Panel data stationarity and cointegration: An application to the determinants of health care expenditure

Vincenzo Atella*

Giorgia Marini†

This version 2005, December

Abstract

Since Newhouse (1977) seminal work, several authors have attempted to empirically study the determinants of health care expenditure. From the literature review, it is clear that research interest has moved from the necessity to find an answer to a question («Is health care expenditure a luxury good?») to the necessity to solve an issue of cointegration between health and its determinants. However, despite the large amount of literature produced, the matter of cointegration is still an open issue and spurious regression problems are still highlighted. This paper re-examines the question of stationarity and cointegration on a panel of OECD countries, giving particular emphasis to panel tests. Results from our methodological analysis can be summarised as follows. Cointegration is strongly affected by the power of a test: as far as tests are developed continuously, results on cointegration will be affected consequently. Thus it will be rather difficult to reach a general agreement and a clear cut conclusion about health care expenditure determinants and their cointegration. Our suggestion for future research in the field of health expenditure and its determinants is to test and verify stationarity and cointegration with the most powerful tests.

JEL classification: C12; C22; C23; I10

Keywords: Stationarity and cointegration, health care expenditure, static and dynamic panels, OECD, panel tests

*SEFEMEC - CEIS, Department of Economics, University of Rome Tor Vergata, Via Columbia 2, 00133 Rome, Italy.
†*Corresponding author.* Current Affiliation: Centre for Health Economics, Alcuin `A' Block, University of York, Heslington, YORK, YO10 5DD, UK. Phone: +44 (0)1904 321421, Fax: +44 (0)1904 321454, E-mail: *gm509@york.ac.uk*

1. Introduction

Since Newhouse (1977) seminal work, several authors have attempted to empirically study the determinants of health care expenditure. These studies can be classified depending on the type of statistical tools and approaches they apply.

A first group of individual country studies (among them, Hansen and King (1996) and Blomqvist and Carter (1997)) investigates the relationship between a country's health care expenditure and its income: $HCE_i = f(Y_i, X_i)$, where HCE_i represents health care expenditure of country *i*, Y_i is its GDP, and X_i is a vector of regressors such as dependency rate, urbanisation, extent of public sector provision, number of practicing physicians, and in-patient care expenditure. $f(\cdot)$ is the functional form linking the three variables. Main results of this stream of research are the following: *(i)* aggregate income is the most important component of health care expenditure; *(ii)* choice of the functional form affects income elasticity; *(iii)* the inclusion of the vector X does not affect the explanatory power of GDP and the value of income elasticity. The vector of additional regressors is included to capture the presence of omitted variables bias in income coefficient, and thus to explain the contrast between high explanatory power of GDP and high values of income elasticity found at aggregated level and low values found in several studies at household level.

A second group based on panel data techniques (among them, Gerdtham (1992), Gerdtham *et al.* (1992b), Gerdtham *et al.* (1998) and Hitiris and Posnett (1992)) tests the presence of country- and time-specific effects, using a relationship between countries' health care expenditure and their income similar to the relationship used in cross-sectional studies: $HCE_{it} = f(Y_{it}, X_{it})$. As for the cross-sectional studies, the main results of this stream of research can be summarized as follows: *(i)* aggregate income is the most important component of health care expenditure; *(ii)* results on income elasticity depend on the choice of the functional form; *(iii)* the inclusion of the vector X does not affect the explanatory power of GDP and the value of income elasticity; *(iv)* there exists either a country- or a time-specific effect, whether the model is either a random or a fixed effects model.

Within the latter stream of research, some studies (among them, Hansen and King (1996), Blomqvist and Carter (1997), McCoskey and Selden (1998), Hitiris (1997), Roberts (2000), and Gerdtham and Lothgren (2000)) test whether the relationship between income and health expenditure (i.e. the income elasticity when the functional form is log-linear) is a cointegrating relation. Conflicting results regarding stationarity and cointegration involve the use of panel tests against time series based tests (Gerdtham and Lothgren (2000)) and the omission of a structural break in the unit roots test (Carrion-i-Silvestre *et al.* (2005)).

Despite the large amount of literature produced however, the matter of cointegration is still an open issue and spurious regression problems are still highlighted (Roberts (2000)). This paper re-examines the question of stationarity and cointegration on a panel of OECD countries, giving particular emphasis to panel tests.[3] The major reason to prefer panel to time series tests is related to the weakness of the latter: the combination of time and cross-section information mitigates the lack of power that the time series based unit root and cointegration tests show, especially when the time series may not be very long but similar data may be available across a cross-section of units.

