second, the income elasticity of health care is larger

[20] I aslo estimated a model in which the intercepts were allowed to differ across countries and carried out the Wald test to see if income elasticity is stable across countries with different income level. We can not the null hypothesis that income elasticity is stable.

than one, which means that health care is a luxury instead of a necessity. This result is consistent with the conclusions of Newhouse (1977), Leu (1986), Gerdtham and Joensson (1991), Gergtham and Joensson (1992) and Hitiris (1992) and explains the phenomena mentioned in section 1 of this paper. Finally, income is not the only determinant of health care expenditure. The results of my empirical analyses confirm the conclusions of Leu (1986) that demographic variables and other variables reflecting the public share of health expenditure also affect the demand for health care, and thus affect health care expenditure. What is more, unlike former research, this paper is the first to include the variable HALE, which reflects the population' s health status. According to my empirical research, this variable affects health care expenditure significantly: a one percent increase in HALE will reduce health care expenditure by about 0.5 percent, if other factors remain unchanged.

At the same time, there are some limitations with my study. Like all the previous studies, my study does not have enough theoretical guidance on selecting explanatory variables for aggregate health expenditure. As I have mentioned in section 3, due to the special characteristics of health care as a commodity, the factors that affect the demand for health care are very complicated. Because of the inavailability of data for other variables, for example, the relative price index for health services or the number of physicians in most developing countries, I cannot include more variables in my model and test their significance, unless I limit my sample to developed countries. Even though my model has very high goodness of fit and the results of the specification tests show no evidence of mis-specification, I cannot conclude that no variables have been omitted. Omitting variables might make my estimate of the income elasticity biased, which will cast doubt on my conclusion that health care is a luxury. As the WHO and other global organizations develop more complete data sets for health care, it would be desirable to include more variables in further research to determine more precisely the determinants of the demand for health care.

What is more, like other cross-country analyses, my research is based on the assumption that countries are homogeneous with respect to health care expenditure. In particular, the coefficients of the estimated equations are assumed to be the same for all countries. This assumption seems unrealistic since different countries may have different levels of development, environmental conditions, production functions, health care systems and so on. I tried to solve the problem of differences across countries by adding geographic dummy variables to my model, but this method is obviously tentative. It would be desirable to add more dummy variables, especially those that can account for differences in health care systems and health care policies across countries, if data become available.

Finally, including the variable HALE may violate the assumption of OLS that the explanatory variables are uncorrelated with the error term of the estimating equation. Theoretically, the HALE of a country might be affected by the country' s health care expenditure. Therefore, the HALE might be a function of health care expenditure while also being an explanatory variable for health care expenditure. If HALE is a function of health care expenditure, it will inevitably be correlated with the error term. It would be desirable to find some other exogenous variables as instrumental variables and carry out the Durbin-Wu-Hausman test to see if there is a correlation between the HALE and the error term. If there is a correlation between the

HALE and error term, the estimate of income elasticity will be biased. This would make the conclusion that health care is a luxury doubtful.

References

Belsley, D.A., E. Kuh, and R.E. Welsch (1980), *Regression Diagnostic*, New York: Wiley.

Blomqvist, A., and R.A.L.Carter (1997), "Is Health Care Really a Luxury?" , *Journal of Health Economics*, Vol. 16, 207-229.

Culyer, A.J. (1971), "Is Medical Care Different?" , in Cooper, M. H. and A.J. Culyer (eds.), *Health Economics,* Harmondsworth: Penguin, 49-74.

Gerdtham, Ulf-G., and Bengt Joensson (1991), "Conversion Factor Instability in International Comparisons of Health Care Expenditure", *Journal of Health Economics*, Vol. 10, 227-234.

Gerdtham, Ulf-G., and Bengt Joensson (1992), "International Comparisons of Health Care Expenditure-Conversion Factor Instability, Heteroscedasticity, outliers and Robust Estimators" , *Journal of Health Economics*, Vol. 11, 189-197.

Greene, W.H. (2000), *Econometric Analysis*, Fourth Edition, New Jersey: Prentice Hall.

Grossman, M. (1972), *The Demand for Health,* New York: National Bureau of Economic Research.

Hansen, P., and King, A. (1996), "The Determinants of Health Care Expenditure: a Cointegration Approach" , *Journal of Health Economics*, Vol. 15, 127–137.

Hitiris, Theo, and John Posnett (1992), "The Determinants and Effects of Health Expenditure in Developed Countries" , *Journal of Health Economics*, Vol.11, 173-181.

Hitiris, T. (1997), "Health Care Expenditure and Integration in the Countries of the European Union" , *Applied Economics,* Vol. 29, 1-6.

