Health Care Expenditure and Income in the OECD Reconsidered: Evidence from Panel Data

Badi H. Baltagi  Francesco Moscone
Syracuse University*  University of Leicester†

February 20, 2009

Abstract

This paper reconsiders the long-run economic relationship between health care expenditure and income using a panel of 20 OECD countries observed over the period 1971-2004. In particular, the paper studies the non-stationarity and cointegration properties between health care spending and income. This is done in a panel data context controlling for both cross-section dependence and unobserved heterogeneity. Cross-section dependence is modelled through a common factor model and through spatial dependence. Heterogeneity is handled through fixed effects in a panel homogeneous model and through a panel heterogeneous model. Our findings suggest that health care is a necessity rather than a luxury, with an elasticity much smaller than that estimated in previous studies.

JEL Classification: C31, C33, H51

Keywords: Health expenditure, income elasticity, cross section dependence, heterogeneous panels, factor models.

1 Introduction

Health care expenditure in the OECD varies substantially over time and across countries. From 1970 to 2004, per capita health expenditure has increased markedly in the OECD with an annual average rate of 11.5 per cent. Such temporal dynamic has been characterized by large differences across countries, leading to marked geographical heterogeneity in the level of spending. For example, a snapshot in 2004 shows that the US, with an average of $6,037\(^1\), has the highest amount of health expenditure, followed by Switzerland ($4,045), Norway ($4,103), and Germany ($3,169). On the other hand, countries that devote less resources to health care are Turkey and Mexico, with an average per

\(^*\)Department of Economics and Center for Policy Research. E-mail: bbaltagi@maxwell.syr.edu.

\(^†\)Department of Economics. E-mail: fmoscone@le.ac.uk.

\(^\dagger\)The levels of total health expenditure per capita shown in U.S. dollars have been adjusted for purchasing power parity (PPP).
capita expenditure of $562 and $655, respectively. As a share of Gross Domestic Product (GDP), health care spending in the OECD has almost doubled over this period, increasing from 4.9 per cent in 1970 to 8.8 per cent in 2004. However, there is a substantial heterogeneity across these OECD countries. In fact, while several countries continued to experience an increase in their share in the 80s and 90s, others have experienced modest declines, possibly associated with reforms aimed at limiting the percentage rise in health care spending as a proportion of GDP. Over time, the shares of health care spending as a percentage of GDP, ranged between 2.5 and 7.0 per cent in the 70s, compared to 5.5 and 15.2 per cent in 2004.

Since the work by Kleiman (1974) and Newhouse (1977), income has been identified as the most important factor explaining differences across countries in the level and growth of health care expenditure. Therefore, earlier research focused on measuring the size of the income elasticity of health care, and on its policy implications for the financing and distribution of health care resources. Advocates of health care being a luxury good, argued that it is a commodity much like any other and is best left to market forces. On the other hand, advocates of health care being a necessity, often support the idea of more government intervention in the health care sector (see Culyer, 1988; and Di Matteo, 2003). We will review the empirical literature for the OECD countries in the next section. Several empirical studies pointed to the possible non-stationarity of health care spending and income, which in turn cast doubt on prior inference on income elasticity obtained from spurious regressions. This literature focused on studying the time series properties of health expenditure and income, and on assessing whether there exists a long-run relationship between them.

A number of non-income determinants of health care spending have been identified in the literature. For example, the age structure of the population has been traditionally flagged as an important factor in explaining variations of health care expenditure across countries (Leu, 1986; Culyer, 1988). Indicators such as the share of young (e.g., under 15 years) and old people (e.g., above 65 or 75 years) over the active or total population have been included in regression models explaining per-capita health spending. Nevertheless, little evidence exists on a significant effect of these variables (Leu, 1986; Hitiris and Posnett, 1992; Di Matteo and Di Matteo, 1998; Grossman, 1972). Another determinant of health expenditure is the extent to which health care expenditure is financed by the government, though only few empirical studies support its effect on health care spending (Leu, 1986; Culyer 1988; Hitiris and Posnett, 1992).

Microeconomic theory emphasize the role of real prices for health care services in determining the demand for health care (Grossman, 1972). A positive effect of relative prices on health spending would support the so-called Baumol (1967) cost disease theory that productivity in the health sector is low relative to other sectors. Hence, prices for health services will rise relative to other prices because wages in low productivity sectors must keep up with wages in high productivity sectors. However, there is no empirical consensus on the effect of real prices on health care spending. See (Hartwig, 2008; Okunade et al., 2004), who report a positive and statistically significant effect, and (Gerdtham
et al., 1992; Murthy and Ukpolo, 1994) who report an insignificant effect. Yet, there are skeptics who do not recommend the use of price indexes in health care, especially across countries that provide health care at no cost or at very low cost, see Berndt et al. (2000). In fact, Hartwig (2008, p.6) argues that "...we have to recognize that medical care price indices can probably not be relied on as deflators or explanatory variables." Given the paucity of data on price across the OECD, the diverse national schemes of price regulation, and the problems with measuring quality of health care in obtaining this medical price index, we decided not use this variable in our empirical analysis (see Section 5).

Since the work by Newhouse (1992), technological progress has been seen as an important driver of health care expenditure. However, very few studies have attempted to study the relationship between technological progress and health care expenditure due to the difficulty of finding an appropriate proxy for changes in medical care technology. A number of proxies have been considered in the literature, such as the surgical procedures and the number of specific medical equipment (Baker and Wheeler 2000; Weil, 1995); the R&D spending specific to health care (Okunade and Murthy, 2002); life expectancy and infant mortality (Dregen and Reimers, 2005). Some other papers have proxied the effect of technical change by adding a time index (Gerdtham and Lothgren, 2000), or time-specific intercepts (Di Matteo, 2004) in the regression specification.