Results from our methodological analysis show that cointegration is strongly affected by the power of a test. As far as tests are developed continuously, results on cointegration will be affected consequently. Thus it will be rather difficult to reach a general agreement and a clear cut conclusion about health care expenditure determinants and their cointegration.

The rest of the paper is organized as follows. Sections 2 and 3 describe the country-by-country and panel stationarity and cointegration tests. Section 4 presents the OECD data and the results for the tests. Section 5 concludes.

2. Unit roots and stationarity tests

2.1 Tests of null of unit roots

A standard approach to test for non-stationarity in time series is to estimate an augmented Dickey-Fuller (ADF) regression

$$\Delta y_{it} = \alpha_i + \delta_i t + \beta_i y_{it-1} + \sum_{j=1}^{p_i} \rho_{ij} \Delta y_{it-j} + \varepsilon_{it}, i = 1,\ldots,N; t = 1,\ldots,T \qquad (1)$$

where $\Delta y_{it} = y_{it} - y_{it-1}$ and t is a linear trend. The number of lags p is such that residuals are serially incorrelated. The null hypothesis $H_0 : \beta_i = 0$ for all i states that the data generating process can be characterized as a difference stationary $I(1)$ process, and it is tested against the alternative $H_A : \beta_i < 0$ for all i, based on the *t*-statistic of the β_i estimate.[4]

Among the panel tests for unit roots we apply the Levin *et al.* (2002) test, which is an extension of the Levin and Lin (1993) test, the Im *et al.* (1997 and 2003) test and the Maddala and Wu (1999) test. From the early paper developed by Engle and Granger (1987), the Levin *et al.* (LLC) test represents a clear cut with other tests based on estimators and statistics having limiting distributions which are complicated functionals of Wiener processes. The main innovation introduced by the LLC test, later adopted by Im *et al.* (IPS) and Maddala and Wu (MW), is an estimator having Gaussian distribution in the limit.

[3]Tests are carried out over the interval 1971-2000 for 20 OECD countries (Austria, Australia, Belgium, Canada, Denmark, Finland, France, Germany, Iceland, Ireland, Italy, Japan, Luxembourg, Netherland, New Zealand, Norway, Spain, Switzerland, UK and USA). The data series on Czech Republic, Greece, Hungary, Korea, Mexico, Poland, Portugal, Slovak Republic, Sweden and Turkey were much shorter than for the other countries so they were excluded from the analysis.

[4]The ADF test is the starting point of most of the unit root panel test described in this section as it is more robust and more powerful than the Phillips-Perron (1988) test.

The structure of the LLC test assumes that each individual unit in the panel shares the same $AR(1)$ coefficient, but allows for individual effects, time effects and possibly a time trend:

$$\Delta y_{it} = \alpha_i + \theta_t + \delta_i t + \beta_i y_{it-1} + \sum_{j=1}^{p_i} \rho_{ij} \Delta y_{it-j} + \varepsilon_{it}, i = 1,\ldots,N; t = 1,\ldots,T \tag{2}$$

The null hypothesis $H_0 : \beta_i = 0$ for all i is tested against the alternative $H_A : \beta_i = \beta < 0$ for all i. Lags of the dependent variable may be introduced to allow for serial correlation in the errors. The test may be viewed either as a pooled DF test or as an ADF test when lags are included, with the null hypothesis of non-stationarity. The model is a two-step regression:

First step We estimate $\hat{\beta}_i$ from the regression of the residuals from regression (2), $\hat{e}_{it}$, on the residuals from regression of y_{it-1} on remaining variables in (2), $\hat{V}_{it-1}$.

Second step We standardise the above residuals with the standard deviation of $\hat{e}_{it}$, $\hat{\sigma}_{e_i}$, and we estimate $\hat{\beta}$ (for all i and t) from the regression of the standardised residuals from regression (2), $\tilde{e}_{it}$, on the standardised residuals from regression of y_{it-1} on remaining variables in (2), $\tilde{V}_{it-1}$.