Human Development Reports Office (2002), *Human Development Reports 2002*, http://hdr.undp.org/reports/global/2002.

Im, K.S., M.H. Pesaran, and Y. Shin (1997), "Testing for Roots in Heterogeneous Panels" , DAE Working Paper No. 9526, University of Cambridge.

Johansen, S. (1991), "Determination of Cointegration Rank in the Presence of a Linear Trend" , *Oxford Bulletin of Economics and Statistics* 54, 383-397.

Leu, R. (1986), "The Public-Private Mix and International Health Care Costs" , in Culyer A. and Joensson B. (eds.), *The Public-Private Mix of Health Services*, Oxford: Blackwell, 41-61.

McCoskey, S.K., and Thomas M. Selden (1998), "Health Care Expenditure and GDP: Panel Data Unit Root Test Results" , *Journal of Health Economics*, Vol. 17, 369-376.

McPake, B., L. Kumaranayake, and C. Normand (2002), *Health Economics-An International Perspective*, New York: Routledge.

Musgrove, P. (1983), "Family Health Spending in Latin America" , *Journal of Health Economics,* Vol. 2, 245-257.

Murthy, Vasudeva N.R. (1992), "Conversion Factor Instability in International Comparisons of Health Care Expenditure: Some Econometric Comments" , *Journal of Health Economics*, Vol. 11, 183-187.

Newhouse, J.P. (1977), "Medical Care Expenditure: a Cross-National Survey" , *Journal of Human Resources*, Vol.12, 115–125.

Organisation for Economic Co-operation and Development (1995), "Factors Affecting Health Spending: a Cross-Country Econometric Analysis" , in *Health Policy Studies*, No. 7, 71-85.

Parkin, D., Alistair McGuire, and Brian Yule (1987) "Aggregate Health Care Expenditures and National Income: Is Health Care a Luxury Good?" , *Journal of Health Economics*, Vol. 6, 109-127.

Pfaff, Martin (1990), "Differences in Health Care Spending Across Countries: Statistical Evidence" , *Journal of Health Politics, Policy and Law,* Vol. 15, 1-67.

Phillips, P.C.B, and P. Perron (1988), "Testing for a Unit Root in Time Series Regression" , *Biometrika,* Vol. 75, 335-346.

Phillips, P.C.B., and M. Loretan (1991), "Estimating Long-run Equilibria" , *Review of Economic Studies*, Vol. 58, 407-436.

Roberts, J. (2000), "Spurious Regression Problems in the Determinants of Health Care Expenditure: a Comment on Hitiris (1997)" , *Applied Economics Letters*, Vol. 7, 279-283.

Ronald M., Andersen (1995), "Revisiting the Behavioral Model and Access to Medical Care: Does It Matter?" , *Journal of Health and Social Behavior*, Vol. 36, 1-10.

Shin, Y. (1994), "A Residual-based Test of the Null of Cointegration against the Alternative of No Cointegration" , *Econometric Theory,* Vol. 10, 91-115.

Sorkin A. L. (1975), *Health Economics*, Lexington: Lexington Books.

Varian, H.R. (1993), *Intermediate Microeconomics-A Modern Approach,* Third Edition, New York: W. W. Norton & Company.

World Health Organization (2002), *World Health Report 2002*, http://www.who.int/whr/en/.

Zweifel, P., and F. Breyer (1997), *Health Economics*, New York: Oxford University Press.

Appendix

List of Sample Countries:
European countries (33 countries):

Albania	Austria	Belgium	Croatia	Czech Republic	Denmark
Estonia	Finland	France	Germany	Greece	Hungary
Iceland	Israel	Italy	Latvia	Lithuania	Luxembourg
Malta	Netherlands	Norway	Poland	Portugal	Romania
Russian Federation	Slovak Republic	Slovenia	Spain	Sweden	Switzerland
Turkey	United Kingdom	Yugoslavia, Fed. Rep.			

Africa countries (42 countries):

Algeria	Cape Verde	Egypt, Arab Rep.	Guinea	Madagascar	Mozambique
Angola	Central African Republic	Equatorial Guinea	Guinea-Bissau	Malawi	Namibia
Benin	Comoros	Ethiopia	Kenya	Mali	Senegal
Botswana	Congo, Dem. Rep.	Gabon	Lesotho	Mauritania	Seychelles
Burkina Faso	Congo, Rep.	Gambia, The	Liberia	Mauritius	Sierra Leone
Cameroon	Djibouti	Ghana	Libya	Morocco	South Africa
Algeria	Cape Verde	Egypt, Arab Rep.	Guinea	Madagascar	Mozambique

