

What Determines the Health Care Expenditure of High Income Countries? A Dynamic Estimation

Yan Feng¹, Toby Watt², Anita Charlesworth², Grace Marsden¹, Adam Roberts², Jon Sussex^{3,4}

¹Office of Health Economics, London, United Kingdom

²The Health Foundation, London, United Kingdom

³RAND Europe, Cambridge, United Kingdom

⁴Cambridge Centre for Health Services Research, University of Cambridge, Cambridge, United Kingdom

Correspondence: Yan Feng, Office of Health Economics, Southside, 7th Floor, 105 Victoria Street, London SW1E 6QT, United Kingdom.

Received: August 11, 2017

Accepted: September 7, 2017

Available online: September 11, 2017

doi:10.11114/aef.v4i6.2586

URL: <https://doi.org/10.11114/aef.v4i6.2586>

Abstract

Constraining the rise in costs continues to be a major focus of health care policy in high income countries. It is important for governments to understand what is driving the rise in health care expenditure and what the impact will be over the coming years. This paper aims to provide an alternative econometric model to ascertain the determinants of health expenditure. Data from the OECD and IMS data bases for 18 OECD countries between 1988 and 2012 is collected. The analysis is at the year and country level. This study applies three methods: (1) panel data models with country fixed effects; (2) a first difference model; (3) a Vector Error Correction Model to account for the long run and short run effects as well as the endogeneity of the explanatory variables. The empirical results suggest that the use of different econometric specifications has a significant impact on both establishing the determinants of health expenditure and their magnitudes. Based on results from the Vector Error Correction Model, the GDP is considered as the only driver for country level health care expenditure growth. A 1% increase in the GDP is associated with a 1.1% increase in the health care expenditure.

Keywords: health expenditure, OECD countries, time-series, vector error correction model

1. Introduction

Controlling the rising costs of health care continues to be a major focus of health care policy internationally. In the UK for example, health expenditure (HE) growth is outstripping that of national income, and without increasing taxes or rearranging the way health care is provided this differential will place a lot of pressure on healthcare funding. Interestingly, there are substantial between-country differences in health care expenditure per capita, even amongst high income countries with generally healthy populations. For instance, in 2011, there were large disparities in per capita spend on healthcare across 18 OECD countries, varying from US\$2,355¹ per capita in Portugal, to more than double this, US\$5,121 in Norway. Such international comparisons can be used to investigate, and hopefully explain, differences in HE and its progression over time. Typically this is done through econometric analysis of macroeconomic data.

It is important for governments to understand what determine the rise in health care expenditure and what the impact will be over the coming years. In existing analyses of the drivers of national HE, the estimated relationship varies considerably. The primary differentiating issue is the model specification. This study aims to develop a modelling approach that could be used to establish the determinants of national HE for 18 high income countries.

Section 2 discusses the research background of the determinants of HE. Section 3 describes the data. Section 4 outlines the three econometric methods for modelling the HE data. The results are presented in Section 5, followed by a discussion in Section 6 and conclusions in Section 7.

¹ Analysis of OECD data from this study, in 2005 US\$.

2. Literature Review

Two streams of literature are reviewed: 1) the determinants of per capita HE and 2) the econometric approaches that have been used to model them. The review was focused on macro-level econometric models.

2.1 Determinants of National Health Care Expenditure

Drivers of HE can be separated into demographic and non-demographic factors (de la Maisonneuve & Oliveira Martins, 2013). Demographic drivers relate to the age and health status of the population, whilst non-demographic drivers include income and other variables (e.g. price of health care, technology and characteristics of a health care system).

2.1.1 Demographic Characteristics

Demography was often included as a determinant of HE, usually measured by the share of the population of an elderly age group, e.g. over 65 years old, or by the average age of the population (de la Maisonneuve & Oliveira Martins, 2013). It could be expected that an aging population would cause an increase in per capita HE. However, a number of studies have assumed that *proximity to death*, rather than age, is the key demographic driver of HE. This theory is consistent with observed data that individual HE tends to increase exponentially as patients approach death (Breyer & Felder, 2006; Seshamani & Gray, 2004; Felder et al., 2000). The mortality rates rise with age, all else equal. At the same time, due to advances in medicine, life expectancy rising. A simplistic proxy like average age, or share of the population over 65, does not take into account the latter. This is why readily measurable demographic characteristics are rarely found to be statistically significant in regression models of per capita HE (Hitiris & Posnett, 1992; Di Matteo & Di Matteo 1998; Leu, 1986).

2.1.2 Non-Demographic Characteristics: Income

The income elasticity of health spending has been debated widely, and is still uncertain (Lago-Penas et al., 2013). Previous results suggested that the income elasticity estimate depends on the countries reviewed, the time period and the estimation method. Lago-Penas et al. (2013) highlighted evidence on both sides of the debate: in the past health care has been found to be both a “luxury” good with an income elasticity greater than one (Liu, et al., 2011; Mehrara, et al., 2010; Parkin, et al., 1987; Leu, 1986; Newhouse, 1977), and a “necessity” with an income elasticity less than one (Baltagi & Moscone, 2010; Chakroun, 2009; Sen, 2005; Gerdtham et al., 1998; Gerdtham, 1992). The review also identified research suggesting that the income elasticity for national HE could be close to 1 (Gerdtham, 1992; Hitiris & Posnett, 1992). The author concluded that whilst the reviewed studies generally found the income elasticity to be positive, consensus is still not reached on whether health care is a necessity or a luxury good.

In the same study, Lago-Penas and colleagues conducted their own analyses (Lago-Penas et al., 2013). They found the estimated short-run elasticity to be around 0.3, and the long-run elasticity to be 1.1, which is similar to the results of Gerdtham’s dynamic analysis that was published in 1992.

Getzen in 2006 provided a comprehensive literature review that identified more than 40 empirical studies about the sources of variation in HE. Getzen (2006) found that as the unit of observation becomes larger (i.e. whether analysis is performed at the individual, the regional or national level), the correlation between HE and GDP increases in strength, i.e. an analysis at the person level finds HE to be inelastic with respect to income, while at country-level finds the opposite. This suggests that the income elasticity depends on the level of aggregation of data.

A recent OECD report in 2013 about national HE projection used 0.8 as the income elasticity, with a value of 1 for a sensitivity analysis (de la Maisonneuve & Oliveira Martins, 2013). This figure was used on the basis of a literature review as well as a linear panel data regression model with country fixed effects and time trends.

2.1.3 Other Non-Demographic Variables

Various other non-demographic drivers of HE were discussed in the literature. The OECD report suggested that the relative price of health services, technological progress and underlying health policies and institutions are likely to be the key other non-demographic drivers for HE (de la Maisonneuve & Oliveira Martins, 2013). Given the lack of appropriate variables representing many non-demographic drivers, the OECD report left an unexplained “residual” factor in its model explaining HE growth. It has been noted that the residual effect is large, and accounts for more than 50% of changes in HE (Medeiros & Schwierz, 2013).