The LLC statistic is given by

$$t_\beta^* = \frac{t_{\beta=0} - N\tilde{T}\hat{S}_{NT}\hat{\sigma}_\varepsilon^{-2} RSE\left(\hat{\beta}\right)\mu_{\tilde{T}}}{\sigma_{\tilde{T}}} \tag{3}$$

where $t_{\beta=0}$ is the *t*-statistic associated to the $\hat{\beta}_i$, under the null hypothesis that $\beta_i = 0$, $\tilde{T} = (T - \bar{p} - 1)$ and $\bar{p} = N^{-1} \sum_{i=1}^{N} p_i$, $\hat{S}_{NT}^2$ is the standardised variance of y_{it}, $\hat{\sigma}_\varepsilon^2$ is the variance of standardised residuals and $RSE\left(\hat{\beta}\right)$ is the residuals standard error in estimating $\hat{\beta}_i$. $\mu_{\tilde{T}}$ and $\sigma_{\tilde{T}}$ are mean and standard deviation adjustment terms computed by Monte Carlo simulations. Under the null that $\beta_i = 0$, the panel test t_β^* is distributed as standard normal: $t_\beta^* \Rightarrow N(0,1)$.

Despite the innovation of the paper, the LLC test contains all the unsolved key elements systematically discussed by the literature on stationarity and cointegration: «*(a)* the demonstration of asymptotic normality; *(b)* the necessity of focusing on the rates at which T and N tend to infinity; *(c)* the issue of homogeneity versus heterogeneity; *(d)* the maintained assumption of independence across cross-section units; and *(e)* the corrections or modifications required to allow for dependent and heteroscedastic error processes and the resulting endogeneity of the regressors.» See Banerjee (1999), p.614.

In this contest, IPS and MW propose important relaxations of assumptions *(c)* and *(d)*. The IPS test represents an extension and a generalization of LLC test since it allows for heterogeneity in the value of β_i under the alternative hypothesis. The model under investigation is represented by equation (2).

The null hypothesis $H_0 : \beta_i = 0$ for all i is tested against the stationary alternative $H_1 : \beta_i < 0$ for $i = 1, \ldots, N_1$ and $\beta_i = 0$ for $i = N_1 + 1, \ldots, N$, where β_i is allowed to differ between groups and the fraction N_1/N is stationary. If the null hypothesis cannot be rejected, panel data series are difference-stationary around a linear trend.

In response to the objections of Pesaran and Smith (1995) on the use of pooled panel estimators for processes which display heterogeneity, IPS propose the use of mean-group *t*-bar statistic to test for the null hypothesis $H_0 : \beta_i = 0$

$$Z_{\bar{t}} = \frac{\sqrt{N}\left(\bar{t}_{NT} - E\left(\bar{t}_{NT}\right)\right)}{\sqrt{Var\left(\bar{t}_{NT}\right)}} \quad (4)$$

where $\bar{t}_{NT} = \frac{1}{N}\sum_{i=1}^{N} t_{iT}(p_i)$ and $t_{iT}(p_i)$ is the individual *t*-statistic for testing $\beta_i = 0$ for all i, while $E(\bar{t}_{NT})$ and $Var(\bar{t}_{NT})$ are obtained by stochastic simulation. Under the null hypothesis of non-stationarity, the standardised *t*-bar statistic converges to a standard normal distribution, $Z_{\bar{t}} \Rightarrow N(0,1)$. IPS demonstrate that the statistic $Z_{\bar{t}}$ has better finite sample performances than the statistic t_{β^*}.

«MW argue that while the IPS test relaxes the assumption of homogeneity of the root across the units, several difficulties still remain: *(a)* T is the same for all the cross-section units; *(b)* the critical values of LLC and IPS are sensitive to the choice of lag lengths in the ADF regressions; *(c)* the results in these papers apply to only a limited class of tests of the unit root hypothesis in panels; *(d)* while IPS allow for a limited amount of cross-correlation across units by allowing for common time effects, MW suggest quite rightly that in many real-world applications the cross-correlations are unlikely to take this simple form.» See Banerjee (1999), p.616.

The MW test is based on the combination of p-values of the individual *t*-statistics for the unit roots. The statistic is obtained in two steps: first, we estimate ADF regression (1), under the null hypothesis of unit roots; then, we build the statistic

$$-2\sum_{i=1}^{N}\ln\left(\pi_i\right) \quad (5)$$

where the p-values π_i associated to the individual *t*-statistics t_{iT} are independent uniform, and $-2\ln(\pi_i)$ is distributed as a chi-squared with 2 degrees of freedom. Therefore the *F* statistics (5) is distributed as a chi-squared with $2N$ degrees of freedom.

MW prove that the Fisher's test is preferred to LLC test and to IPS test for several reasons:

- LLC and IPS tests test a very restrictive hypothesis;
- Fisher test dominates IPS and LLC, both without and with a linear trend. See Maddala and Wu (1999) Table 1, p.640 and Table 2, p.642;

- Fisher test has the highest power. See Maddala and Wu (1999) Table 3, p.644.