Asian countries (65 countries):

Antigua and Barbuda	Armenia	Azerbaijan	Bahamas, The	Bahrain	Brazil
Bangladesh	Barbados	Belarus	Bhutan	Bosnia and Herzegovina	Cote d'Ivoire
Bulgaria	Burundi	Cambodia	Chad	China	Haiti
Dominica	Dominican Republic	Eritrea	Georgia	Grenada	Japan
India	Indonesia	Iran, Islamic Rep.	Ireland	Jamaica	Lebanon
Jordan	Kazakhstan	Kuwait	Kyrgyz Republic	Lao PDR	Nigeria
Malaysia	Maldives	Moldova	Mongolia	Nepal	Rwanda
Oman	Pakistan	Panama	Philippines	Qatar	Singapore
St. Kitts and Nevis	St. Lucia	St. Vincent and the Grenadines	Sao Tome and Principe	Saudi Arabia	Trinidad and Tobago
Sri Lanka	Syrian Arab Republic	Tajikistan	Tanzania	Thailand	Yemen, Rep.
Turkmenistan	Uganda	Ukraine	Uzbekistan	Vietnam	

South American countries (11 countries):

Argentina	Bolivia	Chile	Colombia	Ecuador	Guyana
Paraguay	Peru	Suriname	Uruguay	Venezuela, RB	

North American countries (8 countries):

Belize	Canada	Costa Rica	El Salvador	Guatemala	Honduras
Mexico	United States				

Oceania countries (12 countries):

Australia	Fiji	Kiribati	Marshall Islands	Micronesia, Fed. Sts.	New Zealand
Palau	Papua New Guinea	Samoa	Solomon Islands	Tonga	Vanuatu

Data Resources

Per capita health care expenditure (PHE), converted by GDP PPPs, were collected from *The World Health Report 2002,* Annex 5.
http://www.who.int/whr/2002/annex/en/#excel
Per capita GDP (PGDP), converted by GDP PPPs, were collected from *Human Development Reports 2002*[21] commissioned by the United Nations Development Programme (UNDP).
The share of health care expenditure to total health care expenditure (SPT) was collected from *The World Health Report 2002,* Annex 5.
The percentage of aged people (60 and order) the total population (SOP) was collected from *The World Health Report 2002,* Annex 1.

[21] The Human Development Report (HDR) is an independent report and commissioned by the United Nations Development Programme (UNDP). From 1990 to 2003, an annual report was published each year. In each issue, more than 200 indexes were reported and PGDP was one of them.

Table 1: Total expenditure on health, %GDP

Country	1960	1970	1980	1990	2000
Canada	5.4	7	7.1	9	9.2
Finland	3.8	5.6	6.4	7.8	6.7
Ireland	3.7	5.1	8.4	6.1	6.4
Japan	3	4.5	6.4	5.9	7.6
Norway	2.9	4.4	6.9	7.7	7.6
Spain	1.5	3.6	5.4	6.7	7.5
United Kingdom	3.9	4.5	5.6	6	7.3
United States	5	6.9	8.7	11.9	13.1

Source: OECD Health Data 2003

Table 2. Summary of former research

Author	Data set		Functional Form	Conclusion (Health care is:)	
	No. of countries	Year			
Newhouse (1977)	13	1972	Linear	Luxury	
Leu, R. (1986)	19	1974	Log-linear	Luxury	
				RE[1]	PPP[2]
Parkin, D. et al (1987)	18	1980	Linear	Luxury	Necessity
			Semi-log	Necessity	Necessity
			Exponential	Luxury	Luxury
			Log-linear	Luxury	Luxury
Gerdtham (1991)	22	1985	Log-linear	Luxury	
			Linear	Luxury	
Murthy (1992)	22	1985	Linear Log-linear	Ambiguous	
Gergtham et al (1992)	20	1985	Log-linear	Luxury	
Hitiris (1992)	20	1960-1987	Log-linear	Luxury	
OECD (1995)	20	1970-1991	Log-linear	Necessity	
Blomqvist and Carter (1997)	18	1960-1991	Log-linear	Necessity	
Hitiris (1997)	10	1960-1991	Log-linear	Luxury	

Note: 1. Health care expenditures and GPD are converted from national currencies to US dollars using exchange rates.