To summarize, while income has been recognized as an important determinant of health care spending, there is still no consensus on which other factors may be associated with the remaining largely unexplained variation in per capita health expenditure. Some attribute this failure to identify other non-income determinants to the limited availability of health care data at the macro level, others even blame the weakness of the econometric methods used, or the informal economic theory used to model per capita medical care expenditure (Wilson, 1999).

This paper studies the long-run economic relationship between health care expenditure and income in the OECD countries, ultimately assessing whether health care is a luxury or a necessity. Using a panel of 20 OECD countries followed over the period 1971-2004, we investigate the non-stationarity and cointegration properties between health care spending and income. The dynamics of health expenditure and income and their relationship are investigated by estimating a heterogenous panel model with cross sectionally correlated errors. Initially a factor structure is included in the econometric specification with the intent to synthesize the effects of shocks that may hit health spending and that are not directly measurable by the econometrician, such as advances in medical care technology, policy shifts, new diseases, and shifts in preferences and expectations by users of health services. The factor structure can capture any contemporaneous correlation that arises from the common response of countries to such unanticipated events. We then model cross section dependence by assuming that the regression errors follow a spatial autoregressive process. Indeed, consumption of health care resources of a single country may be related to unobservable general population characteristics of neighbouring countries. A very recent strand of literature has recognized that cross section dependence is an
important characteristic of health data, and has tried to incorporate it in their models (Jewell et al., 2003; Freeman, 2003; Carrion-i-Silvestre 2005; Wang and Rettenmaier, 2006; Chou, 2007). We also check the robustness of our results by including in the regression specification variables recognized by the literature to play an important role, such as government expenditure on health, and the age structure. The aim is to assess income elasticity more accurately, controlling for various alternative forms of cross section dependence, as well as non income determinants of health expenditure.

The plan of the paper is as follows. Section 2 summarizes the prior empirical results on this topic. Section 3 introduces the econometric methods adopted. Section 4 presents the data. Section 5 summarizes our empirical results, and points to some of the limitations of our study. While, Section 6 gives our concluding remarks.

2 Income elasticity in the OECD

This short review summarizes some of the existing studies that have used panel data sets to measure the relationship between health care spending and income in the OECD.

We start with Gerdtham et al. (1992) who estimated a regression for health care spending as a function of GDP and a number of other variables, including institutional and socio-demographic factors. Using data on 20 OECD countries over the period 1960 and 1987, they estimated an income elasticity that is larger than one, thus finding that health care is a luxury good. This finding is in line with previous results based on a single cross section (e.g., Kleiman, 1974; Newhouse, 1977; and Leu (1986). Using the same data, Hitiris and Posnett (1992) estimated a regression model for health care expenditure and income, controlling for unobserved heterogeneity by adding country-specific effects. They measured an income elasticity close to one, thus questioning the luxury attribute of health care raised by Gerdtham et al. (1992). As observed by Hansen and King (1996), one limitation of the above studies is that they have ignored the possibility of non-stationarity in health data and income. Using the same data set as Gerdtham et al. (1992), they computed Dickey Fuller statistics for health care spending, GDP, and residuals from a regression of GDP on health expenditure, for each country separately. While detecting non-stationarity for health care spending and GDP for the majority of OECD countries, they did not find evidence of cointegration among the variables. Using data on 24 OECD countries observed over the period 1960 to 1991, Blomqvist and Carter (1997) computed the Phillips and Perron t-ratios for health care spending and GDP and for regression residuals. The authors conclude that their results cast doubt on pooling and upon the notion of an elasticity larger than one.

McCoskey and Selden (1998) revisited the work by Hansen and King (1996), applying for the first time non-stationarity tests that exploit the panel nature of the OECD data. The low power of country-by-country tests employed in previ-
ous studies is one of the major motivations for the use of panel unit root tests. Specifically, McCoskey and Selden (1998) computed the tests by Im, Pesaran, and Shin (2003), and rejected the joint hypotheses of unit root in all countries for both health care spending and income, though observing that results are sensitive to the inclusion of a time trend in the augmented Dickey-Fuller equation. Using data on 21 countries followed over the years 1960-1997, Gerdtham and Lothgren (2000, 2002) computed the Im, Pesaran, and Shin (2003) test and the panel version of the Kwiatkowski and Phillips (1992) test, with linear trends, and concluded in favour of non-stationarity and cointegration between health care spending and GDP (see also Okunade and Karakus (2001)). Similar results have been obtained by Dregen and Reimers (2005), for the years 1975-2001. These authors estimated the relationship between health care expenditure and GDP controlling for non-income determinants and a proxy of technical progress, concluding that health care expenditure is not a luxury good.

The above results have been criticized by Jewell et al. (2003), who emphasized the importance of controlling for structural breaks and cross section dependence. Further, to mitigate cross section dependence they included time-specific effects common to all countries. Hence, the authors detected one or two breaks for most OECD countries, and found that health care spending and GDP are stationary once these breaks are taken into account. Similar results have been obtained by Carrion-i-Silvestre (2005). He based his inference on the bootstrap distribution of stationarity tests, in order to render the analysis robust to the presence of cross section dependence. Hartwig (2008) reviews this literature and concludes: "Unfortunately, given that the available time series are rather short, which lowers the power of the tests, and that the number of competing tests is huge (and growing), some uncertainty is likely to remain with respect to the properties of the time series analyzed in this field of research."