In particular, under the assumption of no cross-sectional correlation in errors, IPS test is slightly more powerful than Fisher test, and vice versa under the assumption of cross-sectional correlation. Both tests are more powerful than LLC test, for any assumption on the errors. Finally when T is large but N is small, size distortion is smaller with Fisher test than with IPS or LLC tests.

2.2 Tests of null of stationarity

Despite the evidence of increased power in the panel tests of null of unit roots, the non-rejection of the panel unit root tests may still be due to multicollinearity. An alternative way to handle this potential multicollinearity is to switch the null and alternative hypothesis and consider the stationarity test by Kwiatkowski *et al.* (1992). Under the null of this test the series are specified to be trend stationary, and under the alternative the series are difference stationary. Comparing results from the ADF and Kwiatkowski *et al.* (KPSS) tests can give some insight into the stationarity properties of the series. If both the ADF and the KPSS tests fail to reject the null hypotheses or if both reject, we have mixed results and can only conclude that the data are not informative enough. But, if the ADF tests fail to reject the null and the KPSS tests reject, we have more confidence in the results that the series under consideration are in fact non-stationary.

The KPSS test is based on the decomposition of the series into the sum of a deterministic trend, a random walk, and a stationary error: $y_t = \delta t + r_t + \varepsilon_t$ where $r_t = r_{t-1} + u_t$ with $u_t \sim iid(0, \sigma_u^2)$. The initial value r_0 is treated as fixed and serves the role of an intercept. The stationarity hypothesis is simply $\sigma_u^2 = 0$. Since ε_t is assumed to be stationary, under the null hypothesis y_t is trend-stationary.

The one-sided LM statistic is

$$LM_i = \frac{\sum_{t=1}^{T} S_{it}^2}{\hat{\sigma}_{i\varepsilon}^2} \qquad (6)$$

where S_{it} is the partial sum process of the residuals from the regression of y_{it} on an intercept and time trend and $\hat{\sigma}_{i\varepsilon}^2$ is the estimate of the error variance from the same regression.

The Hadri (2000) test is a stationarity test based on the average of the N country-specific KPSS LM-statistic: $\overline{LM_\tau} = N^{-1}\sum_{i=1}^{N} LM_i$. Under the assumption that errors are independent and normally distributed, across i and over t, Handri proves that the statistic $\overline{LM}$ for the null of stationarity has the following limiting distribution:

$$Z_\tau = \frac{\sqrt{N}\left(\overline{LM}_\tau - \xi_\tau\right)}{\zeta_\tau} \Rightarrow N(0,1) \qquad (7)$$

where $\xi_\tau = E\left(\int V_2^2\right)$ and $\zeta_\tau^2 = Var\left(\int V_2^2\right)$ and V is a standard Brownian motion.

3. Cointegration tests

3.1 Tests of null of no-cointegration

As proved by Engle and Granger (1987), «the regression of non-stationary variables is likely to result in a spurious regression problem, by which the apparent close correlation between two variables is the result of common trends rather than evidence of any real economic relationship. Non-stationarity can arise from deterministic and/or stochastic trends in the data, and if variables possess a stochastic trend component they should be differenced to impose stationarity and to avoid misleading inference. In the special case where a linear combination of non-stationary variables is itself stationary, then this combination represents the cointegrating or long-run relationship, which can be specified in levels with short-run dynamics modelled via an error correction process.» See Roberts (2000), p.280.
We thus consider the following system of cointegrated regressions

$$y_{it} = \alpha_i + \delta_i t + \beta_i x_{it} + +\varepsilon_{it}, i = 1,\ldots,N; t = 1,\ldots,T$$

and we estimate an ADF regression where the estimated residuals $\hat{\varepsilon}_{it}$ are regressed on p lagged values $\hat{\varepsilon}_{it-j}$

$$\Delta\hat{\varepsilon}_{it} = \rho\hat{\varepsilon}_{it-1} + \sum_{j=1}^{p}\phi_j\Delta\hat{\varepsilon}_{it-j} + \upsilon_{it} \qquad (8)$$

