

1993

# Is Health Care Really a Luxury?

A G. Blomqvist

RAL Carter

Follow this and additional works at: <https://ir.lib.uwo.ca/economicsresrpt>

 Part of the [Economics Commons](#)

---

## Citation of this paper:

Blomqvist, A G., RAL Carter. "Is Health Care Really a Luxury?." Department of Economics Research Reports, 9311. London, ON: Department of Economics, University of Western Ontario (1993).

35555

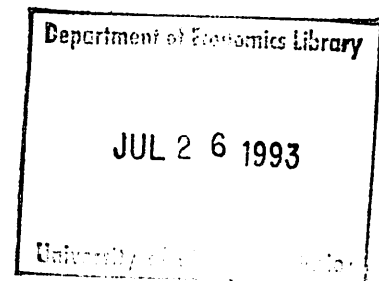
ISSN: 0318-725X  
ISBN: 0-7714-1537-0

RESEARCH REPORT 9311

Is Health Care Really a Luxury?

by

A.G. Blomqvist  
and  
R.A.L. Carter



July 1993

Department of Economics

Social Science Centre

University of Western Ontario

London, Ontario, Canada

N6A 5C2

# Is Health Care Really a Luxury?

Å. G. Blomqvist

R. A. L. Carter \*

Department of Economics, University of Western Ontario,  
London, Ont. N6A 5C2, Canada

Using OECD data we investigate the widely held notion that the income elasticity of expenditure on health care exceeds one. When we employ traditional models and econometric techniques we obtain traditional results. We then expand the model to account for technical change, long run effects and short run dynamics. Modern techniques of inference that take explicit account of the time series nature of the data leave us unable to reject the null hypothesis that the income elasticity is less than one, and lead us to conclude that technological change has been the main factor explaining health care's growing share of GDP.

## 1 Introduction

“As with luxury goods, health spending tends to rise disproportionately as countries become richer...” *The Economist* March 27, 1993, p113.

---

\*We are grateful for diligent research assistance from P. Gogas and, especially, C. Wang. We have received helpful comments from: J. Kimmel, A. Melino, J. Praschnik, P. Schmidt and M. Veall. D. Fretz has given us advice and assistance above and beyond the calls of both friendship and duty. We have benefited from an SSHRC Internal Grant administered by the University of Western Ontario. We are entirely responsible for any errors. Correspondence to: Å. Blomqvist, Department of Economics, University of Western Ontario, London, Ont. N6A 5C2, Canada; tel. (519) 679 2111 ext. 5305; fax (519) 661 3666; e-mail blomqvist@uwo.ca

One of the puzzling "stylized facts" emerging from aggregative data on health care expenditures is that health care apparently is a luxury – that is, the demand for health care seems to have an income elasticity in excess of unity. The hypothesis that this is so has been supported in numerous studies using international cross-section data, and is also consistent with time series evidence from individual industrialized countries: In most cases, health care spending has been growing considerably faster than real income.

To those who think of health care as something that is (or ought to be) provided in accordance with some more or less well-defined criterion of "need", the concept of health care as a luxury certainly appears counter-intuitive. If the per-unit cost of care and the incidence of illness are similar across populations, a poor population would be expected to devote a *higher* share of its income to health care than a rich one. This is also the pattern that emerges when data on health care spending by families and individuals within a country are related to family or individual income. In such studies, the estimated income elasticity of demand is typically well below unity.

The discrepancy between income elasticities estimated from individual or family data and those estimated from aggregative cross-section or time series data is not easily explained. The relatively low values of the former are consistent with the idea that the public-sector insurance programs which exist in every industrialized country ensure that those who fall ill receive relatively similar standards of treatment regardless of their income. Even in the United States where ability to pay (for health insurance or health care) probably has more significance as a determinant of health care spending than in any other industrialized country, the existence of publicly financed programs such as Medicare and Medicaid certainly can be expected to weaken the link between family income and the standard of care. To the extent that this type of redistribution explains the relatively low apparent income elasticities estimated on the basis of individual or family data one would expect them to be somewhat lower than those estimated from aggregative international data. However, even allowing for this effect, the idea that health care spending should behave as a luxury good, when aggregate data are used, appears puzzling.

One possible explanation has been based on the suggestion that the per-unit cost of health care tends to rise, relative to other goods and services, as income rises; an intuitively appealing justification for this suggestion is based on the idea that health care is a relatively labour-intensive commodity,

and that the relative price of labour-intensive commodities tends to rise with income<sup>1</sup>. This idea received some support from the results reported in Parkin et al. (1987). In their study, they found that in a cross-section of data from 18 countries for which explicit estimates by the OECD of a purchasing-power-parity (PPP) index for health care was used to deflate aggregate health care expenditures, the elasticity of aggregate health care spending with respect to real GDP (deflated by a GDP PPP index) was 0.90. However, subsequent work by Gerdtham and Jönsson (1991) casts doubt on the robustness of this result: They obtained, using a 1985 cross section of OECD data, an estimated income elasticity of 1.43, whether per capita health care spending was deflated by a health care PPP index or by a general PPP index (p. 231). In work published in Swedish, they found no statistically significant correlation between the relative price of health care and real GDP per capita (Gerdtham and Jönsson (1990), p. 654). Moreover, in Gerdtham, Sjøgaard, Jönsson, and Anderson (1992) (henceforth GSJA), the authors find no significant effect of the relative price variable in their cross-section regressions.

Like most earlier contributions to this literature, the estimates by Parkin et al. and Gerdtham and Jönsson were essentially based on two-variable regressions of some measure of health care spending on GDP. The main reason for this, of course, is the fact that when the data are limited to a few international cross-sections or short time series from individual countries, the scope for testing richer specifications of the determinants of aggregate health care spending is limited. However, thanks to the work of the OECD (1990), a relatively rich data base consisting of reasonably comparable time series for 24 countries is now opening up the possibility of testing more elaborate specifications by pooling cross-sections of time series. In particular, since there are very substantial differences among countries in the financing and organization of health services production, evidence on how the cost of health care varies among systems that differ in this respect may be of considerable interest.