In the next section, we review a number of methods to study the long-run relation between health care expenditure and income. Our regression specification incorporates global shocks, spatial spillovers and unobserved heterogeneity across countries.

3 Methods

3.1 The econometric model

Blomqvist and Carter (1997, p.226) argued that their most important finding is that ". pooling restrictions are of very doubtful validity. Even allowing for different country intercepts, there is considerable evidence, albeit somewhat questionable, against the hypotheses of equal income elasticities and a common trend reflecting technological progress." In response to this statement, we consider the following linear heterogeneous panel regression model:

$$h_{it} = \alpha_i + d_t + \beta_i x_{it} + u_{it}, \quad i = 1, \ldots, N; t = 1, \ldots, T,$$

(1)
where $h_{it}$ indicates the logarithm of real per-capita health care expenditure in the $i$th country at time $t$, $x_{it}$ is a $k \times 1$ set of regressors including income, public expenditure on health, and the age structure; $\alpha_i$ is a country-specific intercept, $d_t$ is a time dummy, and $u_{it}$ is the error term. In this paper we consider two alternative ways of incorporating cross section dependence in equation (1). The first model assumes that the errors have the following multifactor structure

$$u_{it} = \gamma' f_t + \varepsilon_{it}, \quad (2)$$

in which $f_t$ is the $m \times 1$ vector of unobserved common effects and $\varepsilon_{it}$ is a country-specific error assumed to be independently distributed. From (2), correlation arises because the responses to common external forces or perturbations is similar, though not identical, across countries. Notice that common factors induce a correlation between pairs of statistical units that does not depend on how close they are in the geographical space. In model (1), we allow $x_{it}$ to be correlated with the unobserved effects $f_t$. Therefore, common factors can impact health expenditure not only directly via the factor structure (2), but also indirectly by affecting the regressors.

The second model we consider for the error term $u_{it}$ is the following Spatial Autoregressive (SAR) process:

$$u_{it} = \delta \hat{u}_{it} + \varepsilon_{it}, \quad (3)$$

where

$$\hat{u}_{it} = \sum_{j=1}^{N} s_{ij} u_{jt}, \quad (4)$$

with $s_{ij}$ being the generic $(i,j)$th element of a $N \times N$ spatial weights matrix $S$ (Anselin, 1988). In our empirical work we will adopt weights based on the inverse of the distance expressed in kilometers across countries.

A recent strand of literature in health economics has suggested a number of reasons why one should expect geographical concentration of unobservable risk factors of health care spending (Revelli, 2006, Moscone and Knapp, 2005, and Moscone et al. 2007). Further, there exists a growing medical literature that has detected spatial geographical concentration of various health conditions such as mortality (Lorant et al., 2001), cancer mortality (Thouez et al., 1997), child leukemia (Alexander, 1993), childhood cancer (Gatrell and Whitelegg, 1993), and asthma (Hsiao, 2000).

The estimation and testing approach to equation (1) with multifactor errors (2) is based on the Common Correlated Effects (CCE) method advanced by Pesaran (2006). Accordingly, we estimate the slope parameters $\beta_i$ applying standard panel techniques to the following equation

$$h_{it} = \alpha_i + d_t + \beta_i' x_{it} + \gamma_0 z_t + e_{it}, \quad (5)$$

where $z_t = (\bar{h}_t, \bar{x}_t)'$, with $\bar{h}_t$ and $\bar{x}_t$ being the cross section averages of the dependent variable and regressors respectively. In our analysis we will compute
CCE Pooled (CCEP) estimator for the average of the slope coefficients (Pesaran, 2006). Heterogeneity is captured by the individual specific fixed effects, $\alpha_i$, the time dummies, $d_i$, and the loadings, $g_i$. For comparison purposes, we will also compute the Fixed Effects (FE) estimator with period dummies. We observe that the CCEP and the FE frameworks differ in that the latter assumes that $g_i$ is zero, and the $\beta_i’s$ are the same. The estimation and testing strategy of equation (1) with fixed effects and spatially correlated errors (3) is based on Maximum Likelihood (ML) estimation techniques.

3.2 Testing for unit roots

Consider the $p^{th}$ order augmented Dickey Fuller regression

$$\Delta q_{it} = a_i + b_i q_{i,t-1} + c_i t + \sum_{j=1}^{p} d_{ij} \Delta q_{i,t-j} + u_{it},$$

where $q_{it}$ is either the logarithm of real per-capita health spending, the logarithm of the $j$th regressor $x_{j, it}$, or regression residuals from equation (1). $u_{it}$ are errors that we assume to have a single factor structure, where the idiosyncratic component follows a spatial autoregressive process as in (3). When testing for unit roots, the null hypothesis is

$$H_0 : b_i = 0, i = 1, \ldots, N,$$

against the alternative that

$$H_1 : b_i < 0, i = 1, \ldots, N_1; b_i = 0, i = N_1 + 1, \ldots, N,$$

where $N_1$ is such that $N_1/N$ is nonzero and tends to a fixed constant as $N$ goes to infinity. Pesaran (2007) proposes to test (7) against (8) by computing the simple average of the $t$-ratios of the OLS estimates of $b_i$ in equation (10), namely,

$$CIPS = \frac{1}{N} \sum_{i=1}^{N} \hat{t}_i,$$

where $\hat{t}_i$ is the OLS $t$-ratio of $b_i$ in the following Dickey Fuller (CADF) regression augmented with the cross section averages $\bar{q}_{i-1}$ and $\Delta \bar{q}_{i-j}$, for $j = 0, \ldots, p$

$$\Delta q_{it} = a_i + b_i q_{i,t-1} + c_i t + \sum_{j=1}^{p} d_{ij} \Delta q_{i,t-j} + g_i \bar{z}_t + e_{it}.$$

where $\bar{z}_t = (\bar{q}_{i-1}, \Delta \bar{q}_t, \Delta \bar{q}_{i-1}, \ldots, \Delta \bar{q}_{i-p})'$. The critical values for the CIPS tests are given in Tables 2(a)-2(c) in (Pesaran, 2007).