Prices in the health sector relative to the rest of the economy are clearly a driver of HE. If inflation in any sector is higher than that for the rest of the economy, then real expenditure in that sector increases, other things being equal. In health, as in other areas of the economy where labour costs feature as a large part of the whole and productivity growth is constrained by the face to face nature of many of the services delivered, this increase is predominantly driven by what’s known as the “Baumol effect” or “cost-disease” (Baumol, 1967). The relatively high cost increases in these sectors reflect the negative productivity differential and the equalisation of wages across the economy i.e. prices for health services will rise relative to other prices because wages in low productivity sectors must keep up with wages in the economy as a whole.

New technologies are being developed and launched all the time. One of the key categories of new technologies in health care is new pharmaceuticals and medical devices. Governments and other health care payers have traditionally seen new medicines as one of the sources for increasing cost, at least in the short term. This is either because the new drug or other technology improves the quality of health care, but is more expensive than the treatment it is replacing, or it provides treatment for a need that could previously not be met. It counters to most industries, where new technologies tend to reduce costs. The existence of new technologies, especially medicines, also has the effect of reinforcing patients' expectations for care. On the other hand, new medicines and other technologies are also a source of productivity growth: better patient outcomes per pound spent.

One would expect systematic characteristics of each country's healthcare system to affect their HE. For instance the manner of service provision (e.g. the existence of a gatekeeper to the provision of secondary services) or the manner of health financing, provider payment mechanisms (e.g. whether the systems follow reimbursement, contract or integrated systems). In the latest publication of 2016 by de la Maisonneuve and colleagues, they explored the policies and institutional determinants of health spending in the OECD countries between 2000 and 2010 (de la Maisonneuve et al., 2016). There are 20 health policies and institutional indicators included in their empirical analysis, from both the supply-side aspects (e.g. provider payment, provider competition) and the demand-side aspect (e.g. gatekeeping, cost-sharing). The results suggested that the estimate of the income elasticity of public health expenditures becomes larger than one when health policy and institutional indicators were included as the independent variables of modelling the public health expenditure. Those 20 indicators were dropped out when the fixed effects method applied, as they are all time-invariant. The fact that these attributes tend to be either fixed or at least sticky over time means one may control for them using fixed effects analysis.

2.2 Econometric Methods

Gerdtham & Jönsson (2000) performed an extensive literature review of the analysis of aggregate data on the determinants of HE in OECD countries. The analysis of macroeconomic HE has mostly been based on simple demand theory – namely calculating income elasticity of demand by regressing HE on gross domestic product (GDP). Early versions of these analyses were performed on cross-sectional data. Newhouse (1977) used 1971 data from 13 countries. He regressed GDP on HE and found that 92% of the variation in HE was explained by GDP. On this basis, he concluded firstly that other factors such as prices, demographics and quality had little effect on HE, and secondly that health is a luxury good².

Parkin et al. (1987) argued that Newhouse's conclusions on income elasticity were based on microeconomic principles applied to macroeconomic data, causing aggregation problems. Parkin and colleagues also highlighted the likelihood of misspecification of the model, due to probably both omitted variable bias and inadequate functional form. Attempts to include more variables in the multivariate cross-section analysis were made by Gerdtham et al. (1992) and Leu (1986).

The first attempts to refine the functional form of the model arrived with improvements in panel data methods. These enabled researchers to account for and test the presence of country specific, time invariant determinants of healthcare expenditure. The static panel data model allows for variable intercepts to represent country fixed effects as well as a time trend. It is established in the literature that in dealing with macro data from OECD countries, it is appropriate to take the fixed effects regression approach (Gerdtham & Jönsson, 2000) either with a time trend or without. A number of studies in the 1990s using panel data methods (Roberts, 1999; Gerdtham et al., 1998; Gerdtham et al., 1992; Hitiris & Posnett, 1992).

Gerdtham (1992), accepting the likely dynamic nature of the determinants of HE, confronted issues of lags and the dynamic response of HE to changes in exogenous regressors. This dynamic analysis was furthered later in the 1990's with the use of co-integration analysis. It became apparent that several of the variables used in previous panel data studies were non-stationary (e.g. HE and GDP), following major findings that applying OLS to non-stationary variables may lead to "spurious regressions" (Engle & Granger, 1987). Non-stationary variables are said to follow a "unit-root", meaning that the level of variable y_t at time t depends on the all of the values of y that came before it, i.e. (y_0, \dots, y_{t-1}) .

Many macroeconomic variables including HE and GDP might be non-stationary at levels but stationary at first difference. In this case one can estimate the relevant coefficients by running a regression analysis in first differences. However, this process may result in a loss of information regarding the long-run relationships between variables, so while in some cases it is acceptable, it may not be the best course of action. This is particularly true if there exists a long run equilibrium relationship between variables and therefore an error correction mechanism involved in determining HE. If this mechanism exists and is omitted, then the use of first differences in a standard linear regression is a misspecification and is likely to result in biased estimates.

² Coefficients for income elasticity varied from 1.13-1.31 (Newhouse, 1977).

When two or more variables are integrated of the same order it is possible that they follow a co-integrating relationship i.e. a linear combination of these non-stationary variables is itself stationary. Granger (1981) was able to show that an OLS regression on a set of co-integrated variables will still produce consistent estimates. It is possible to test for such integration using the residuals from an OLS regression, such as in the test employed by Engle & Granger (1987)³.

Hansen & King (1996) performed individual country-by-country unit root and co-integration testing. Based on their results they could not (generally) reject the unit root hypothesis for HE or GDP. Nor could they reject the null hypothesis that HE and GDP are not co-integrated, which provides no support for the existence of equilibrium co-integrating relationships between HE and GDP. They then suggested that panel data estimations of HE and GDP relationships are spurious. Blomqvist & Carter (1997) performed individual country by country and panel data unit root tests and individual country-by-country co-integration tests. They reached the same results as Hansen & King (1996) concerning the unit-root hypothesis, but they rejected the no co-integration hypothesis. Roberts (1999) performed similar tests and found evidence of a co-integrating relationship between HE and GDP to be inconclusive. McCoskey & Selden (1998) presented unit root test results for time series on per capita national HE and GDP in the OECD. Their results rejected the unit root hypothesis.