The most ambitious existing attempt along these lines is by GSJA who specify a model in which per capita health care spending depends on per capita real income and some ten different institutional and demographic variables, such as: the proportion of inpatient expenditure to total health

---

<sup>1</sup>Note that this explanation is based on the implicit assumption that the demand for health care at the aggregate level is price-inelastic.

care spending, the percent of total health care costs paid for by government; urbanization, female labour force participation, the proportion of elderly people in the population; and a dummy variable representing predominance or otherwise of the fee-for-service method of paying physicians. They also include an index of the relative price of health care services as an explanatory variable.

To estimate the parameters of this model, GSJA use three cross-sections of 19 countries each, observed some six or seven years apart. They also allow for the possibility of shifts over time by including dummy variables to separate the three cross-sections. Although continuous time series of 15 years or more are available for some of the variables in OECD (1990), data for others had to be obtained from other sources which, in some cases, don't provide complete time series; this, presumably, is one of the reasons the estimation is based on the three cross-sections only.

The GSJA equations generally show a high degree of explanatory power (with  $R^2$ s above .9), and provide a number of interesting insights into the question of what factors modify the income-health spending relationship. Most interesting from our point of view are their point estimates of the income elasticity. Although generally lower than the estimate in Gerdtham and Jönsson (1990) using 1985 data, they cluster around 1.3 and remain significantly above unity in most versions of their basic equation. Interestingly, they also find statistically significant upward shifts in health spending over time: For example, the estimated shift dummy for 1987 in their "preferred" specification suggests that on average, per capita health care spending has grown at a little above 1.5% per year in real terms, after correcting for the influence of per capita income and other variables.

The work we undertake in the present paper can be seen as complementary to that of GSJA. Our approach is to use essentially the entire time series of annual cross-section data assembled by OECD to derive alternative estimates of the income elasticity of health care spending: our sample consists of 28 annual observations for 18 countries. The nature of this data set forces us to use estimation techniques specifically developed for time series analysis. Although this makes the analysis more complicated, it also has the advantage that we can rigorously investigate the GSJA result that the health care-income relationship has been shifting over time, by testing for the presence of an independent time trend in the health expenditure series.

Like much of the literature on aggregate productivity, we will interpret

the existence of an independent time trend as being principally due to the influence of changing technology. At first glance, it might seem that the effect of technical progress ought to be to reduce health care spending: Technical progress can, in general, be seen as something that lowers the effective cost of producing an industry's output, and if the price elasticity of demand is less than one (as it is generally thought to be in the health services industry), total spending should decrease. However, the demand for health services can be thought of as a derived demand, the fundamental commodity that is valued by consumers being "good health". Therefore, if technical change takes the form of progress in our ability to transform health services into "good health", rather than reducing the resource cost of producing health services, then the demand curve for health services would shift out, so that total spending could increase even if the price elasticity were less than one. Many observers have argued that in health care, most technological change is of the former, cost-increasing, kind. Baumgardner (1991) provides a theoretical model which explains how the prevalence of third-party financing of health care spending can be used to explain why this is so.

A drawback of using the full OECD time series data set is, of course, that we have to omit many of the institutional and demographic variables considered by GSJA, because of the difficulty of constructing comparable time series for them. To partially compensate for this, we have, in some versions of our basic equation, used country-specific dummy variables. Clearly, this is a very crude way to account for the effects of the relevant differences among national health care systems; for one thing, their characteristics may change over time. Nevertheless, institutional characteristics have some degree of stability<sup>2</sup>, and the estimated coefficients on the country dummies may be of interest in themselves.

## 2 The Data

As in most of the earlier literature, the basic model we postulate is a simple relationship between health care expenditure in real terms, and real per capita income. Our principal data source is OECD (1990) which gives annual time series, spanning the years 1960 to 1987, for the 24 OECD countries

---

<sup>2</sup>In fact, in the GSJA paper, four of the ten explanatory variables (other than the time dummies) have time-independent values for each country in their sample.

on: total populations, health care expenditure and gross domestic product in national currencies, and the PPP index for GDP, with the U.S. as base. Unfortunately, the data series on Luxembourg, New Zealand, Portugal and Turkey were much shorter than for the other countries so they were excluded from the analysis. Real, U.S. dollar, per capita versions of the health care expenditure and income variables were obtained by deflating the nominal values by the product of: population, the PPP index and the implicit price deflator for U.S. GDP. In common with most authors in this field, we postulated a log-linear relationship between per capita health care spending and real income. Thus we defined  $h_{i,t}$  and  $y_{i,t}$  as the natural logarithms of real, US dollar, per capita values of health care expenditure and income, respectively, in country  $i$  in year  $t$ .

As in GSJA, we attempted to control to some extent for a factor that is often cited as a major contributor to rapid growth in health care spending in many countries, namely an aging population. For this purpose we used data on the number of people 65 years of age, or older, in each of the countries. These data are published in the UN Demographic Yearbook. While this publication appears annually, the data we wanted appear irregularly, even for the countries in the OECD. The data for Ireland and Spain were so sparse that we excluded them, leaving a total of 18 countries to be analyzed. For a number of other countries, missing observations were filled in using alternative data sources or interpolation: See Appendix A.

### 3 Replicating Previous Results

We wondered whether the application of OLS to cross-sections of data on all of the 18 countries at one or several years would yield results similar to those obtained in previous research. In particular we were interested in whether we would find estimates of the income elasticity of demand for health care which were “significantly” above one. For this purpose we formulated the model

$$h_{i,t} = \alpha_0 + \alpha_1 y_{i,t} + \alpha_2 a_{i,t} + u_{i,t} \quad \text{for } i = 1, \dots, 18. \quad (1)$$

We obtained the OLS point estimates and standard errors for all the years for which no interpolation of  $a_{i,t}$  was necessary. Some of the results are shown in Table 1 in which the row labelled “RESET” contains values of the statistic for a test that the true regression is linear and the row labelled “LM” contains



values of  $18 \times R^2$  from the regression of the squared residuals on the squared fitted values, which is a diagnostic against heteroskedasticity . Both these test statistics are asymptotically  $\chi_1^2$  under the null hypothesis.

Although the heteroskedasticity diagnostics may appear large, the asymptotic p-values are all above .05. The other diagnostics are small and the goodness of fit is somewhat higher than is often seen in cross-section studies. Thus, these results seem to support the accepted view that the income elasticity of demand for health care is quite a bit larger than one. They also seem to show that from 1960 until 1970 increases in the proportion of persons aged 65 or over had no significant impact on the demand for health care. However, in 1975 the estimated positive impact of this proportion grew to more than twice its standard error and the estimated income elasticity became somewhat larger than in the earlier years.