The CIPS test has been designed for testing the unit root hypothesis when the variable under study has a factor structure. However, Monte Carlo experiments have indicated that this test is robust also to the presence of other sources of cross section dependence such as the SAR process (3) (see Baltagi et al., 2007).
3.3 Cross section dependence tests

We now briefly review some statistics of cross section dependence that we use in our empirical work. A statistic that captures the overall amount of cross section dependence in the data, at a descriptive level, is the following average pairwise correlation coefficient

\[ \bar{p} = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \rho_{ij}, \]  

where \( \rho_{ij} \) is given by

\[ \rho_{ij} = \frac{\sum_{t=1}^{T} q_{it} q_{jt}}{\sqrt{\left(\sum_{t=1}^{T} q_{it}^2\right) \left(\sum_{t=1}^{T} q_{jt}^2\right)}}. \]

and \( q_{it} \) are regression residuals from equations (1) or (6).

We also consider a diagnostic test of cross section independence based on the above pairwise correlation coefficients. In particular, we consider the CD\(_{LM}\) test based on the Lagrange Multiplier statistic (Frees, 1995)

\[ CD_{LM} = \sqrt{\frac{1}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left( T \rho_{ij}^2 - 1 \right) \left( T \rho_{ij}^2 - 1 \right) \left( T \rho_{ij}^2 - 1 \right)}. \]  

Under the null hypothesis of no cross section dependence, the \( CD_{LM} \) tends to a \( N(0, 1) \) with \( T \to \infty \) and then \( N \to \infty \).

In our empirical study we also test for spatial correlation. In particular, we compute the following Moran’s \( I \) test statistic (Kelejian and Prucha, 2001)

\[ I = \frac{\sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{j=1}^{N} s_{ij} \hat{u}_{it} \hat{u}_{jt}}{\left( T \sum_{i=1}^{N} \sum_{j=1}^{N} (s_{ij} + s_{ji})^2 \sigma_i^2 \sigma_j^2 \right)^{1/2}}, \]

where \( \sigma_i^2 = \hat{\sigma}_i^2 = \frac{1}{T} \sum_{t=1}^{T} \hat{u}_{it}^2 \), and \( s_{ij}, i, j = 1, ..., N \), are the spatial weights. The Moran’s \( I \) is asymptotically normally distributed as \( N \) goes to infinity, for fixed \( T \). Spatial statistics such as the Moran’s \( I \) differ from the CD statistic (12) since they exploit information on the spatial ordering of the data, giving more importance to countries that are closer to each other.

4 Data description

Our analysis uses annual data on 20 OECD countries from 1971 to 2004 (\( T = 34 \)), gathered from the OECD Health Data Set 2007. We collected information
on per-capita total health care expenditure and per-capita income estimated in GDP purchasing power parity, and expressed in US Dollars. We also gathered data for the following variables that have been identified by the literature as having a role in determining health care expenditure: public expenditure on health care computed as government expenditure over total health care expenditure; the dependency rates for old and young people, defined as the population aged 65 and over divided by the population aged 15-64, and the population aged 0-14 divided by the population aged 15-64, respectively. All variables are expressed in natural logarithm. As shown in Table 1, the sample of 20 countries and 34 years decreases by few units when public expenditure on health care and the age structure are added to the regression.

After a preliminary exploratory data analysis, our empirical study is structured as follows: we first check whether our variables are non-stationary; we then estimate the income elasticity controlling for a set of regressors and for unobserved common factors; finally, we test whether our variables form a cointegrating set and therefore if they are linked in the long-run.

5 Empirical analysis

Table 2 reports the average correlation coefficient and the $C_{LM}$ tests for the first differences of the logarithm of all the variables, regressed on a country-specific intercept. Results indicate the presence of cross section correlation between pairs of countries for all variables. The Moran $I$ statistic suggests the presence of geographical concentration of health care spending and its determinants. Therefore, these tests show that pairs of countries in our data set are correlated to each other for all variables, and that in some cases, these display a spatial pattern. These two sources of correlation will be taken into account when studying the time series properties of the variables as well as when estimating the health care spending equation. Failure to do so may lead to misleading inference, particularly if the source of cross section dependence is correlated with the regressors (Andrews, 2005).