A recently published study in 2016 by Hauck & Zhang estimated the drivers of HE growth in 34 OECD countries between 1980 and 2012. The econometric model in this paper accounted for the heterogeneous effects of unobserved common factors on HE growth across countries. Furthermore, the method addressed model uncertainty that related to choice of regressors by Bayesian Model Averaging. The study found 16 significant drivers for HE growth, including the growth in GDP per head. The estimated elasticity of HE growth with respect to GDP per head is 0.77. A methodological limitation of this study is that the country-specific regressors are assumed to be stationary and exogenous. It does not allow for unit roots and other dynamic nature of the time series data.

This study adds empirical evidence to the discussion about the determinants of national HE. First, it uses a recent and 25 years' panel data set, looking at similarly high income OECD countries. Second, it uses a wider range of variables than in other studies, looking for new and richer equilibrium relationships.

3. Data

Data used in the empirical analyses are collected annually, from 1988 to 2012, for 18 OECD countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden and the United Kingdom. The country selection is driven by data availability for key variables, i.e. GDP per head and HE per head. The statistics summary and data source are reported in Table 1.

Table 1. Statistics summary and data source

Variable names	Mean	SD	Min	Max	Obs	Source
Total health spend ^a	62914.2	70250.0	3637.9	339130.6	450	OECD
GDP ^b	693264.6	705385.3	55235.3	3009809.0	450	OECD
Population ^c	24260	23932	3344	82534	450	OECD
Log health spend per head ^d	7.8	0.4	6.7	8.6	450	Derived
Log GDP per head ^e	10.2	0.2	9.5	11.0	450	Derived
Share over 65 ^f	0.2	0.0	0.1	0.3	450	OECD
Standardised mortality rate ^g	772.8	120.6	506.3	1115.0	450	OECD
Global pharma spend ^h	417478.5	201395.5	161050.4	760694.3	450	IMS
Relative health price ⁱ	1.0	0.1	0.4	1.6	425 ^k	OECD
Average wage ^j	37062.5	7327.9	14937.0	52838.0	425 ^l	OECD

^a It refers to private and public HE in million (US\$ at 2005 prices)

^b GDP in million (US\$ at 2005 prices)

^c Population per 1000

^d Log of HE per head (US\$ at 2005 prices)

^e Log of GDP per head (US\$ at 2005 prices)

^f Proportion of population over 65 years old

^g Death rate per 100,000 of the population

^h Global spend on pharmaceuticals in million (US\$ at 2005 prices)

ⁱ Health and social value deflator divided by GDP deflator

^j Average annual wage per head (US\$ at 2013 prices)

^k Information for Ireland is missing

^l Information for New Zealand is missing

³ Extensions of this test to panel data have been created by Johansen (1991) and McCoskey & Kao (1998). The Johansen fisher panel co-integration test has been applied to the empirical analysis of this study.

The selected variables in the empirical analysis followed the findings from the literature review in Section 2. HE is the key variable of interest. GDP is used as a proxy for income. The demographic characteristics of a country are measured by the proportion of the population aged above 65 years old and the standardised mortality rate. The standardised mortality rate could also be used as a measure of quality for health care. The relative health prices is measured by two variables, i.e. the national annual average wage per head, and the health and social value deflator adjusted by the GDP deflator. Technology is proxied by global pharmaceutical spend.

The data set is complete for GDP per head and global pharmaceutical spend (n=450). There are missing values reported in other four variables, i.e. HE (n=446), proportion of population over 65 years old (n=423), standardised mortality rate (n=434) and relative health price (n=341). It is assumed that there is a linear relationship between the missing value and available data points within a country for each variable. All missing values based on this rule for the four variables are imputed. The econometric analysis of the best model in this study is based on the data set with imputation. Sensitivity analysis is also provided to check the robustness of the results, use data set without imputation.

4. Econometric Methods

In the econometric analysis, three separate techniques were applied. In the first model, a fixed effects method was used with error terms clustered at country level. However, this raised concerns about the stationarity of variables. In the second model, a first difference method was used with error terms clustered at country level. In the third model, a group of Vector Error Correction (VEC) models was implemented followed by the VEC residual portmanteau test. This approach addressed the key characteristics of the data, i.e. non-stationarity, endogeneity, serial correlation between error-terms, and the assumed long run equilibria between variables. The Pantula principle was applied for selection and to compare the performance between models. The performance of Vector Autoregression (VAR) on models that show no long term equilibrium has also been checked.

4.1 Fixed Effects Model

The first model uses a fixed effect method with error terms clustered at country level. The specification is reported by equation (1).

$$\log\left(\frac{HE_{it}}{N_{it}}\right) = f_i + \alpha Demo_{it} + \beta \log\left(\frac{P_{it}}{P_{Yi}}\right) + \gamma \log Q_{it} + \delta \log\left(\frac{Y_{it}}{N_{it}}\right) + \epsilon \log Tech_t + \theta t + \varepsilon_{it} \quad (1)$$

Where $\frac{HE_{it}}{N_{it}}$ refers to HE per head at country i and year t , which is derived from total HE_{it} and population N_{it} ; f_i refers to country fixed-effects, allowing HE to vary across countries; the demographic effect $Demo_{it}$ is captured by the share of the population over 65 years old; the relative price of health services $\frac{P_{it}}{P_{Yi}}$ is calculated as the health and social care value deflator P_{it} divided by the whole economy value deflator P_{Yi} ⁴; the quality of health care Q_{it} is proxied by the standardised mortality rate; $\frac{Y_{it}}{N_{it}}$ is GDP per capita for country i at year t ; technology is proxied by total global expenditure on pharmaceuticals $Tech_t$; t is a time trend; and ε_{it} is a random error term.

There are two assumptions for equation (1). First, all independent variables are weakly exogenous. Second, the error terms are assumed to be independent and identically distributed (i.i.d).

Wooldridge test for residual autocorrelation in panel data was applied after each regression analysis (Drukker, 2003; Wooldridge, 2002).

4.2 First-Difference Model

Assuming error terms in equation (1) are independent and identically distributed is a strong assumption for modelling time series data, as non-stationarity is a commonly observed characteristic for time series data. Non-stationary error terms violate the assumptions required for OLS to produce consistent estimates which can lead to meaningless results⁵.

Many macroeconomic variables including (typically) HE and GDP are non-stationary but integrated of order one –

⁴ The OECD average wage by country was used as an alternative proxy.