Therefore, the test of no-cointegration is a test of the null hypothesis of no-cointegration $H_0 : \rho = 1$ against the alternative that variables are cointegrated $H_A : \rho < 1$. Under the null hypothesis, the *t*-statistic $t_\rho = \hat{\rho}/se(\rho)$ is non-standard and critical values for ADF tests are not reliable since the size of the test will be distorted. The distribution of t_ρ is found with Monte Carlo simulation by Phillips and Ouliaris (1990).
Among the panel tests of null of no-cointegration we perform the LLC and the IPS tests, which are modifications of the residual-based tests for cointegration regression in panel data proposed by Kao (1999) and Pedroni (1999), respectively.[5] Both tests are applied to the residuals of the cointegration relation and are performed by a standardised statistic for the equation (8). The LLC statistic is a modification of the *t*-statistic t_β^* (3)

$$t_\rho^* = \frac{t_{\rho=0} - N\tilde{T}\hat{S}_{NT}\hat{\sigma}_\varepsilon^{-2}RSE(\hat{\rho})\mu_{\tilde{T}}}{\sigma_{\tilde{T}}}$$

where $t_{\rho=0}$ is the *t*-statistic associated to $\hat{\rho}$. Under the null hypothesis of no-cointegration, the panel

[5]Pedroni's (1995, 1997) results for cointegrating tests in heterogeneous panels with a single regressor are published in

test t_ρ^* is distributed as standard normal $t_\rho^* \Rightarrow N(0,1)$. On the other hand, the IPS statistic is a modification of the t-bar statistic $Z_{\bar{t}}$ (4)

$$Z_{\bar{t}_\rho} = \frac{\sqrt{N}\left(\bar{t}_\rho - E\left(\bar{t}_\rho\right)\right)}{\sqrt{Var\left(\bar{t}_\rho\right)}}$$

where $\bar{t}_\rho$ is the mean of the estimated statistic. Under the null hypothesis of no-cointegration, the standardised t-bar statistic converges to a standard normal distribution $Z_{\bar{t}\rho} \Rightarrow N(0,1)$.

3.2 Tests of null of cointegration

Starting from Engle and Granger (EG) error correction representation of a VAR(p)-model

$$\Delta\mathbf{y}_t = \boldsymbol{\alpha} + \boldsymbol{\zeta}_0\mathbf{y}_{t-1} + \boldsymbol{\zeta}_1\Delta\mathbf{y}_{t-1} + \ldots + \boldsymbol{\zeta}_{p-1}\Delta\mathbf{y}_{t-p+1} + \boldsymbol{\varepsilon}_t \tag{9}$$

where $\mathbf{y}_t$ denotes a $(n \times 1)$ vector and follows a $VAR(p)$ in levels, each variable y_{it} is $I(1)$ and errors are normally distributed, Johansen (1995) introduces a likelihood ratio test, the trace statistic, to test the cointegrating rank of the matrix $\boldsymbol{\zeta}_0$. The Johansen test of cointegration is a test of null hypothesis of h cointegrating relations among the elements of $\mathbf{y}_t$ against the alternative that there are n cointegrating relations, where n is the number of elements of $\mathbf{y}_t$ and $h \le n$. Under the null H_0, the VAR is restricted by the requirement that $\boldsymbol{\zeta}_0 = -\mathbf{BA}'$, for $\mathbf{B}$ an $(n \times h)$ matrix and $\mathbf{A}'$ an $(h \times n)$ matrix, and only h linear combinations of the levels of $\mathbf{y}_{t-1}$ can be used in the regression (9). The largest value achieved by the associated log-likelihood function is

$$L_0^* = -\frac{Tn}{2}\log(2\pi) - \frac{Tn}{2} - \frac{T}{2}\log\left|\hat{\boldsymbol{\Sigma}}_{\mathbf{UU}}\right| - \frac{1}{2}\sum_{i=1}^{h}\log\left(1 - \hat{\lambda}_i\right)$$