Table 3 reports the results of first generation panel unit root tests which do not account for cross-country dependence. The first column of Table 3 reports the Im, Pesaran and Shin (2003) (IPS) $W_{1\text{bar}}$ statistic for the logarithm of our variables when the ADF regression has an intercept only. Interestingly, all the variables considered reject the panel unit root hypothesis with the exception of the public and old people variables. The second column of Table 3 reports the IPS statistic for the logarithm of our variables when the ADF regression has an intercept and a linear time trend. In this case, only income and old people do not reject the null of panel unit root. The third column of Table 3 shows results for the Breitung (2000) t-statistic for the intercept and trend case.2

\[ \text{Note that the null hypothesis for the IPS and Breitung tests are the same. However, the alternatives are not. The IPS statistic has an alternative described in (7) where some fraction of the countries are stationary while others are not. Breitung (2000) shows that his test exhibits better power than the IPS test in the presence of country specific trends.} \]
all cases, the lag order $p$, was selected using the SIC criterion. Results from the Breitung t-statistic do not reject the null of panel unit root for all variables.\(^3\)

Table 4 shows the CIPS statistics for the logarithm of our variables. We report these results for lag orders $p = 0; 1; 2; 3$. As we can see from the table, most of the variables are non-stationary when adding an intercept only, and when including an intercept and a linear trend. On the other hand, they are stationary when the unit root tests are applied to the first differences of these variables. Given the sizeable amount of cross country dependence detected by the CD tests reported in Table 2, we believe that the CIPS unit roots tests give more reliable inference than those from first generation panel unit root tests, and we conclude that the variables under study are non-stationary.

In order to check the sensitivity of our panel unit root results, we run these tests again but now removing one country at a time from the sample. Table 5 report the CIPS statistics for the variables health care expenditure, income, and public expenditure in the intercept and trend case. By and large, these results show that if we drop any country from the analysis, the results of the CIPS tests are similar to those reported in Table 4. The variable that exhibits the most sensitivity is income.

Table 6 shows results from FE, Spatial MLE, and CCEP estimation when income is the only variable included in the regression (Panel A), as well as when public expenditure and dependency rates are added (Panel B). If we focus on the FE estimates (column (I) Panels A, B), the income elasticity is smaller than one, suggesting the necessity nature of health care.\(^4\) The variables public expenditure and dependency rate for old people are not significant in both regressions, thus confirming similar findings in previous studies on the OECD countries (Hitiris and Posnett, 1992). On the other hand, the variable dependency rate for young people has a significant and positive influence on health care expenditure for the regression reported in Panel B.

For the MLE accounting for spatial correlation (column II Panels A, B), the parameter estimates for the income elasticity are close to their FE non spatial counterpart. However, the estimates of the other control variables are different, with the old people variable becoming significant. Interestingly, once one controls for the period effects, the estimated spatial coefficients are negative ranging from -0.41 to -0.46. These may be capturing the indirect effects of unobservable neighbouring variables such as environmental risks which are difficult to measure on health care spending.

The CCEP estimates (column III Panels A, B) give the lowest estimates of the income elasticity, especially when we control for non-income variables. These results corroborate the hypothesis that health care is a necessity good. Given

\(^3\)We have also computed other first generation panel unit root tests. The Levin Lin and Chu (2002), and the Maddala and Wu (1999) Fisher type tests. All of which reject the null of a panel unit root for health care expenditures and income in the case of individual effects and a linear trend. The Hadri (2000) test on the other hand, which reverses the null and the alternative, rejects the null of stationarity for all variables.

\(^4\)One interesting point to observe is that if we omit the time dummies, the FE estimate of the income elasticity becomes larger than one. The time dummies however are significant and we chose to include them.
the sizeable amount of correlation across countries detected in our exploratory data analysis, we believe that the CCEP approach, incorporating the effect of unobservable common factors, is more appropriate for estimating equation (1). Table 6 also reports the statistics \( CD_{LM} \), and Moran’s \( I \) applied to the residuals of the CCEP, spatial MLE and FE regressions. These indicate the presence of a general form of cross section dependence.

Table 7 checks the sensitivity of our income elasticity estimates from the health care expenditure equation by removing one country at a time from the sample. This is done controlling for non income factors. By and large, this income elasticity is robust and lower than one, regardless of which country was dropped from the sample. The only exception is when we remove Ireland, there the income elasticity becomes larger than one for the spatial MLE and FE estimates but not for the CCEP estimate.

Table 8 performs the CIPS panel unit roots tests on the residuals from the estimated equations reported in 6. The CCEP residuals from the first and second regression are stationary for \( p = 0, 1, 2 \), suggesting the existence of a long-run economic relationship between health expenditure and income whether one controls for public expenditure and dependency rates, or not. In contrast, for the FE regressions, we do not reject the unit root hypothesis in the residuals, for \( p = 0, 1, 2, 3 \) whether we control for public expenditure and dependency rates, or not. Hence, there is a marked difference between the CCEP and FE (non spatial and spatial).

Table 9 reports the error correction models attached to the CCEP estimation, which has shown a cointegration relation between the variables. The coefficient attached to \( h_{i,t-1} - \hat{\beta}_0 x_{i,t-1} \) measures the speed of adjustment of health care spending to a deviation from the long-run equilibrium relation between expenditure and its determinants. As expected, this coefficient is negative and significant in both regressions. Notice that short run changes in public expenditure, and dependency rates do not seem to have significant effects on health expenditure.