These changes suggest that more careful attention should be paid to the time series nature of the series. Also, this model contains many parameters compared to the number of observations and it ignores contemporaneous correlation between the errors of different cross sections. A popular way to address these problems is to pool the cross sections over time (see Gerdtam et. al. (1992)) and then use a generalized-least-squares (GLS) procedure which allows for first order autoregressive errors for each country, differing error variances between countries, and contemporaneous correlation between the errors of different countries; see Kmenta (1986) and White et. al. (1987). Of course, these results are conditional on the interpolated-extrapolated values for  $a_{i,t}$ . The GLS procedure stacks the values of  $h_{i,t}$ ,  $y_{i,t}$  and  $a_{i,t}$  one country above another, Australia at the top, U.S.A. at the bottom, into the vectors  $h$ ,  $y$  and  $a$  each of which contains 504 entries (18 countries  $\times$  28 years). The model is now

$$h = \gamma_0 + \gamma_1 y + \gamma_2 a + v \quad (2)$$

The residuals from first stage OLS fits of  $h_{i,t}$  on a constant,  $y_{i,t}$  and  $a_{i,t}$  are used to form the estimates of the first order autoregressive coefficients for each country, the country error variances, and the contemporaneous correlations between country errors. These first stage estimates are used to build up an estimate of the covariance matrix of  $v$ . We found it necessary to use the first order autocorrelations of the country residuals to estimate the autoregressive coefficients as some of the OLS estimates from the direct autoregressions of

the residuals exceeded one<sup>3</sup>. We shall return to this point later. The results of the GLS procedure are given in Tables 2.

These results are similar to those in Table 1 in that they seem to support the accepted belief that the demand for health care is elastic with respect to income. But now increases in the proportion of older people seems to have a significant (at the 1% level) positive effect on the demand for health care, as intuition would suggest. The fairly wide dispersion among the country error variances, shown in Table B.1 of Appendix B, is not surprising. But the large positive estimates of the autocorrelations, also in Table B.1, together with the result that some OLS estimates of the autoregressive coefficients exceeded one are reason to doubt the adequacy of this model.

## 4 Our Model

In formulating our model we sought to rectify some of the deficiencies of the simple model and econometric approach corresponding to equation (2). Chief among these was the failure to deal with the time series structure of per capita health care expenditure and income: See the discussion at the end of the previous section. *A priori*, both these variables are nonstationary but they may be linked by a long run equilibrium or co-integrating relationship (see e.g. Engle and Granger (1987)) which would imply an error correction model for the short run dynamics. Since  $a_{i,t}$  is bounded between zero and one we believe it to be stationary.

A second weakness of most earlier work is its failure to adequately deal with the impact of technical change on health care expenditure. Such change is often modelled as a linear time trend which suggests that the nonstationarity of health care expenditure should arise, at least in part from a deterministic trend. Then if income and health care expenditure are to be co-integrated income must also contain a deterministic trend.

Finally, we have retained the age distribution variable. However, we caution the reader that the interpolation and extrapolation to which some of the demographic data have been subject makes them suspect<sup>4</sup>.

---

<sup>3</sup>These computations were done with SHAZAM. The data transformations and interpolation discussed in the previous section were done with TSP. The results in Table 1 were produced with J. Breslaw's GAUSSX shell for GAUSS.

<sup>4</sup>It has been suggested to us that, although 65 is the cut-off age conventionally used

As noted previously, different countries with the same real per capita incomes and age distributions may spend different amounts on health care because of differing institutional arrangements regarding its financing and delivery. In a log-linear specification this implies different intercepts for different countries.

Our model combines these ideas with the idea of pooling explored in the previous section. Its first component is a long run relationship linking the nonstationary levels of the variables and the deterministic trend.

$$h = \sum_1^{18} \lambda_{0,i} d_i + \lambda_1 y + \lambda_2 a + \lambda_3 s + w \quad (3)$$

The first 17 of the  $d_i$  are dummy variables with the value one for each of the 28 observations corresponding to country  $i$ , ordered alphabetically, as in Table B.1, and zero elsewhere. The last one,  $d_{18}$ , is one for each observation. Since the last country in the ordering is the U.S.A., the first 17  $\lambda_{0,i}$  coefficients show deviations in the other countries levels as compared to the intercept which shows the U.S. level. The variable  $s$  captures the deterministic trend by running from 1 to 28 for each country. The error vector  $w$  is a realization of a weakly dependent, heterogeneous process whose form need not concern us if (3) is indeed a co-integrating relationship. We have included  $a$  at this stage to allow for the possibility that it is nonstationary and cointegrated with  $h$ . If, as we suspect *a priori*,  $a$  is stationary the value of  $\lambda_2$  would be zero and  $a$  would be absorbed into  $w$ .

The second component models short run dynamic behaviour:

$$\Delta h = \beta_0 + \beta_1 \Delta y + \beta_2 \Delta a + \beta_3 \hat{w}_{-1} + e \quad (4)$$

where  $\Delta$  is the first difference operator and  $\hat{w}_{-1}$  is the lagged estimate of  $w$ . If (3) is a co-integrating regression then  $\beta_0 = \lambda_3$  and  $\beta_3 < 0$ . In this case we believe *a priori* that  $\beta_1 < \lambda_1$ , since we predict that the long run adjustment of health care expenditure to changes in income will be larger than the short run adjustment. Also, we suspect that  $\lambda_2$  may be 0 since we doubt that  $a$  is nonstationary. Of course, that would not imply  $\beta_2 = 0$  since there could still be short run effects of changes in the age distribution on changes in

---

to define the aged population, health services costs tend to rise more dramatically at a higher age (say, 75) than at 65. In view of the potential data problems and the mixed results with respect to the age variable's significance (see below), we have not thought it worthwhile to experiment with alternative cut-off values.

health care expenditure. On the other hand, if we believed that (3) was not a co-integrating relationship we would restrict  $\beta_3$  to be zero because then  $w$  and, hence,  $\hat{w}_1$  would be nonstationary.