⁵ Granger (1981) showed that an OLS regression on a set of co-integrated variables will still produce consistent estimates. One is able to test for cointegration using the residuals from an OLS regression, such as in the test employed by Engle & Granger (1987).

meaning that if one takes first differences (subtract the previous value of each variable from the current value) then the variable becomes stationary. A panel regression is performed with all variables take their first differences. The specification of the first difference model is reported in equation (2).

$$\Delta \log \left(\frac{HE_{it}}{N_{it}} \right) = \alpha \Delta Demo_{it} + \beta \Delta \log \left(\frac{P_{it}}{P_{Yi}} \right) + \gamma \Delta \log Q_{it} + \delta \Delta \log \left(\frac{Y_{it}}{N_{it}} \right) + \epsilon \Delta \log Tech_t + \theta t + \varepsilon_{it} \quad (2)$$

4.3 Vector Error Correction Model (VECM)

4.3.1 Issues with Specification in Equation (2)

There are two issues with the specification in equation (2). First, it contains the strong assumption that all of the independent variables included are weakly exogenous i.e. that there is no correlation between the variables and the error terms. But one may argue that, for instance, the proportion of people at age above 65 years is an endogenous variable in modelling HE at country level. A higher proportion of people age above 65 years old in a country might lead to more HE in that country, while a large spend on health might have a positive impact on the overall health of a country and therefore contribute to an increase in the proportion of people aged over 65.

Second, the equation (2) assumes that there is no serial correlation between error terms. Again, this is a strong assumption. The error terms in modelling HE include a number of unobserved variables. With 25 years' panel data, the errors might show autocorrelations because of the unobserved variable(s). Therefore, an autocorrelation test should be applied to check the serial correlation of residuals after all regressions.

Furthermore, model 2 uses only short run data variation to analyse the drivers of HE. Any long run information contained in the data is taken out by the first difference. If there is long term equilibrium between variables, but the effect is omitted from the model, then the use of first differences in a standard linear regression is a misspecification and is likely to result in biased estimates.

4.3.2 Specification of the VECM

In order to account for the three issues with model 2, the VECM is considered. This model is a representation of the VAR model that is appropriate to use if there is at least one long-run equilibrium relationship between the variables present. Both models assume that the variables included are endogenous. For all VECM specifications a residual autocorrelation test was applied. The general specification of the VECM is presented by equation (3) (Engle & Granger, 1987). For an accessible introduction to empirical model building in economics see Granger (1999).

$$\Delta y_{it} = \sum_{k=1}^p \Gamma_{ik} \Delta y_{it-k} + \alpha(u_{it-1}) + \gamma + \theta t + \varepsilon_{it} \quad (3)$$

$$\varepsilon_{it} \sim N(0, \Sigma_u)$$

This model takes into account both the short run dynamics among the variables collected in the 6 x 1 vector $y_{it} =$

$\left[\frac{HE_{it}}{N_{it}}, Demo_{it}, \frac{P_{it}}{P_{Yi}}, Q_{it}, \frac{Y_{it}}{N_{it}}, Tech_t \right]'$, and the long-run structure presented by the vector of co-integrating residuals u_{it-1} :

$$\beta y_{it-1} + \mu + \rho t = u_{it-1} \quad (4)$$

Γ_{ik} is a matrix of autoregression coefficients, p is the number of lag intervals, α is a matrix of coefficients adjusting for short run dynamics to the long run equation, γ is a vector of intercept terms and t is the time trend, β is the matrix of coefficients for long run equilibrium, μ is a vector of intercept terms for long run equilibrium.

There are five specifications of the model represented in equation (3) with different assumptions on the inclusion of an intercept term or time trend in the cointegration equation (CE) and the short run dynamics (or VAR). The first specification of deterministic trend assumes that there is no intercept or trend in the CE ($\mu = 0$ and $\rho = 0$) and the VAR ($\gamma = 0$ and $\theta = 0$). The second assumes that there is no trend in the CE or the VAR, but there is an intercept in the CE. The third assumes that there is an intercept in the CE and both an intercept and a trend in the VAR. The fourth assumes that there is a trend and intercept in the CE, and an intercept but no trend in the VAR. The fifth specification assumes that there is an intercept and quadratic trend in the CE, and a linear trend in the VAR. Each of these specifications is performed for each combination of variables.

4.3.3 VECM Process

It takes three steps to specify a VECM. First, one needs to check whether any long-run equilibrium exist, using the Johansen Fisher Panel Co-integration test. Then, if the test suggests that at least one equilibrium does exist, the task is to choose the number of lag intervals (k). Allowing k to equal to 1, 2 or 3 have been experimented in this study. Finally, five specifications with different assumptions of intercept and trends in the CE and VAR that described in the previous section are implemented.

15 econometric models (3 choices of lag intervals x 5 structures of deterministic trend) are tested, one for each choice of variables. In total, there are 6 sets of variables and therefore 90 econometric specifications are experimented. The 6 sets of variables are the same as those in the first difference models. For specifications that show at least one long run equilibrium relationship, the autocorrelation test is applied to check for serial correlation in the residuals. For specifications that suggest no long run equilibrium, the VAR model (a model that only accounts for the short run dynamics) is applied, followed by an autocorrelation test.

4.3.4 Comparing the Performance between Models

The Pantula principle is applied to compare the performance between models. The comparisons of performance are only made between models where the residual portmanteau test suggests no autocorrelation for up to eight lags. The key statistics that are compared include the specification of the deterministic trend, the number of co-integrating relationships, the number of insignificant coefficients, the log likelihood statistics, the AIC statistics, the Schwarz Criterion statistics and the results from autocorrelation tests.

4.3.5 Sensitivity Analysis

The best model in this study, i.e. VECM, is applied in the sensitivity analysis by applying the data set with no imputed values.

The programmes used for data analysis and modelling were Eviews 8 and STATA/MP 13.1.

5. Results

5.1 Unit Root Tests

The results from Augmented Dickey–Fuller (ADF) Fisher unit root tests are reported in Table 2. The results suggest that all variables are non-stationary in levels. All variables are stationary at first difference in levels. Apart from the proportion of population aged above 65, all variables are stationary at first difference of logs.

Table 2. ADF Fisher unit root test

Variables	ADF Fisher Chi-square	Probability
In level		
Health spend per head	7.54	1.00
GDP per head	10.19	1.00
Share over 65	42.58	0.21
Standardised mortality rate	10.28	1.00
Global pharma spend	0.09	1.00
Relative health price	45.38	0.09
Average wage	44.24	0.11
In log		
Health spend per head	22.38	0.96
GDP per head	13.82	1.00
Share over 65	54.67	0.02
Standardised mortality rate	4.46	1.00
Global pharma spend	15.49	1.00
Relative health price	70.80	0.00
Average wage	59.83	0.00
In first difference		
Health spend per head	176.36	0.00
GDP per head	187.80	0.00
Share over 65	57.00	0.01
Standardised mortality rate	407.76	0.00
Global pharma spend	122.54	0.00
Relative health price	140.34	0.00
Average wage	149.05	0.00
In first difference of log		
Health spend per head	147.46	0.00
GDP per head	179.22	0.00
Share over 65	48.44	0.08
Standardised mortality rate	384.22	0.00
Global pharma spend	205.17	0.00
Relative health price	162.85	0.00
Average wage	176.96	0.00

^a Individual intercept are included in the unit root test equation.