6 Concluding remarks

This paper investigated the long-run economic relationship between health care expenditure and income in the OECD countries. Using a panel of 20 OECD countries followed over 34 years, we have studied the non-stationarity and cointegration properties of health care expenditure and GDP, ultimately measuring income elasticity of health care. This paper contributes to the literature adopting tests that allow one to explicitly to control for cross-country dependence and unobserved heterogeneity. Our analysis indicates that health care expenditure and most of its determinants are non-stationary, and that they are linked in the long-run. Our results show that health care is a necessity rather than a luxury, with an elasticity much smaller than that estimated in other OECD studies. As for non-income determinants, our analysis indicates a role for the percentage of young people in explaining health expenditure variations.
References


Table 1: Description of the variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>$N$</th>
<th>$T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_{it}$</td>
<td>Per-capita health expenditure</td>
<td>20</td>
<td>34</td>
</tr>
<tr>
<td>$y_{it}$</td>
<td>Per-capita GDP</td>
<td>20</td>
<td>34</td>
</tr>
<tr>
<td>$Publ_{it}$</td>
<td>% government expenditure</td>
<td>17</td>
<td>34</td>
</tr>
<tr>
<td>$Old_{it}$</td>
<td>Dependency rate, old people</td>
<td>19</td>
<td>34</td>
</tr>
<tr>
<td>$Young_{it}$</td>
<td>Dependency rate, young people</td>
<td>19</td>
<td>34</td>
</tr>
</tbody>
</table>

Table 2: Cross section dependence in the first differences of the logarithm of all variables

<table>
<thead>
<tr>
<th></th>
<th>$\bar{p}$</th>
<th>$CD_{LM}$</th>
<th>$I$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta h_{it}$</td>
<td>0.483</td>
<td>73.843*</td>
<td>24.490*</td>
</tr>
<tr>
<td>$\Delta y_{it}$</td>
<td>0.552</td>
<td>97.273*</td>
<td>29.432*</td>
</tr>
<tr>
<td>$\Delta Publ_{it}$</td>
<td>0.092</td>
<td>2.854*</td>
<td>16.748*</td>
</tr>
<tr>
<td>$\Delta Old_{it}$</td>
<td>0.184</td>
<td>28.826*</td>
<td>11.270*</td>
</tr>
<tr>
<td>$\Delta Young_{it}$</td>
<td>0.296</td>
<td>104.399*</td>
<td>20.062*</td>
</tr>
</tbody>
</table>

Notes: $\bar{p}$, $CD_{LM}$ and $I$ are computed as in (11), (12), and (13), respectively.
All variables (expressed in first differences) have been regressed on a country-specific intercept.
"**" indicates that the coefficient is significant at the 5% level.

Table 3: Im Pesaran and Shin and Breitung panel unit root tests

<table>
<thead>
<tr>
<th></th>
<th>IPS $W_{thor}$</th>
<th>Breitung t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept only</td>
<td>Intercept and trend</td>
</tr>
<tr>
<td>$h_{it}$</td>
<td>-5.456*</td>
<td>-4.017*</td>
</tr>
<tr>
<td>$y_{it}$</td>
<td>-6.793*</td>
<td>-0.079</td>
</tr>
<tr>
<td>$Publ_{it}$</td>
<td>-1.599</td>
<td>-1.929*</td>
</tr>
<tr>
<td>$Young_{it}$</td>
<td>-6.925*</td>
<td>-7.822*</td>
</tr>
<tr>
<td>$Old_{it}$</td>
<td>0.086</td>
<td>-1.276</td>
</tr>
</tbody>
</table>

Notes: "**" indicates that the test is significant at the 5% level.
Table 4: CIPS panel unit roots tests

<table>
<thead>
<tr>
<th></th>
<th>number of lags</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>With an intercept only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h_{it}$</td>
<td>-2.190</td>
<td>-1.927</td>
<td>-1.986</td>
<td>-2.211*</td>
</tr>
<tr>
<td>$y_{it}$</td>
<td>-1.455</td>
<td>-1.791</td>
<td>-1.459</td>
<td>-1.616</td>
</tr>
<tr>
<td>$Publ_{it}$</td>
<td>-1.752</td>
<td>-1.850</td>
<td>-1.925</td>
<td>-1.791</td>
</tr>
<tr>
<td>$Young_{it}$</td>
<td>-1.760</td>
<td>-2.442*</td>
<td>-1.912</td>
<td>-2.098</td>
</tr>
<tr>
<td>$Old_{it}$</td>
<td>-1.367</td>
<td>-1.802</td>
<td>-2.005</td>
<td>-1.919</td>
</tr>
<tr>
<td>$\Delta h_{it}$</td>
<td>-4.822*</td>
<td>-3.774*</td>
<td>-2.942*</td>
<td>-2.510*</td>
</tr>
<tr>
<td>$\Delta y_{it}$</td>
<td>-4.733*</td>
<td>-3.984*</td>
<td>-2.889*</td>
<td>-2.504*</td>
</tr>
<tr>
<td>$\Delta Publ_{it}$</td>
<td>-5.334*</td>
<td>-3.646*</td>
<td>-2.950*</td>
<td>-2.796*</td>
</tr>
<tr>
<td>$\Delta Young_{it}$</td>
<td>-2.718*</td>
<td>-2.216*</td>
<td>-1.982</td>
<td>-2.048</td>
</tr>
<tr>
<td>$\Delta Old_{it}$</td>
<td>-2.616*</td>
<td>-1.997</td>
<td>-1.697</td>
<td>-1.706</td>
</tr>
<tr>
<td></td>
<td>With an intercept and trend</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h_{it}$</td>
<td>-2.750*</td>
<td>-2.516</td>
<td>-2.437</td>
<td>-2.399</td>
</tr>
<tr>
<td>$y_{it}$</td>
<td>-2.375</td>
<td>-2.918*</td>
<td>-2.365</td>
<td>-2.510</td>
</tr>
<tr>
<td>$Publ_{it}$</td>
<td>2.295</td>
<td>-2.134</td>
<td>-2.150</td>
<td>-2.383</td>
</tr>
<tr>
<td>$Young_{it}$</td>
<td>-1.824</td>
<td>-2.285</td>
<td>-2.110</td>
<td>-2.199</td>
</tr>
<tr>
<td>$Old_{it}$</td>
<td>-1.070</td>
<td>-1.767</td>
<td>-1.968</td>
<td>-2.003</td>
</tr>
</tbody>
</table>