The first step in the econometric analysis of this model was to investigate the nature of the nonstationarity of  $h$ ,  $y$  and, in spite of our prior belief,  $a$ . We used two procedures for this. The first was a variant of the popular Said-Dickey (1984) test . The relevant equations for this test were

$$\Delta h = \delta_0 + \delta_1 h_{-1} + \delta_2 s + \delta_3 \Delta h_{-1} + \delta_4 \Delta h_{-2} + u_h, \quad (5)$$

$$\Delta y = \theta_0 + \theta_1 y_{-1} + \theta_2 s + \theta_3 \Delta y_{-1} + \theta_4 \Delta y_{-2} + u_y. \quad (6)$$

and

$$\Delta a = \phi_0 + \phi_1 a_{-1} + \phi_2 s + \phi_3 \Delta a_{-1} + \phi_4 \Delta a_{-2} + u_a \quad (7)$$

Estimates of the coefficients of these equations and of their standard errors were calculated by both OLS and the pooled GLS technique discussed earlier. Both sets of results appear in Table 3. The rows labelled LM contain values of the LM statistic for tests for first order serial correlation of the errors. They are asymptotically standard normal under the null of uncorrelated errors.

It is not clear that tables of critical values of the for test of the null hypotheses that  $\delta_1$ ,  $\theta_1$  and  $\phi_1$  are zero are applicable here given the way these variables were constructed. Nevertheless, straightforward reference to tables by MacKinnon (1991) would lead to the rejection of the null hypothesis of a unit root, after detrending, for both  $h$  and  $y$ , but not for  $a$ <sup>5</sup>.

A disadvantage of the Said-Dickey test is the form of the null hypothesis which, in the frequentist paradigm, can be rejected only in the face of strong evidence for the alternative. An alternative test by Kwiatkowski, Phillips, Schmidt and Shin (1993) (KPSS) examines the residuals from the regression of the variable on a constant and a linear trend,  $s$  in this case. This test is attractive because the null hypothesis is that the series is stationary after detrending. It has the advantage of being insensitive to (stationary) autocorrelation or heteroskedasticity in the residuals but it has the disadvantage of being sensitive to the number of lags used to calculate the long run variance. Also, it is not designed to deal with variables constructed as ours have been.

---

<sup>5</sup>We also conducted the test with  $\phi_2$  restricted to be zero. The resulting t-ratios for  $\phi_1$  were -1.94 for OLS and -1.36 for GLS. Asymptotic critical values may be obtained from the standard normal distribution since the t-ratios for  $\phi_0$  exceeded 3.5 in both cases.

Nevertheless, applying the test using a lag of 22 (which is just less than the square root of the sample size) gave values of the test statistic of .304 for  $h$ , .417 for  $y$  and .128 for  $a$ . The 5% right tail asymptotic critical value is .146. Thus, these values would lead to rejection of stationarity, after detrending, for both  $h$  and  $y$  but not for  $a$  – that is, the opposite result as that from the Said-Dickey test.

Contradictory results such as these are not uncommon; see DeJong et. al. (1992) and Kwiatkowski et. al. (1993) . If we conclude that  $h$ ,  $y$  and  $a$  are trend stationary then OLS would be consistent for the parameters of an equation like (3), with the detrended versions of  $h$ ,  $y$  and  $a$  in place of their levels and without  $s$ . Moreover, standard central limit theorems<sup>6</sup> would apply so OLS would provide a basis for asymptotic inferences about the parameters of such an equation. But these results would be identical to OLS applied to (3) as shown.

On the other hand, if we conclude that  $h$ ,  $y$  and  $a$  are nonstationary, even after detrending, OLS, although super-consistent (see Engle and Granger (1987)) for the coefficients of (3), would be asymptotically inefficient and would not provide consistent estimates of standard errors. In this case we could use Park's (1992) canonical cointegrating regression (CCR) estimator, which is consistent for both the coefficients and their standard errors. It allows the inclusion of deterministic trends, like  $s$ , and dummy variables, like the  $d_i$ .

In the face of this uncertainty about the true time series character of our variables it seemed prudent to estimate the coefficients of (3) under both scenarios. This will also allow us to see whether our conclusions are robust to changes in the assumed nature of the nonstationarity. Results using both OLS and GLS, under the assumption of trend stationarity, are presented in Table 4. The row labelled U.S.A. contains results for the intercept, which is the level for the U.S.A., while the other country rows show deviations from this intercept.

At first glance these results, especially those for GLS, seem to indicate that the income elasticity is much lower than is commonly supposed. The GLS results also suggest that there is a positive long run impact on health

---

<sup>6</sup>Standard central limit theorems would also apply in two other cases: first, if  $a$  is stationary so that the number of deterministic trends is equal to the number of random, nonstationary regressors (see Park (1992)) secondly, if we imagined both the number of countries and the number of years growing large at the same rate (see Quah (1989)).

care expenditure of increases in the proportion of old people in the population. The results for the country dummies are what one would expect: Compared to the U.S.A. , other countries (with the possible exception of Sweden) have more negative intercepts.

However, we are reluctant to put too much faith in these results because of several warning signs. First, the value of  $R^2$  for the OLS results was .978 but the value of the Durbin-Watson statistic was only .345. An inequality in this direction between these statistics indicates that the regression is "spurious"; i.e. it contains variables which follow a random walk. A second warning in the same direction is the first order autocorrelation of the OLS residuals of .825, which is, of course, directly related to the low value of the Durbin-Watson statistic. Finally, several of the autocorrelations appearing in Table B.2, in Appendix B, are over .9.

These warnings suggest that the application of OLS and GLS to (3) is inappropriate because even after detrending some of the variables involved are nonstationary. Hence, we turned our attention to the CCR estimator. This procedure requires the specification of a lag length for the estimation of the long run variance (we used 22) and the partitioning of the right side variables into those which are cointegrated with  $h$  and those which are deterministic, like  $s$  or the  $d_i$ . Our preliminary testing convinced us that  $h$  and  $y$  are most likely cointegrated but left us uncertain about the status of  $a$ . Therefore, we used CCR to estimate the coefficients of (3) assuming both  $y$  and  $a$  were cointegrated with  $h$  and again omitting  $a$ . The results<sup>7</sup> appear in Table 5.