2. Health care expenditures and GPD are converted from national currencies to US dollars using GDP PPPs.

Table 3: Descriptive statistics (grouped by regions)

Var	STAT	ASIA	AFRICA	OCE	NAM	LAT	EUR	TOTAL
PHE	Mean	323	135	516	1126	464	1460	557
	Median	199	53	266	434	323	1469	237
	std.Dev.	395	173	675	1570	344	921	776
PGDP	Mean	6012	3407	7596	12004	5953	18316	8135
	Median	3816	1881	3735	7128	4799	17367	4668
	Std.Dev.	6416	3504	8440	12050	3084	10316	8831
HALE	Mean	54.79	42.78	57.82	61.41	58.95	66.72	54.93
	Median	57.32	41.89	56.46	61.29	58.59	69.24	57.41
	Std.Dev.	8.68	9.18	6.65	5.85	65.83	4.64	11.19
SPT	Mean	55.78	58.25	73.52	53.83	56.04	73.74	61.00
	Median	58.80	59.65	73.70	47.15	55.80	74.30	63.40
	Std.Dev.	20.53	16.22	9.97	11.93	12.69	16.61	18.23
SOP	Mean	8.42	5.51	7.69	8.91	8.69	18.74	9.69
	Median	7.10	5.15	6.90	7.15	7.00	19.90	6.90
	Std.Dev.	4.61	1.27	4.08	4.80	3.60	4.19	5.97

Note: Each value of PHE and PGDP has been rounded to ones place.

Table 4. The outliers for the model (1): linear function form

RSTUDENT	DFFITS	DFBETA-PGDP	DFBETA-HALE	DFBETA-SOP	DFBETA-SPT
United States	United States	United States	United States	United States	United States
Qatar	Qatar	Luxembourg	Qatar	Singapore	Luxembourg
Luxembourg	Luxembourg	Qatar	Norway	Switzerland	Switzerland
Singapore	Singapore	Singapore	Iceland	Qatar	Kuwait
Switzerland	Switzerland	Switzerland	Kuwait	Kuwait	Singapore
Israel	Israel	Israel	Germany	Lebanon	
Germany	Kuwait	Kuwait	Georgia	Luxembourg	
Kuwait	Mexico	Iceland	Singapore	Georgia	
	Iceland	Romania	Switzerland		
	Belize	Latvia			
	Germany	Germany			
		Norway			

Table 5. The Outliers for Model (2): Semi-Log Functional Form

RSTUDENT	DFFITS	DFBETA-PGDP	DFBETA-HALE	DFBETA-SOP	DFBETA-SPT
Botswana	Luxembourg	Luxembourg	Botswana	Norway	Lebanon
South Africa	Ecuador	Japan	Luxembourg	Turkey	Tajikistan
Liberia	Botswana	Botswana	Angola	Japan	Moldova
Lebanon	Japan	Bahamas, The	Zimbabwe	Moldova	Bhutan
Tajikistan	Moldova	Turkey	Burundi	Georgia	Georgia
Luxembourg	Yugoslavia	Switzerland	Cape Verde	Yugoslavia	Liberia
Namibia	Liberia	Romania	Swaziland	Iran	South Africa
Moldova	South Africa	Norway	Comoros		Sudan
Ecuador			Liberia		Indonesia
Iran,			Lebanon		Vietnam
			Iran		Maldives
					Gambia
					Dominican
					Saudi Arabia

Table 6. The outliers for the model (3): Log-linear Functional Form

RSTUDENT	DFFITS	DFBETA -LNPGDP	DFBETA -LNHALE	DFBETA -LNSOP	DFBETA -LNSPT
Azerbaijan	Congo, D. R.	Angola	Angola	Azerbaijan	Bhutan
Bhutan	Ecuador	Bhutan	Cape Verde	Iran.	Cambodia
Congo, D. R.	Equatorial Guinea	Congo, D. R.	Cote d'Ivoire	Jordan	Congo,D. R.
Ecuador	Iran	Cote d'Ivoire	Ecuador	Norway	Cote d'Ivoire
Equatorial Guinea	Kiribati	Ecuador	Equatorial Guinea	Qatar	Indonesia
Iran	Lebanon	Equatorial Guinea	Iran		Kenya
Lebanon	Liberia	Kiribati	Jordan		Kiribati
Liberia	Romania	Lebanon	Malawi		Lebanon
Romania		Liberia	Tunisia		Liberia
		Mongolia	Zimbabwe		Nigeria
		Qatar			Singapore
		Romania			Sudan
		Singapore			Turkmenistan
		South Africa			
		Zimbabwe			