Notes: "***" indicates that the test is significant at the 5% level.
<table>
<thead>
<tr>
<th>Country</th>
<th>Per-capita health expenditure</th>
<th>Per-capita GDP</th>
<th>% government expenditure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of lags</td>
<td>Number of lags</td>
<td>Number of lags</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Sweden</td>
<td>-2.813* -2.561 -2.493 -2.448</td>
<td>-2.348 -2.871* -2.400 -2.514</td>
<td>- - - -</td>
</tr>
<tr>
<td>Switzerland</td>
<td>-2.762* -2.546 -2.519 -2.465</td>
<td>-2.455 -3.033* -2.445 -2.591</td>
<td>- - - -</td>
</tr>
</tbody>
</table>

Notes: we only report the intercept and the trend case.

*** indicates that the test is significant at the 5% level.
Table 6: Estimation results: the determinants of health expenditure

<table>
<thead>
<tr>
<th>Panel A</th>
<th>(I): FE</th>
<th>(II) Spatial MLE</th>
<th>(III) CCEP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ceff.</td>
<td>Std.err.</td>
<td>Ceff.</td>
</tr>
<tr>
<td>$y_{it}$</td>
<td>0.899*</td>
<td>0.044</td>
<td>0.896*</td>
</tr>
<tr>
<td>$\delta$</td>
<td>-0.350*</td>
<td>0.110</td>
<td></td>
</tr>
<tr>
<td>CD statistics$^\text{a})$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$CD_{LM}$</td>
<td>80.877*</td>
<td></td>
<td>80.696*</td>
</tr>
<tr>
<td>$I$</td>
<td>-0.996</td>
<td></td>
<td>-1.059</td>
</tr>
<tr>
<td>Panel B</td>
<td>(I): FE</td>
<td>(III) Spatial MLE</td>
<td>(III): CCEP</td>
</tr>
<tr>
<td></td>
<td>Ceff.</td>
<td>Std.err.</td>
<td>Ceff.</td>
</tr>
<tr>
<td>$y_{it}$</td>
<td>0.867*</td>
<td>0.047</td>
<td>0.854*</td>
</tr>
<tr>
<td>$Publ_{it}$</td>
<td>0.083</td>
<td>0.090</td>
<td>0.034</td>
</tr>
<tr>
<td>$Old_{it}$</td>
<td>0.094</td>
<td>0.051</td>
<td>0.133*</td>
</tr>
<tr>
<td>$Young_{it}$</td>
<td>0.200*</td>
<td>0.053</td>
<td>0.183*</td>
</tr>
<tr>
<td>$\delta$</td>
<td>-0.417*</td>
<td>0.116</td>
<td></td>
</tr>
<tr>
<td>CD statistics$^\text{a})$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$CD_{LM}$</td>
<td>48.693*</td>
<td></td>
<td>51.198*</td>
</tr>
<tr>
<td>$I$</td>
<td>-3.833*</td>
<td></td>
<td>-4.145*</td>
</tr>
</tbody>
</table>

The superscript "*" indicates that the coefficient is significant at the 5% level. $^\text{a)}$: The CD statistics are computed on $\hat{e}_{it} = h_{it} - \hat{\alpha} - \bar{d}_{i} - \hat{\beta}'x_{it}$. 