A test of the null hypothesis that  $w$  is stationary, i.e. that (3) really is a cointegrating regression, is given by Shin (1992) based on residuals from any asymptotically efficient estimates of the coefficients of (3); we used CCR residuals  $\hat{w}$ . Shin estimated exact critical values, which depend upon the number of stochastic regressors in the equation, by Monte Carlo. The estimated 10% critical value for a cointegrating regression with a level, a linear trend and one stochastic regressor is .097. With two stochastic regressors, as well as a constant and a trend, the estimated 10% critical value is .081. The values of the test statistic are given in the last line of Table 5. We concluded that  $w$  was indeed stationary: i.e. (3) is indeed a cointegrating relationship.

These results are not greatly different from those reported in Table 4. The

---

<sup>7</sup>These results were obtained using S. Ouliaris' COINT procedures for GAUSS.

estimates of the income elasticity are less than those commonly obtained. The asymptotic p-values for  $H_0 : \lambda_1 \leq 1.0$  against  $H_1 : \lambda_1 > 1.0$  were .2791 when h was assumed cointegrated with both a and y, and .3280 when a was omitted. Thus we are unable to reject  $H_0$  in either case. Also the level dummies for all countries except Canada and Sweden were significantly less than zero at the 5% level. One difference between these results and the earlier ones is that now a has no significant long run impact on h: The asymptotic p-value for  $H_0 : \lambda_2 = 0$  against  $H_1 : \lambda_2 \neq 0$  was .4733 when h was assumed cointegrated with both a and y. This is consistent with a being stationary.

To investigate short run dynamics the value of  $\hat{w}_{-1}$  from the CCR without a was used in (4). Of course the unrestricted GLS estimate might fail to obey the restriction  $\beta_0 = \lambda_3$  implied by cointegration. However, it is easy to impose this restriction and both sets of estimates appear in Table 6. The rows of that table labelled Q(4), Q(7) and Q(14) give values of the portmanteau statistic for the residuals at the lag length in parentheses. They are asymptotically  $\chi^2$ , with degrees of freedom equal to the lag length, under the null hypothesis of uncorrelated errors. In performing the pooled GLS the first stage OLS estimates of the autoregressive coefficients were all absolutely less than one. The estimated autocorrelations and variances are presented in Table B.3 in Appendix B. (Note that the entries in Table B.3 are almost all much lower than the corresponding entries in Tables B.1 or B.2. The only exception is the estimated variance for Denmark which is slightly larger in Table B.3 than in Table B.1.)

The results in Table 5 show that the country dummies are important causal variables in the long run equation. Omitting these variables would be a specification error which we expect would have an impact on the estimates of the remaining coefficients. To see the effect of this omission in our case we use the same CCR procedure as produced the results in Table 5, but without the country dummies: The results are in Table 7. Wald statistics for testing the validity of the restrictions appear in the row with that label. They are asymptotically  $\chi^2$  under the null that the restrictions are valid with degrees of freedom equal to the number of restrictions imposed.

The large values of both the Wald statistic and the Shin statistic, which is larger than the estimated 5% critical values, in Table 7 confirm that omitting the country dummies was indeed a specification error. What is interesting is the extent to which this specification error increased the point estimates of the income elasticities.

Table 7 also shows the effect of omitting the time trend  $s$  from equation (3). In this case too we obtain a much increased estimate of the income elasticity. In contrast to the Table 5 results, the estimates of the country dummy coefficients are absolutely smaller, some are even positive, with bigger standard errors. Also  $a$  now appears to have an important positive *long run* effect on  $h$ . But this can only be true if  $a$  can grow indefinitely in the same way that  $h$  and  $y$  can: This is something we find *a priori* incredible, given the bounded nature of  $a$ . We believe instead that the influence of  $\lambda_3$ , the  $s$  coefficient, has been absorbed into the estimates of  $\lambda_1$  and  $\lambda_2$ . The value of the Shin statistic in this case does not lead us to reject of the null hypothesis of cointegration but the large value of the Wald statistic does convince us that omitting the trend was a specification error.

## 5 Conclusions

Perhaps the most significant aspect of the results in Tables 5 and 6 is that our point estimate of the long-run income elasticity of aggregate health care spending with respect to real income is close to unity, and that the upper limit of our 95% confidence interval is less than 1.25, a level lower than the point estimates in many earlier studies. Our estimate of the short run elasticity (from Table 6) is even smaller: We are 95% confident that it lies between .5604 and .6926. Our conclusion is that many earlier estimates have been biased upward because of the lack of attention to the time series structure of the data and the omission of country-specific factors correlated with income, on the one hand, and the confounding of time trend and real income effects, on the other.

As we see it, the significance of this is that there is nothing inevitable about health care spending accounting for a growing share of total resources over time. Holding institutional factors constant, and in the absence of cost-increasing technological change, our results would predict an essentially constant share.

Although our specification does not allow us to draw any conclusions about which institutional factors are responsible for varying country-specific effects, our results in Table 7 for the country dummies suggest that they are important. The analysis in GSJA (Table 4, p. 303) points to factors such as the share of inpatient spending in total health care cost, the share



of public-sector financing, and use of the fee-for-service method of paying physicians, as important determinants. Countries where the institutional organization leads to particularly low spending include Japan, the United Kingdom, and Belgium. Although the United States, as the country with the highest per capita income in the sample, can be expected to have the highest per capita health care spending, our results show that even controlling for income, the U.S. system is significantly more expensive than those in most other countries. However, the coefficients for the country dummies for several countries, notably Canada and Sweden, indicate no significant country-specific effect in comparison with the United States.

Our results show that increases in the proportion of the population aged 65 years or older have a large positive effect on health care expenditure in the short run but no effect in the long run. However, we are unsure how much of this result is due to the interpolation and extrapolation of the a series. It is also interesting to note that when we analyzed (3) without a time trend (but with the country dummies), we obtained a large positive estimate, with a large t-ratio, of the coefficients of the age variable (as well as larger estimates of the income elasticity than those in Tables 4 and 5). This raises the possibility that there is some confounding of age effects and the time trend.

Nevertheless, we interpret our estimates as giving strong evidence that technological change has had the effect of increasing real expenditure on health care by some 2.5% per year, even after correcting for the effect of income growth. This finding strikes us as extremely important. As observed previously, technological change may, in principle, shift the income-health expenditure relationship either up or down. Moreover, as argued in Baumgardner (1991), the direction of the effect may depend on institutional arrangements and incentives affecting the nature of medical research and the speed with which different kinds of technology is adopted. Seen in this light, our results can be thought of as lending support to those who argue that in the health care sector, industrialized countries have not come up with very good answers to the questions of how to promote the appropriate use of new technological possibilities, and of how allocate R&D in directions that promise the most favourable cost-effectiveness combinations. Finding better ways of exploiting the potential of new technology, therefore, represents one of the most important challenges for future health care policy.