^b Schwarz Information Criterion is used to automatically select the lag length in the unit root test.

^c Probability at 5% level is used as the threshold for accepting/rejecting the unit root null hypothesis.

5.2 Regression Results

In general, the results show that the specification of the model has significant impact on the results.

Results from six fixed effects regressions are reported between column 1 and column 6 in Table 3. The log of GDP per head is used to represent national income. Results from the 6 models suggest that the estimated value of income elasticity fluctuates between 0.74 and 0.87, all of which are statistically significant at the 5% level. The demographic characteristics are proxied by the proportion of the population aged above 65; this is not significant at the 10% level in explaining HE in any models. Quality of health care is measured by the log of mortality rate. Results in columns 3-6 suggest that the effect is not statistically significant at the 10% level. In column 5, the impact of relative health care price is found not to be statistically significant in explaining the HE. As an alternative measure of health price, the effect of average wage is also found not to be significant in column 6. Global pharmaceutical spend is used to measure health care technology. Results in column 5 suggest that a 1% increase in the global pharmaceutical spends is associated with a 0.17% increase in HE. The effect is statistically significant at the 5% level. The effect is also reported as statistically significant at the 5% level in column 6 with a magnitude of 0.21%.

Wooldridge test for autocorrelation in panel data strongly rejects the null hypothesis of no serial correlation for all 6 models. The results are reported in the last row of Table 3.

Table 3. Fixed effects method

Variables	Column 1 (t)	Column 2 (t)	Column 3 (t)	Column 4 (t)	Column 5 (t)	Column 6 (t)
log GDP	0.8683** (14.62)	0.8732** (12.89)	0.8707** (14.73)	0.8743** (12.92)	0.8131** (9.84)	0.7386** (9.13)
Share over 65		0.1643 (0.19)		0.1219 (0.15)	-0.2038 (-0.36)	0.8490 (1.02)
log mortality rate			-0.2226 (-1.25)	-0.2213 (-1.26)	-0.2427 (-0.98)	-0.1805 (-1.04)
log relative health price					0.2102 (1.67)	
log wage						0.3630 (1.69)
log pharm spend					0.1742** (2.93)	0.2051** (2.95)
Year trend	0.0177** (10.10)	0.0174** (8.38)	0.0140** (4.21)	0.0138** (3.80)	0.0001 (0.01)	-0.0026 (-0.34)
Constant	-36.5808** (-11.69)	-36.0630** (-10.18)	-27.6667** (-3.70)	-27.3350** (-3.46)	-1.2511 (-0.08)	0.0799 (0.01)
Number of obs	450	450	450	450	425	425
R2	0.8749	0.8773	0.8788	0.8805	0.8756	0.8849
Prob > F ^a	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

** p<0.05; * p<0.1

^a Wooldridge test for autocorrelation in panel data with null hypothesis of no first-order autocorrelation.

Columns 1-6 in Table 4 report results from the first-difference models. The estimated elasticity of HE with respect to income (0.33-0.44) is much lower than in the fixed effects models. The proportion of the population aged above 65 has a positive effect on HE in all models and is statistically significant at the 5% level in three models (reported in columns 2, 4 and 5). Mortality rate is not significant in explaining HE in any of these models. In column 5, the impact of relative health care price is found to be statistically significant in explaining the HE. As an alternative measure of health price, the effect of average wage is also found to be significant in column 6. Estimates for the impact of technology and relative prices on HE that are reported in columns 5 and 6 suggest positive and significant effects at the 5% level.

Table 4. First-difference method

Variables	Column 1 (t)	Column 2 (t)	Column 3 (t)	Column 4 (t)	Column 5 (t)	Column 6 (t)
D(log GDP)	0.4009** (4.84)	0.3894** (4.82)	0.4014** (4.85)	0.3899** (4.84)	0.4442** (5.80)	0.3319** (5.02)
D(Share over 65)		0.0357** (3.09)		0.0362** (3.03)	0.0335** (2.57)	0.0302 (1.63)
D(log mortality rate)			0.0282 (0.60)	0.0402 (0.85)	0.0061 (0.12)	0.0241 (0.46)
D(log relative health price)					0.2268** (3.48)	
D(log wage)						0.5099** (5.48)
D(log pharm spend)					0.1446** (4.15)	0.1440** (4.15)
Year trend	-0.0005* (-1.81)	-0.0003 (-1.03)	-0.0005 (-1.73)	-0.0003 (-0.94)	-0.0000 (-0.08)	-3.72e-06 (-0.01)
Constant	1.0576* (1.85)	0.5562 (0.93)	1.0348* (1.78)	0.51625 (0.84)	-0.0095 (-0.02)	-0.0425 (-0.06)
Number of obs	432	432	432	432	408	408
R2	0.1125	0.1273	0.1129	0.1280	0.1875	0.2299

** p<0.05; * p<0.1

90 different specifications of the VECM are experimented, of which two show at least one long run equilibrium between variables, and no residual autocorrelation up to lag 8 with a 10% significance level. The first of these VECMs includes two variables: log GDP per head and log health spend per head. The second includes three variables: log GDP per head, log health spend per head and log mortality rate. Both models contain two lag intervals for the short run

dynamics ($k=2$) with a trend and intercept in the co-integrating equation and an intercept but no trend in the VAR. The key statistics about the performance of the two models are reported in Table 5. The results from the VECM residual portmanteau tests of the two models are reported in Appendices A and B.

Table 5. Key statistics for two VECM

Key statistics	VECM 1	VECM 2
Number of lag interval	2	2
Deterministic trend	A trend and intercept in the CE and an intercept but no trend in the VAR	A trend and intercept in the CE and an intercept but no trend in the VAR
Co-integration rank	1	2
Number of insignificant coefficients	4 out of 12	15 out of 27
Log likelihood	1668.88	2584.34
AIC	-8.35	-12.88
Schwarz Criterion	-8.20	-12.52
No residual autocorrelations up to lag ($p<0.05$)	21	8
No residual autocorrelations up to lag ($p<0.1$)	18	8

Applying the Pantula principle to compare the performances between two models, the first model (with log GDP per head and log health spends only) might suggest a better performance than the second (with log GDP head, log health spend and log mortality rate). The first model reports a smaller number of co-integrating relationships, a smaller proportion of insignificant variables in the VAR and a larger number of lags with no-serial correlation in residuals than the second model. The second model shows a higher log likelihood statistic and lower AIC and Schwarz Criterion statistics. This is expected as the second model includes more variables than the first model.

Results from the best VECM are reported in Table 6. The magnitude of the effect of the time trend variable is rather small ($8.49e-05$) and not statistically significant at the 10% level. To increase the efficiency of the estimates, restrictions to time trend are imposed with a coefficient of zero. Results from the best model with this restriction are reported in Table 7.