19
Table 7: Sensitivity analysis for the income elasticity

<table>
<thead>
<tr>
<th>Country</th>
<th>Excluded</th>
<th>(I): FE</th>
<th></th>
<th>(II) Spatial MLE</th>
<th></th>
<th>(III) CCEP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Coeff.</td>
<td>Std.err.</td>
<td>Coeff.</td>
<td>Std.err.</td>
<td>Coeff.</td>
</tr>
<tr>
<td>Australia</td>
<td></td>
<td>0.863*</td>
<td>0.048</td>
<td>0.836*</td>
<td>0.048</td>
<td>0.441*</td>
</tr>
<tr>
<td>Austria</td>
<td></td>
<td>0.897*</td>
<td>0.045</td>
<td>0.875*</td>
<td>0.045</td>
<td>0.466*</td>
</tr>
<tr>
<td>Canada</td>
<td></td>
<td>0.855*</td>
<td>0.048</td>
<td>0.851*</td>
<td>0.047</td>
<td>0.456*</td>
</tr>
<tr>
<td>Denmark</td>
<td></td>
<td>0.826*</td>
<td>0.044</td>
<td>0.826*</td>
<td>0.044</td>
<td>0.475*</td>
</tr>
<tr>
<td>Finland</td>
<td></td>
<td>0.878*</td>
<td>0.046</td>
<td>0.871*</td>
<td>0.046</td>
<td>0.475*</td>
</tr>
<tr>
<td>Germany</td>
<td></td>
<td>0.851*</td>
<td>0.050</td>
<td>0.832*</td>
<td>0.048</td>
<td>0.488*</td>
</tr>
<tr>
<td>Iceland</td>
<td></td>
<td>0.875*</td>
<td>0.047</td>
<td>0.853*</td>
<td>0.047</td>
<td>0.359*</td>
</tr>
<tr>
<td>Ireland</td>
<td></td>
<td>1.086*</td>
<td>0.053</td>
<td>1.058*</td>
<td>0.054</td>
<td>0.437*</td>
</tr>
<tr>
<td>Japan</td>
<td></td>
<td>0.911*</td>
<td>0.053</td>
<td>0.912*</td>
<td>0.052</td>
<td>0.410*</td>
</tr>
<tr>
<td>Luxembourg</td>
<td></td>
<td>0.841*</td>
<td>0.052</td>
<td>0.840*</td>
<td>0.051</td>
<td>0.492*</td>
</tr>
<tr>
<td>Netherland</td>
<td></td>
<td>0.848*</td>
<td>0.047</td>
<td>0.832*</td>
<td>0.047</td>
<td>0.469*</td>
</tr>
<tr>
<td>New Zealand</td>
<td></td>
<td>0.848*</td>
<td>0.051</td>
<td>0.822*</td>
<td>0.051</td>
<td>0.478*</td>
</tr>
<tr>
<td>Norway</td>
<td></td>
<td>0.807*</td>
<td>0.049</td>
<td>0.808*</td>
<td>0.049</td>
<td>0.431*</td>
</tr>
<tr>
<td>Spain</td>
<td></td>
<td>0.844</td>
<td>0.045</td>
<td>0.833*</td>
<td>0.045</td>
<td>0.403*</td>
</tr>
<tr>
<td>United Kingdom</td>
<td></td>
<td>0.868*</td>
<td>0.048</td>
<td>0.852*</td>
<td>0.048</td>
<td>0.431*</td>
</tr>
<tr>
<td>United States</td>
<td></td>
<td>0.874*</td>
<td>0.046</td>
<td>0.862*</td>
<td>0.044</td>
<td>0.443*</td>
</tr>
</tbody>
</table>

Note: This table reports estimates of the income elasticity from equation $h_{it} = \hat{\alpha}_i + \beta x_{it} + u_{it}$ where $x_{it}$ contains $y_{it}$, $PUBL_{it}$, $OLD_{it}$, and $YOUNG_{it}$. The superscript "*" indicates that the coefficient is significant at the 5% level.
Table 8: CIPS panel unit roots tests on the residuals (intercept only case)

<table>
<thead>
<tr>
<th>Number of lags</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Regression I</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{u}_{it}$ (FE)</td>
<td>-1.611</td>
<td>-1.617</td>
<td>-1.767</td>
<td>-1.382</td>
</tr>
<tr>
<td>$\hat{u}_{it}$ (Spatial MLE)</td>
<td>-2.226*</td>
<td>-1.973</td>
<td>-1.940</td>
<td>-2.085</td>
</tr>
<tr>
<td>$\hat{u}_{it}$ (CCEP)</td>
<td>-2.674*</td>
<td>-2.353</td>
<td>-2.271</td>
<td>-2.344*</td>
</tr>
<tr>
<td><strong>Regression II</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{u}_{it}$ (FE)</td>
<td>-1.758</td>
<td>-1.614</td>
<td>-1.380</td>
<td>-1.509</td>
</tr>
<tr>
<td>$\hat{u}_{it}$ (Spatial MLE)</td>
<td>-1.945</td>
<td>-1.848</td>
<td>-1.752</td>
<td>-1.870</td>
</tr>
<tr>
<td>$\hat{u}_{it}$ (CCEP)</td>
<td>-2.430*</td>
<td>-2.027</td>
<td>-2.152</td>
<td>-2.216*</td>
</tr>
</tbody>
</table>

Notes: Tests are computed on residuals $\hat{u}_{it} = h_{it} - \alpha_i - \delta_t - \hat{\beta} x_{it}$

**"*" indicates that the test is significant at the 5% level.**

Table 9: CCEP estimation of Error Correction Models

<table>
<thead>
<tr>
<th></th>
<th>Regression I</th>
<th>Regression II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>Std.err.</td>
</tr>
<tr>
<td>$h_{i,t-1} - \hat{\beta} x_{i,t-1}$</td>
<td>-0.282*</td>
<td>0.040</td>
</tr>
<tr>
<td>$\Delta h_{i,t-1}$</td>
<td>0.181*</td>
<td>0.055</td>
</tr>
<tr>
<td>$\Delta y_{it}$</td>
<td>0.289*</td>
<td>0.083</td>
</tr>
<tr>
<td>$\Delta Publi_{it}$</td>
<td>-</td>
<td>0.290</td>
</tr>
<tr>
<td>$\Delta Old_{it}$</td>
<td>-</td>
<td>-0.254</td>
</tr>
<tr>
<td>$\Delta Young_{it}$</td>
<td>-</td>
<td>0.344</td>
</tr>
<tr>
<td><strong>CD statistics</strong>(§)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$CD_{LM}$</td>
<td>66.151</td>
<td>57.534</td>
</tr>
<tr>
<td>$I$</td>
<td>-0.807</td>
<td>-2.249</td>
</tr>
</tbody>
</table>

Notes: **"*** indicates that the coefficient is significant at the 5% level.

(§): The CD statistics are computed on $\hat{e}_{it} = \Delta h_{it} - \hat{\alpha}_i - \hat{\theta} \left( h_{it} - \hat{\beta} y_{it} \right) - \hat{\theta} \Delta h_{i,t-1} - \hat{\delta} \Delta x_{i,t-1}$.  

21