## A Interpolation of Age Data

Missing observations on the proportion of the population aged 65 years or older were obtained: for Australia from *Australian Demographic Statistics Quarterly*, for Canada from *Annual Estimates of Population by Sex and Age for Canada and the Provinces 1971 to 1979*, for the U.K. from *Statistical Abstract for the United Kingdom* and for the U.S. from the Citibase databank. Table A.1 shows the years for which data were missing for each of the 18 remaining countries. Complete series were available for: Australia, Canada, Japan, U.K. and U.S.A. Age data were available for every country for the years: 1960 to 1965, 1967, 1968, 1970, 1971, 1975 and 1978.

Let  $a_{i,t}$  be the proportion of the population of country  $i$  aged 65 years or over in year  $t$ . Plots of  $a_{i,t}$  over the period 1960 to 1987 appeared to be dominated by local trends rather than by random shocks. This suggested the possibility of interpolating the missing value(s), if there was an observation following the gap, or extrapolating them, if they occurred at the end of the sample, using trends which are linear only locally, rather than globally. The first step in the interpolation was to use the Holt-Winters method (see e.g. Harvey (1981)) to create a smoothed series,  $\tilde{a}_{i,t}$ , running from the beginning of the sample, 1960, up to the year immediately preceding the gap. Let  $r_{i,t}$  be the value of the local trend in  $a_{i,t}$ . Then

$$\tilde{a}_{i,t} = \eta a_{i,t} + (1 - \eta)\tilde{a}_{i,t-1} + r_{i,t} \quad (8)$$

and

$$r_{i,t} = \xi(\tilde{a}_{i,t} - \tilde{a}_{i,t-1}) + (1 - \xi)r_{i,t-1}, \quad (9)$$

where  $\tilde{a}_{i,0} = a_{i,0}$ ,  $r_{i,1} = a_{i,1} - a_{i,0}$  and  $t = 0$  in 1960. We set  $\eta = \xi = .7$  for all cases.

If the  $a_{i,t}$  series stopped at year  $t$  and began again in year  $t + m$  the missing values were interpolated by

$$\hat{a}_{i,t+j} = \frac{(m+1-j)\tilde{a}_{i,t} + ja_{i,t+m}}{m+1} \quad \text{for } j = 1, \dots, m. \quad (10)$$

If the  $a_{i,t}$  series ended  $m$  years before the end of the sample, when  $t = T$ , the missing observations were extrapolated by

$$\hat{a}_{i,T-m+j} = \tilde{a}_{i,T-m} + jr_{i,T-m} \quad \text{for } j = 1, \dots, m. \quad (11)$$

The results for France and West Germany are shown on Figures 1 and 2.

## **B First Stage Estimates of Autocorrelations and Variances**

This appendix presents the estimates of the first stage OLS estimates of the autocorrelations and variances which are used to form the estimated covariance matrix for the GLS estimates of the coefficients of equations (2), (3) and (4) in Tables B.1, B.2 and B.3, respectively. To save space the estimated contemporaneous covariances between country errors are not shown.

## References

Australian Bureau of Statistics, Australian Demographic Statistics Quarterly (Canberra, Australia).

Baumgardner, J. R., 1991, The Interaction between forms of insurance contract and types of technical change in medical care, *Rand Journal of Economics* 22, 36-53.

Central Statistical Office, 1982, Statistical abstract for the United Kingdom (His Majesty's Stationary Office, London).

DeJong, D. N. , J. C. Nankervis, N. E. Savin and C. H. Whiteman, 1992, Integration versus trend stationarity in time series, *Econometrica* 60, 423-434.

Engle, R. F. and C. W. J. Granger, 1987, Co-integration and error correction: representation, estimation and testing, *Econometrica* 55, 251-276.

Gerdtham, U. G., and B. Jönsson, 1990, Pris och kvantitet i internationella jämförelser av hälso-och sjukvårdsutgifterna, *Ekonomisk Debatt* 7, 650-659.

Gerdtham, U. G., and B. Jönsson, 1991, Conversion factor instability in international comparisons of health care expenditure, *Journal of Health Economics* 10, 227-234.

Gerdtham, U.-G., J. Sjøgaard, B. Jönsson and F. Andersson, 1992, A pooled cross-section analysis of the health care expenditure of the OECD countries, in: Zweifel P. and H. E. Frech III, eds. *Health Economics Worldwide* (Kluwer Academic Publishers, Netherlands).

Harvey, A. C., 1981, *Time Series Models* (Philip Allan, Oxford).

Kmenta, J., 1986, *Elements of Econometrics*, 2nd ed. (Macmillan, New York).

Kwaitkowski, D., P. C. B. Phillips, P. Schmidt and Y. Shin, 1993, Testing the null hypothesis of stationarity against the alternative of a

unit root: how sure are we that economic time series have a unit root, *Journal of Econometrics* 54, 159-178.

MacKinnon, J. G., 1991, Critical values for cointegration tests, in: Engle, R. F. and C. W. J. Granger, eds. *Long-Run Economic Relationships: Readings in Cointegration* (Oxford University Press, Oxford) 267-276.

OECD, 1990, *Health Care Systems in Transition*, OECD Social Policy Studies No.7 (OECD, Paris).

Park, J. Y., 1992, Canonical cointegrating regressions, *Econometrica* 60, 119-144.

Parkin, D., A. McGuire, and B. Yule, 1987, Aggregate health care expenditure and national income, *Journal of Health Economics* 6, 109-127.

Quah, D., 1989, International patterns of growth: I. persistence in cross-country disparities, mimeo, MIT.

Said, S. E. and D. A. Dickey, 1984, Testing for unit roots in autoregressive-moving average models of unknown order, *Biometrika* 71, 599-608.

Shin, Y., 1992, A residual-based test of the null of cointegration against the alternative of no cointegration, mimeo, Michigan State University.

Statistics Canada, 1979, *Annual Estimates of Population by Sex and Age for Canada and the Provinces 1971 to 1979* (Queen's Printer, Ottawa).

United Nations, *Demographic Yearbook* (United Nations, New York).