Table 6. Best model without coefficients restrictions

Cointegration equation		Cointegration equation 1	
LnHE (-1)		1.00	
LnGDP (-1)		-1.09	
		(0.13)	
		[-8.35]	
Trend		8.49E-05	
		(0.00)	
		[0.37]	
Intercept		3.38	
Error correction		D(lnHE)	D(lnGDP)
Cointegration equation 1		-0.06	-0.02
		(0.01)	(0.01)
		[-4.53]	[-1.70]
D(lnHE(-1))		0.12	-0.04
		(0.05)	(0.04)
		[2.49]	[-0.87]
D(lnHE(-2))		-0.03	0.09
		(0.05)	(0.04)
		[-0.55]	[2.09]
D(lnGDP(-1))		0.14	0.25
		(0.06)	(0.05)
		[2.33]	[4.66]
D(lnGDP(-2))		0.33	0.01
		(0.06)	(0.05)
		[5.39]	[0.24]
Intercept		0.02	0.01
		(0.00)	(0.00)
		[7.44]	[4.80]
R-squared		0.23	0.09
Adj. R-squared		0.22	0.08
Sum sq. resids		0.41	0.31
S.E. equation		0.03	0.03
F-statistic		23.67	7.66
Log likelihood		800.26	852.21
IC		-4.01	-4.27
Schwarz SC		-3.95	-4.21
Mean dependent		0.03	0.02
S.D. dependent		0.04	0.03
Determinant	resid		
covariance (dof adj.)		7.72E-07	
Determinant	resid		
covariance		7.49E-07	
Log likelihood		1668.88	
AIC		-8.35	
Schwarz criterion		-8.20	

^a Standard errors in () & t-statistics in []

^b Sample adjusted for a period between 1991 and 2012.

^c There are 396 observations included after adjustments.

The results in Table 7 suggest that there is a long-run equilibrium between HE per head and GDP per head. The relationship is statistically significant at the 5% level. The income elasticity estimated in this long-run equilibrium model is 1.10. This implies that a 1% increase in GDP per head leads to HE per head increasing by 1.10% in the long run.

Table 7. Best model with coefficients restrictions

Cointegration equation		Cointegration equation 1	
LnHE (-1)		1.00	
LnGDP (-1)		-1.10	
		(0.13)	
		[-8.45]	
Intercept		3.46	
Error correction		D(lnHE)	D(lnGDP)
Cointegration equation 1		-0.06	-0.02
		(0.01)	(0.01)
		[-4.52]	[-1.70]
D(lnHE(-1))		0.12	-0.04
		(0.05)	(0.04)
		[2.48]	[-0.88]
D(lnHE(-2))		-0.03	0.09
		(0.05)	(0.04)
		[-0.58]	[2.08]
D(lnGDP(-1))		0.14	0.25
		(0.06)	(0.05)
		[2.33]	[4.66]
D(lnGDP(-2))		0.34	0.01
		0.06	0.05
		[5.39]	[0.25]
Intercept		0.02	0.01
		(0.00)	(0.00)
		[7.45]	[4.80]
R-squared		0.23	0.09
Adj. R-squared		0.22	0.08
Sum sq. resids		0.41	0.31
S.E. equation		0.03	0.03
F-statistic		23.63	7.65
Log likelihood		800.19	852.20
IC		-4.01	-4.27
Schwarz SC		-3.95	-4.21
Mean dependent		0.03	0.02
S.D. dependent		0.04	0.03
Determinant	resid		
covariance (dof adj.)		7.73E-07	
Determinant	resid		
covariance		7.49E-07	
Log likelihood		1668.81	
AIC		-8.35	
Schwarz criterion		-8.20	

^a Standard errors in () & t-statistics in []

^b Sample adjusted for a period between 1991 and 2012.

^c There are 396 observations included after adjustments.

In the short run equation of HE, the coefficient for the co-integrating equation is negative and statistically significant at the 5% level. This implies that if there is a positive departure from the long run equilibrium, then HE will fall in order to return to the equilibrium level. However, GDP does not respond to any departure from the long run equilibrium between GDP and HE. This is reported by the statistically insignificant coefficient of the cointegration equation in the short run equation of GDP.

In the short run, both GDP and HE respond to the lag of GDP and the lag of HE positively. HE per head responds to its own one year lag, one year lag of GDP per head and two year lag of GDP per head, significantly at the 5% level. GDP per head responds to its own one year lag and two year lag of HE per head significantly at the 5% level.

For the specifications that suggest there is no long-run equilibrium, the VAR model is applied and followed by an autocorrelation test. The results suggest that there is no VAR specification that results in no autocorrelation up to 8 lags. For these specifications there is serial correlation in the residuals, implying that there are dynamic relationships the VAR cannot account for. To apply the VAR model to the data set would therefore likely lead to misspecification.

There was no imputation required for GDP per head data. There were four missing observations reported in the HE data in 2012. A sensitivity analysis was conducted with VECM on a data set without imputation by dropping year 2012. The results suggest that the elasticity of income in the long run is statistically significant at the 5% level with a coefficient of

1.06. This is similar to the results in Table 7, i.e. 1.10. The coefficients of short run effects are also similar to the results in Table 7. An alternative method for inputting missing data is the Multiple Imputation technique (Hauck & Zhang, 2016). One might not expect the different methods to have a large impact on the results as this is only relevant for 4 observations.

6. Discussion

The three separate analyses performed in this study gave rise to rather different results for the magnitude of the national income elasticity of demand for health care expenditure.

In a regression of levels with country fixed effects the results suggest that the income elasticity is less than one, with a high R^2 term, but few statistically significant variables other than GDP. From this result one might conclude, as did in Newhouse (1977), that income is the foremost driver of HE. However, this first regression analysis should be treated with cautious, as in dealing with time series data there is always a risk that variables are non-stationary. Performing OLS regression on non-stationary variables can lead to spurious inferences and should be avoided.

Having found through ADF tests that many of the variables used in this study are non-stationary and in fact appear to follow a unit root. The first difference is calculated for all variable – as time series data that follows a unit root might show “difference stationary”. The results of this set of regressions have a slightly different interpretation, due to the short-term information that is used. The income elasticity estimate here is much lower than in the standard panel regression because it denotes only the instant reaction of HE to changes in GDP. The estimated coefficients are therefore lower in the first difference regression, as well as the explanatory power of the model, implying that HE is not as responsive to changes in GDP instantaneously as it is in the long run.

Using the VECM a long-run dynamic equilibrium relationship between GDP and HE was found. The results suggest that the long run elasticity of income is over, but close to 1. They also suggest that if there is a positive departure from the long run equilibrium, HE will, relatively slowly, fall in order to return to the equilibrium level. However, GDP does not respond to any departure from the long run equilibrium between GDP and HE.