Table 7.OLS regression result for three models: Sample size = 170

Variable or Statistic	Coefficient (*p*-value)	
	Linear Model	Semi-log Model
PGDP	0.07904 (0.00)	9.39E-05 (0.00)
HALE	-7.490911 (0.0174)	0.084134 (0.00)
SOP	19.41565 (0.0088)	0.004525 (0.806)
SPT	-2.681116 (0.0243)	-0.002125 (0.569)
AFR	39.89253 (0.183)	0.431163 (0.022)
EUR	95.34502 (0.247)	0.390059 (0.112)
LAT	122.9746 (0.0006)	0.661755 (0.012)
NAM	346.1475 (0.0762)	0.593089 (0.051)
OCE	171.0335 (0.0003)	0.533936 (0.041)
C	289.4548 (0.0414)	-0.804513 (0.133)
Income elasticity ζ	1.082289	0.537210
R^2	0.826349	0.903621
Log-likelihood	-194.374	-80.738
Akaike infomation criterion	2.416166	0.987234
Schwarz criterion	2.619071	1.483677
	Log-linear Model	
LNPGDP	1.104741 (0.00)	
LNHALE	-0.541682 (0.0428)	
LNSOP	0.251325 (0.0046)	
LNSPT	-0.22437 (0.0079)	
AFR	-0.037029 (0.6731)	
EUR	0.244841 (0.0236)	
LAT	0.311823 (0.01)	
NAM	0.353313 (0.0122)	
OCE	0.408321 (0.0007)	
C	-1.31114 (0.1643)	
Income elasticity ζ	1.104741	
R^2	0.962479	
Log-likelihood	-64.1412	
Akaike infomation criterion	0.872249	
Schwarz criterion	1.056708	

Table 8. Result of the diagnostic tests

Functional form		Linear	Semi-log	Log-linear
Jarque-Bera test:				
Statistics		1449.5	2.2263	20.989
(p-value)		(0.0000)	(0.3285)	(0.000028)
White's test:				
Statistics		76.29578	21.58315	11.49820
(p-value)		(0.0000)	(0.087595)	(0.569157)
RESET Test:				
Statistics (p-value)	γ=2	21.38236 (0.000000)	29.49243 (0.000000)	1.716644 (0.183002)
	γ=3	18.36778 (0.000000)	24.77553 (0.000000)	1.394710 (0.246489)
	γ=4	27.20617 (0.000000)	18.58659 (0.000000)	1.039637 (0.388665)

Note: γ is the number of fitted terms.

Table 9: Estimated Coefficients with and without outliers

Variable	Coefficient (t-Statistic)			
	Model (3)		Model (6)	
	With outliers	Without outliers	With outliers	Without outliers
LNPGDP	1.104741**	1.108558**	1.102107**	1.103792**
	(35.16856)	(40.62248)	(35.88759)	(42.713)
LNHALE	-0.541682*	-0.548359*	-0.48856*	-0.604**
	(-2.04226)	(-2.380805)	(-2.09713)	(-3.08977)
LNSOP	0.251325**	0.258932**	0.252122**	0.268377**
	(2.872927)	(3.418371)	(2.89009)	(3.69697)
LNSPT	-0.22437**	-0.220479**	-0.22751**	-0.17242*
	(-2.69165)	(-3.045212)	(-2.74729)	(-2.45687)
AFR	-0.037029	0.002954		
	(-0.42265)	(0.038948)		
EUR	0.244841*	0.251629**	0.249059*	0.259599**
	(2.285906)	(2.734662)	(2.341421)	(2.939983)
LAT	0.311823*	0.432944**	0.318824**	0.448336**
	(2.606412)	(4.051208)	(2.697746)	(4.389938)
NAM	0.353313*	0.369847**	0.359811**	0.392804**
	(2.536159)	(3.093729)	(2.605328)	(3.426052)
OCE	0.408321**	0.427335**	0.417097**	0.429451**
	(3.442548)	(4.199866)	(3.580842)	(4.445366)
C	-1.31114	-1.363686	-1.50315	-1.32776
	(-1.39729)	(-1.687092)	-1.83549	(-1.94404)
R^2	0.962479	0.971114	0.962437	0.973982
Adjusted R^2	0.960369	0.969436	0.960571	0.972639
F-statistic	456.0328	578.9824	515.6453	720.3059
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)

Note: Coefficients with superscript "**" are statistically significant at 1% significance level;
Coefficients with superscript "*" are statistically significant at 5% significance level;
All constant terms (Cs) are statistically significant at the 10% significance level.

Table 10. Wald test for the stability of income elasticity, with or without outliers

Variable	Coefficient/ Statistics (*p*-value)	
	With outliers	Without outliers
h_2	1.056479 (0.000)	1.066479 (0.000)
h_3	0.014084 (0.1955)	0.014084 (0.3941)
h_4	0.018718 (0.1327)	0.018718 (0.4179)
h5	0.038926 (0.082)	0.038926 (0.1934)
Wald Test	3.285440 (0.349674)	3.130121 (0.371995)