White, K. J., S. A. Haun and N. G. Horsmann, 1987, *SHAZAM User's Reference Manual, Version 6* (Department of Economics, University of British Columbia, Vancouver).

Year	$\hat{\alpha}_0$	$se(\hat{\alpha}_0)$	$\hat{\alpha}_1$	$se(\hat{\alpha}_1)$	$\hat{\alpha}_2$	$se(\hat{\alpha}_2)$	$\bar{R}^2$	D-W	RESET	LM
1960	-3.75	.207	1.28	.126	.157	2.27	.887	2.20	.0222	1.76
1965	-3.51	.209	1.26	.123	-.831	1.66	.869	2.16	.115	2.93
1970	-3.88	.325	1.47	.130	-.320	1.79	.853	2.60	.150	1.74
1975	-4.54	.186	1.67	.0728	2.21	1.07	.914	2.44	.150	.659

Variable	Coeff.	Estimate	Std. Error
Constant	$\gamma_0$	-4.1281	.053507
y	$\gamma_1$	1.3972	.027938
a	$\gamma_2$	3.5200	0.33693
Buse's $R^2$		.8844	
Durbin-Watson		1.7585	

Table 3  
Said-Dickey Results

		OLS		GLS	
Variable	Coeff.	Estimate	t-Ratio	Estimate	t-Ratio
Constant	$\delta_0$	.557E-01	5.70	.860E-01	8.79
h <sub>-1</sub>	$\delta_1$	-.134E-01	-3.47	-.189E-01	-6.64
s	$\delta_2$	-.134E-02	-3.47	-.138E-02	-6.61
$\Delta h_{-1}$	$\delta_3$	.794E-01	1.73	.111	2.47
$\Delta h_{-2}$	$\delta_4$	.662E-01	1.51	.259E-01	.625
LM		.281		-.223	
Constant	$\theta_0$	.859E-01	8.79	.841E-01	10.5
y <sub>-1</sub>	$\theta_1$	-.253E-01	-5.43	-.228E-01	-7.24
s	$\theta_2$	-.398E-03	-1.98	-.668E-03	-3.58
$\Delta y_{-1}$	$\theta_3$	.220	4.78	.181	4.06
$\Delta y_{-2}$	$\theta_4$	-.653E-01	8.79	-.841E-01	10.5
LM		-.116		-.228	
Constant	$\phi_0$	.924E-03	3.62	.700E-03	5.06
a <sub>-1</sub>	$\phi_1$	-.323E-02	-1.36	-.154E-02	-1.37
s	$\phi_2$	-.598E-05	-7.50	.471E-05	.998
$\Delta a_{-1}$	$\phi_3$	.484	10.5	.502	11.6
$\Delta a_{-2}$	$\phi_4$	.196	4.21	.120	3.39
LM		-.838		-.624	

Table 4  
OLS and GLS Results for Equation (3)

variable	coefficient	OLS		GLS	
		estimate	std error	estimate	std error
y	$\lambda_1$	1.112	.04514	.8666	.03465
a	$\lambda_2$	-1.181	.6812	1.864	.4990
s	$\lambda_3$	.02412	.001571	.02777	.001112
Australia	$\lambda_{0,1}$	-.2662	.02857	-.3187	.05562
Austria	$\lambda_{0,2}$	-.1048	.04478	-.3499	.04465
Belgium	$\lambda_{0,3}$	-.3506	.03900	-.5389	.06277
Canada	$\lambda_{0,4}$	-.09197	.02670	-.1131	.05171
Denmark	$\lambda_{0,5}$	-.2083	.03437	-.4998	.09082
Finland	$\lambda_{0,6}$	-.2341	.03250	-.3665	.04703
France	$\lambda_{0,7}$	-.1216	.03597	-.3087	.02958
Germany	$\lambda_{0,8}$	-.09463	.03684	-.2974	.04496
Greece	$\lambda_{0,9}$	-.5268	.05811	-.8402	.06882
Iceland	$\lambda_{0,10}$	-.3217	.02865	-.3680	.03165
Italy	$\lambda_{0,11}$	-.3237	.03453	-.4951	.04312
Japan	$\lambda_{0,12}$	-.4053	.03740	-.4861	.03120
Netherlands	$\lambda_{0,13}$	-.1560	.02861	-.2904	.06553
Norway	$\lambda_{0,14}$	-.2750	.03670	-.4926	.06839
Sweden	$\lambda_{0,15}$	.04051	.04095	-.2003	.05610
Switzerland	$\lambda_{0,16}$	-.2806	.02688	-.3795	.08070
U.K.	$\lambda_{0,17}$	-.3645	.03834	-.5721	.04113
U.S.A.	$\lambda_{0,18}$	-3.077	.1083	-2.803	.08194



Table 5  
Cointegrating Regression Results

Variable	Coeff.	a Cointegrated		a Omitted	
		Estimate	Std. Error	Estimate	Std. Error
y	$\lambda_1$	1.056	.1003	1.045	.1010
a	$\lambda_2$	-1.127	1.498		
s	$\lambda_3$	.02588	.003540	.02430	.003088
Australia	$\lambda_{0,1}$	-.2790	.06212	-.2719	.06140
Austria	$\lambda_{0,2}$	-.1331	.09946	-.1862	.07654
Belgium	$\lambda_{0,3}$	-.3771	.08609	-.4171	.07035
Canada	$\lambda_{0,4}$	-.1095	.05886	-.08251	.05431
Denmark	$\lambda_{0,5}$	-.2306	.07544	-.2626	.06251
Finland	$\lambda_{0,6}$	-.2633	.07179	-.2644	.07319
France	$\lambda_{0,7}$	-.1385	.07969	-.1802	.06558
Germany	$\lambda_{0,8}$	-.1217	.08070	-.1558	.06352
Greece	$\lambda_{0,9}$	-.5946	.1289	-.6140	.1278
Iceland	$\lambda_{0,10}$	-.3368	.06278	-.3298	.06258
Italy	$\lambda_{0,11}$	-.3572	.07582	-.3712	.07357
Japan	$\lambda_{0,12}$	-.4400	.08267	-.4141	.07872
Netherlands	$\lambda_{0,13}$	-.1745	.06230	-.1829	.06296
Norway	$\lambda_{0,14}$	-.2988	.08089	-.3356	.06357
Sweden	$\lambda_{0,15}$	-.01796	.08951	-.02979	.05877
Switzerland	$\lambda_{0,16}$	-.2834	.05894	-.3082	.05204
U.K.	$\lambda_{0,17}$	-.3863	.08480	-.4297	.06579
U.S.A.	$\lambda_{0,18}$	-2.958	.2380	-3.026	.2254
Sum Squared Residuals		3.791		3.804	
Shin Test Statistic		.02656		.02483	