These results support other dynamic equilibrium models presented in Section 2 of the literature review as well as the results from model 2. There is a low, but statistically significant short run response to changes in the GDP. This short run elasticity is much lower than that found in the first difference model, at 0.14 rather than around 0.40.

There are some limitations with the analysis, particularly concerning the data quality. Firstly, the log of relative health price reports large number of missing values (109/450 observations missing). The VECM can only be run on a fully balanced panel and therefore missing data need to be imputed: a linear assumption was applied to the data. Secondly, it is questionable that whether the variables selected in this study are good proxies. This is a common limitation of statistical analysis on macroeconomic health data. Thirdly, one might hypothesise that recent austerity measures, widely applied in high income countries since around 2008 would change the long run relationship between HE and its determinants. Unfortunately the lack of availability of very recent data makes the analysis of any hypothesised structural break due to recent austerity measures difficult. Many high income countries, especially the UK, have been attempting to reduce the proportion of their GDP spent on health. If they are successful, one would expect a change in the relationship between GDP and HE. However, given that the data are only available to 2012, there is no enough recent observations to explore this question. Finally, the analyses with the VECM pooled data from 18 countries. The pooled mean group estimation has been experimented which allows the short run coefficients and error variances to differ across countries. However, the reliability of the results is constrained by the short time period in the data, i.e. 25 years. There is a trade-off which needs to be made between choosing a model that takes into account the heterogeneity between countries and a model that uses pooled data but produces more efficient estimates. It should be noted that Hauck & Zhang (2016) explored an econometric model that allows for unobserved heterogeneity in the effects of common shocks in modelling HE growth in 34 OECD countries.

7. Conclusion

This study provides empirical evidence about the determinants of national HE in high income countries. The results from the Vector Error Correction model suggest that the key driver of HE at national level is GDP, with a 1% increase in GDP per head associated with a 1.10% increase in HE per head. The empirical results also show that the use of different econometric specifications has a significant impact on both establishing the determinants of HE and their magnitudes. A key contribution of this study is to account for the dynamic nature of the macro level data while exploring the determinants of the national HE. Future methodological research might consider exploring methods that address both the dynamic nature of the macro level time series data as well as cross-country heterogeneity.

List of Abbreviations

OECD: Organisation for Economic Cooperation and Development; **HE:** Health Expenditure; **GDP:** Gross Domestic Product; **VECM:** Vector Error Correction Model; **VAR:** Vector Autoregression; **CE:** Cointegration Equation; **ADF:** Augmented Dickey–Fuller

Acknowledgements

We would like to thank Professor Stephen Birch and participants of the UK Health Economists' Study Group meeting, Manchester, January 2016, for their helpful comments. We also would like to thank Professor Peter Burridge for his useful comments on the econometric analysis. This project is funded by a research grant from the Health Foundation (London, UK). The opinions expressed in this paper are those of the authors but not necessarily reflect the views of the funder.

References

- Baltagi, B. H., & Moscone, F. (2010). Health care expenditure and income in the OECD reconsidered: Evidence from panel data. *Economic Modelling*, 27(4), 804-811. <https://doi.org/10.1016/j.econmod.2009.12.001>
- Baumol, W. J. (1967). Macroeconomics of unbalanced growth: the anatomy of urban crisis. *The American economic review*, 57(3), 415-426.
- Blomqvist, A. G., & Carter, R. A. L. (1997). Is health care really a luxury? *Journal of Health Economics*, 16(2), 207-230. [https://doi.org/10.1016/S0167-6296\(96\)00534-6](https://doi.org/10.1016/S0167-6296(96)00534-6)
- Breyer, F., & Felder, S. (2006). Life expectancy and health care expenditures: a new calculation for Germany using the costs of dying. *Health Policy*, 75(2), 178-186. <https://doi.org/10.1016/j.healthpol.2005.03.011>
- Chakroun, M. (2009). Health care expenditure and GDP: An international panel smooth transition approach. *International Journal of Economics*, 4(1), 189-200.
- de la Maisonneuve, C., & Oliveira Martins, J. (2013). A Projection Method for Public Health and Long-Term Care Expenditures. Economics Department Working Papers, 1048. Paris: OECD.
- de la Maisonneuve, C., Moreno-Serra, R., Murtin, F., & Oliveira Martins, J. (2016). The drivers of public health spending: Integrating policies and institutions. OECD Economics Department Working Papers, No. 1283, OECD Publishing, Paris. <https://doi.org/10.2139/ssrn.2762212>
- Di Matteo, L., & Di Matteo, R. (1998). Evidence on the determinants of Canadian provincial government HEs: 1965-1991. *Journal of Health Economics*, 17(2), 211-228. [https://doi.org/10.1016/S0167-6296\(97\)00020-9](https://doi.org/10.1016/S0167-6296(97)00020-9)
- Drukker, D. M. (2003). Testing for serial correlation in linear panel-data models. *The Stata Journal*, 3(2), 168-177.
- Engle, R., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, Estimation and Testing. *Econometrica*, 55(2), 251-276. <https://doi.org/10.2307/1913236>
- Felder, S., Meier, M., & Schmitt, H. (2000). Health care expenditure in the last months of life. *Journal of Health Economics*, 19(5), 679-695. [https://doi.org/10.1016/S0167-6296\(00\)00039-4](https://doi.org/10.1016/S0167-6296(00)00039-4)
- Gerdtham, U. G. (1992). Pooling international HE data. *Health Economics*, 1(4), 217-231. <https://doi.org/10.1002/hec.4730010404>
- Gerdtham, U. G., & Jönsson, B. (2000). International comparisons of HE: theory, data and econometric analysis. In A. J. Culyer & J. P. Newhouse, eds. *Handbook of Health Economics*. Amsterdam: Elsevier, 11-53. [https://doi.org/10.1016/S1574-0064\(00\)80160-2](https://doi.org/10.1016/S1574-0064(00)80160-2)
- Gerdtham, U. G., Jönsson, B., MacFarlan, M., & Oxley, H. (1998). The determinants of HE in the OECD countries: a pooled data analysis. In P. Zweifel, ed. *Health, The Medical Profession, and Regulation*. Boston: Kluwer Academic Publishers, 113-134. https://doi.org/10.1007/978-1-4615-5681-7_6
- Gerdtham, U. G., Sjøgaard, J., Andersson, F., & Jönsson, B. (1992). An econometric analysis of HE: a cross-sectional study of the OECD countries, *Journal of Health Economics*, 11(1), 63-84. [https://doi.org/10.1016/0167-6296\(92\)90025-V](https://doi.org/10.1016/0167-6296(92)90025-V)
- Gerdtham, U. G., Sjøgaard, J., Jönsson, B., & Andersson, E. (1992). A pooled cross-section analysis of the HE of the OECD countries, in: P. Zweifel and H. Frech, eds., *Health Economics Worldwide*. Kluwer Academic Publishers: Dordrecht. https://doi.org/10.1007/978-94-011-2392-1_14
- Getzen, T. E. (2006). Aggregation and the Measurement of Health Care Costs. *Health Services Research*, 41(5), 1938-1954. <https://doi.org/10.1111/j.1475-6773.2006.00558.x>
- Granger, C. W. J. (1981). Some Properties of Time Series Data and Their Use in Econometric Model Specification,