Table 6  
Error Correction Results

Variable	Coeff.	Unrestricted		Restricted	
		Estimate	Std. Error	Estimate	Std. Error
Constant	$\beta_0$	.03924	.001889	.02430 <sup>a</sup>	.003088 <sup>a</sup>
$\Delta y$	$\beta_1$	.49026	.03559	.6265	.03371
$\Delta a$	$\beta_2$	2.425	.7445	4.933	.6862
$\hat{w}_{-1}$	$\beta_3$	-.2424	.01982	-.2423	.02269
Buse's $R^2$		.4365		.4667	
Durbin-Watson		1.924		1.906	
Q(4)		2.56		2.17	
Q(7)		4.29		3.73	
Q(14)		11.27		11.78	
				<sup>a</sup> from Table 5	

Table 7  
Cointegrating Regression Results  
Omitting Country Dummies and Trend

Variable	Coeff.	$d_i$ Omitted		s Omitted	
		Estimate	Std. Error	Estimate	Std. Error
y	$\lambda_1$	1.500	.07053	1.600	.09427
a	$\lambda_2$	.5649	.8307	4.157	1.677
s	$\lambda_3$	.009263	.003256		
Australia	$\lambda_{0,1}$			-.02476	.06838
Austria	$\lambda_{0,2}$			-.02304	.1301
Belgium	$\lambda_{0,3}$			-.2714	.1124
Canada	$\lambda_{0,4}$			.07579	.06889
Denmark	$\lambda_{0,5}$			-.1745	.09902
Finland	$\lambda_{0,6}$			.03413	.07932
France	$\lambda_{0,7}$			-.05530	.1043
Germany	$\lambda_{0,8}$			-.08476	.1061
Greece	$\lambda_{0,9}$			-.007483	.1378
Iceland	$\lambda_{0,10}$			-.07479	.06865
Italy	$\lambda_{0,11}$			-.1292	.09239
Japan	$\lambda_{0,12}$			.01514	.07207
Netherlands	$\lambda_{0,13}$			.01589	.07574
Norway	$\lambda_{0,14}$			-.2619	.10644
Sweden	$\lambda_{0,15}$			-.05586	.1163
Switzerland	$\lambda_{0,16}$			-.3358	.07653
U.K.	$\lambda_{0,17}$			-.3349	.1114
U.S.A.	$\lambda_{0,18}$	-4.145	.1530	-4.558	.1382
Sum Squared Residuals		9.824		5.566	
Wald Statistic		2772		47.64	
Shin Test Statistic		.2179		.02286	

Table A.1	
Years of Missing Data on Number of People Aged 65 Years or Older	
Austria	1979, 1982, 1984, 1987
Belgium	1972, 1980, 1985, 1986, 1987
Denmark	1974, 1977
Finland	1969, 1972, 1977, 1982, 1987
France	1969, 1973, 1974, 1976, 1977, 1979, 1980, 1984
Germany	1969, 1972, 1974, 1983, 1987
Greece	1972, 1977, 1980, 1982, 1985, 1986, 1987
Iceland	1972, 1979, 1981, 1985, 1986, 1987
Italy	1969, 1973, 1979, 1983, 1985, 1986
Japan	1981
Netherlands	1973, 1982, 1983
Norway	1966, 1985
Sweden	1977, 1982, 1984
Switzerland	1972, 1973, 1981, 1983, 1984

**Table B.1**  
**Estimated Autocorrelations and Variances**  
**by OLS for Equation (2)**

	est. autocorr.	est. variance
Australia	.71314	.0040170
Austria	.93005	.0027625
Belgium	.96496	.0028675
Canada	.93128	.0018356
Denmark	.85126	.0031718
Finland	.93243	.0015145
France	.95916	.00081149
Germany	.89798	.0013296
Greece	.95684	.0017410
Iceland	.75129	.0066335
Italy	.87442	.0026736
Japan	.71320	.0033574
Netherlands	.97767	.0012537
Norway	.95131	.0026922
Sweden	.97115	.0014098
Switzerland	.98459	.0030127
U.K.	.98323	.0012665
U.S.A.	.96357	.0014264

**Table B.2**  
**Estimated Autocorrelations and Variances**  
**by OLS for Equation (3)**

	est. autocorr.	est. variance
Australia	.8197	.003433
Austria	.7179	.002041
Belgium	.8899	.001976
Canada	.9268	.001123
Denmark	.9175	.002912
Finland	.8867	.001048
France	.5394	.0004595
Germany	.8196	.001181
Greece	.8995	.001161
Iceland	.5875	.004830
Italy	.8605	.001745
Japan	.5166	.003372
Netherlands	.9507	.0009285
Norway	.8977	.002049
Sweden	.9001	.001164
Switzerland	.9239	.002065
U.K.	.8028	.001246
U.S.A.	.8243	.0004816

Table B.3  
 Estimated Autocorrelations and Variances  
 by OLS for Equation (4)

	Unrestricted		Restricted	
	est. autocorr.	est. variance	est. autocorr.	est. variance
Australia	.38595	.001864	.33847	.001901
Austria	.09485	.001792	.06170	.001926
Belgium	.14761	.001599	.21033	.001850
Canada	.49477	.000714	.44863	.000814
Denmark	.36179	.003192	.34187	.003031
Finland	.51958	.000713	.51934	.000710
France	-.09045	.000466	.03163	.000530
Germany	.40001	.000900	.32557	.000896
Greece	.35012	.001305	.25321	.001204
Iceland	-.16543	.003295	-.12798	.003713
Italy	.12765	.001396	.21701	.001364
Japan	-.09402	.002463	-.12796	.002544
Netherlands	.62740	.000561	.61246	.000596
Norway	.33668	.001773	.33680	.001830
Sweden	.41745	.000918	.41250	.000982
Switzerland	.16832	.001303	.20129	.001328
U.K.	.16028	.000975	.05987	.000905
U.S.A.	.13529	.000239	.22411	.000302