- Journal of Econometrics*, 16(1), 121-130. [https://doi.org/10.1016/0304-4076\(81\)90079-8](https://doi.org/10.1016/0304-4076(81)90079-8)
- Granger, C. W. J. (1999). *Empirical Modelling in Economics*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511492327>
- Hansen, P., & King, A. (1996). The determinants of health care expenditure: a cointegration approach. *Journal of Health Economics*, 15(1), 127-137. [https://doi.org/10.1016/0167-6296\(95\)00017-8](https://doi.org/10.1016/0167-6296(95)00017-8)
- Hauck, K., & Zhang, X. H. (2016). Heterogeneity in the effect of common shocks on healthcare expenditure growth. *Health Economics*, 25(9), 1090-1103. <https://doi.org/10.1002/hec.3329>
- Hitiris, T., & Posnett, J. (1992). The determinants and effects of HE in developed countries. *Journal of Health Economics*, 11(2), 173-181. [https://doi.org/10.1016/0167-6296\(92\)90033-W](https://doi.org/10.1016/0167-6296(92)90033-W)
- Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica*, 59(6), 1551-1580. <https://doi.org/10.2307/2938278>
- Lago-Penas, S., Cantarero-Prieto, D., & Blazquez-Fernandez, C. (2013). On the relationship between GDP and health care expenditure: A new look. *Economic Modelling*, 32, 124-129. <https://doi.org/10.1016/j.econmod.2013.01.021>
- Leu, R. E. (1986). The public-private mix and international health care costs, in Culyer A.J. & Jönsson, B. eds., *Public and Private Health Services: Complementarities and Conflicts*. Oxford and New York: Blackwell, 41-63.
- Liu, D., Li, R., & Wang, Z. (2011). Testing for structural breaks in panel varying coefficient models: with an application to OECD HE. *Empirical Economics*, 40(1), 95-118. <https://doi.org/10.1007/s00181-010-0375-6>
- McCoskey, S. K., & Kao, C. (1998). A residual-based test of the null of cointegration in panel data. *Econometric Reviews*, 17(1), 57-84. <https://doi.org/10.1080/07474939808800403>
- McCoskey, S. K., & Selden, T. M. (1998). Health care expenditure and GDP: Panel data unit root test results, *Journal of Health Economics*, 17(3), 369-376. [https://doi.org/10.1016/S0167-6296\(97\)00040-4](https://doi.org/10.1016/S0167-6296(97)00040-4)
- Medeiros, J., & Schwierz, C. (2013). Estimating the drivers and projecting long-term public HE in the European Union: Baumol's "cost disease" revisited. European Commission, 507.
- Mehrara, M., Musai, M., & Amiri, H. (2010). The relationship between HE and GDP in OECD countries using PSTR. *European Journal of Economics, Finance and Administrative Sciences*, 24, 1450-2275.
- Newhouse, J. P. (1977). Medical-care expenditure: a cross-national survey. *Journal of Human Resources*, 12(1), 115-125. <https://doi.org/10.2307/145602>
- Parkin, D., McGuire, A., & Yule, B. (1987). Aggregate health care expenditures and national income: is health care a luxury good? *Journal of Health Economics*, 6(2), 109-127. [https://doi.org/10.1016/0167-6296\(87\)90002-6](https://doi.org/10.1016/0167-6296(87)90002-6)
- Roberts, J. (1999). Sensitivity of elasticity estimates for OECD health care spending: analysis of a dynamic heterogeneous data field. *Health Economics*, 8(5), 459-472. [https://doi.org/10.1002/\(SICI\)1099-1050\(199908\)8:5<459::AID-HEC454>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1099-1050(199908)8:5<459::AID-HEC454>3.0.CO;2-U)
- Sen, A. (2005). Is health care a luxury? New evidence from OECD data. *International Journal of Health Care Finance and Economics*, 5(2), 147-164. <https://doi.org/10.1007/s10754-005-1866-4>
- Seshamani, M., & Gray, A. (2004). Ageing and health-care expenditure: the red herring argument revisited. *Health Economics*, 13(4), 303-314. <https://doi.org/10.1002/hec.826>
- Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Appendix A: VEC residual portmanteau tests for autocorrelations in VECM 1

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	Df
1	0.560139	NA*	0.561557	NA*	NA*
2	1.513069	NA*	1.519324	NA*	NA*
3	8.995525	0.2530	9.058899	0.2485	7
4	16.74397	0.1157	16.88641	0.1113	11
5	20.94269	0.1387	21.13882	0.1325	15
6	22.48464	0.2608	22.70450	0.2506	19
7	24.17165	0.3944	24.42186	0.3808	23
8	26.23046	0.5058	26.52312	0.4897	27
9	35.64412	0.2590	36.15570	0.2403	31
10	42.21125	0.1874	42.89296	0.1688	35
11	45.74577	0.2124	46.52848	0.1901	39
12	47.84562	0.2825	48.69394	0.2546	43
13	51.02558	0.3184	51.98184	0.2860	47
14	55.97575	0.2935	57.11343	0.2584	51
15	60.31003	0.2898	61.61835	0.2513	55
16	71.42972	0.1286	73.20623	0.1010	59
17	74.92969	0.1444	76.86319	0.1126	63
18	82.04618	0.1020	84.31857	0.0749	67
19	88.22841	0.0811	90.81237	0.0566	71
20	94.76888	0.0613	97.70074	0.0403	75
21	97.13941	0.0811	100.2040	0.0539	79

Appendix B: VEC Residual portmanteau tests for autocorrelations in VECM 2

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	Df
1	0.553394	NA*	0.554795	NA*	NA*
2	2.971187	NA*	2.984861	NA*	NA*
3	17.73194	0.2193	17.85829	0.2133	14
4	27.21503	0.2469	27.43815	0.2378	23
5	34.09356	0.3672	34.40464	0.3533	32
6	46.52221	0.2555	47.02450	0.2394	41
7	60.98906	0.1372	61.75168	0.1231	50
8	65.64742	0.2574	66.50609	0.2344	59

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.