where $\hat{\lambda}_i$ is the i-th eigenvalue of the matrix $\hat{\boldsymbol{\Sigma}}_{\mathbf{VV}}^{-1}\hat{\boldsymbol{\Sigma}}_{\mathbf{VU}}\hat{\boldsymbol{\Sigma}}_{\mathbf{UU}}^{-1}\hat{\boldsymbol{\Sigma}}_{\mathbf{UV}}$. $\hat{\boldsymbol{\Sigma}}_{\mathbf{UU}} \equiv (1/T)\sum_{t=1}^{T}\hat{\mathbf{u}}_t\hat{\mathbf{u}}_t'$ is the variance matrix of residuals $\mathbf{u}$ from regression of $\Delta\mathbf{y}_t$ on vectors $\Delta\mathbf{y}_{t-1}, \Delta\mathbf{y}_{t-2}, \ldots, \Delta\mathbf{y}_{t-p+1}$, $\hat{\boldsymbol{\Sigma}}_{\mathbf{VV}} \equiv (1/T)\sum_{t=1}^{T}\hat{\mathbf{v}}_t\hat{\mathbf{v}}_t'$ is the variance matrix of residuals $\mathbf{v}$ from regression of $\mathbf{y}_{t-1}$ on vectors $\Delta\mathbf{y}_{t-1}, \Delta\mathbf{y}_{t-2}, \ldots, \Delta\mathbf{y}_{t-p+1}$, and $\hat{\boldsymbol{\Sigma}}_{\mathbf{VU}} \equiv \hat{\boldsymbol{\Sigma}}_{\mathbf{UV}} \equiv (1/T)\sum_{t=1}^{T}\hat{\mathbf{u}}_t\hat{\mathbf{v}}_t'$ is the covariance matrix of residuals $\mathbf{v}$ and $\mathbf{u}$.

Under the alternative H_A, every linear combination of $\mathbf{y}_t$ is stationary, in which case $\mathbf{y}_{t-1}$ would appear in (9) without constraints and no restrictions are imposed on $\boldsymbol{\zeta}_0$. The value of the

Pedroni (2004).

log-likelihood function in absence of constraints is

$$L_A^* = -\frac{Tn}{2}\log(2\pi) - \frac{Tn}{2} - \frac{T}{2}\log\left|\hat{\boldsymbol{\Sigma}}_{\mathbf{UU}}\right| - \frac{1}{2}\sum_{i=1}^{n}\log\left(1-\hat{\lambda}_i\right)$$

The likelihood ratio test of H_0 against H_A is based on the trace statistic $L_A^* - L_0^* = -\frac{1}{2}\sum_{i=h+1}^{n}\log\left(1-\hat{\lambda}_i\right)$. Under the null hypothesis the likelihood ratio statistic $L_A^* - L_0^*$ is non-standard and critical values for the test are not reliable. Alternative critical values are provided by Osterwald and Lenum (1992).

Larsson *et al.* (2001) enlarge on this test, proposing a panel test that utilizes the information available in the panel data. The relevant innovation ported by this test is the presence of multiple cointegrating vectors. The Larsson *et al.* (LLL) test is a likelihood panel test of cointegration rank based on the average of the N individual rank trace statistics developed by Johansen. LLL consider the heterogeneous panel VAR

$$\Delta\mathbf{y}_{it} = \boldsymbol{\alpha}_i + \boldsymbol{\zeta}_{i0}\mathbf{y}_{it-1} + \boldsymbol{\zeta}_{i1}\Delta\mathbf{y}_{it-1} + \ldots + \boldsymbol{\zeta}_{ip-1}\Delta\mathbf{y}_{it-p+1} + \boldsymbol{\varepsilon}_{it} \quad (10)$$

where errors are normally distributed. The cointegration test is a test of null of rank at most equal to h against the alternative of rank equal to p. The standardised statistic used as a basis for the panel test is defined by

$$\Upsilon_{\overline{L_A^*-L_0^*}} = \frac{\sqrt{N}\left(\overline{L_A^*-L_0^*} - E(Z_k)\right)}{\sqrt{Var(Z_k)}} \quad (11)$$

where $\overline{L_A^* - L_0^*}_{NT} = \frac{1}{N}\sum_{i=1}^{N}(L_A^* - L_0^*)_{iT}$ is the likelihood ratio test and $E(Z_k)$ and $Var(Z_k)$ are mean and variance of the asymptotic trace statistic, defined on a Brownian motion of dimension $k = (p-h)$. Under the null hypothesis that $rank \leq h$, the statistic $\Upsilon_{\overline{L_A^*-L_0^*}}$ is distributed as standard normal $Z_{t^*} \Rightarrow N(0,1)$.

As the KPSS test is to be contrasted with the tests in the unit roots literature because it reverses the role of the null and the alternative hypothesis, similarly the common trend test by Nyblom and Harvey (2000) has to be contrasted with the EG procedure and the Johansen test. Nyblom and Harvey (NH) test the validity of a specific value of the rank of the covariance matrix of the disturbances driving the multivariate random walk, which is equal to the number of common trends in the set of series. This test is very simple, since it does not require the specification of a model (in contrast to the Johansen approach).

The test considers the null hypothesis that multiple time series are stationary, or stationary around a deterministic trend, against the alternative that a multivariate random walk component is present. The multivariate random walk with noise model, for a univariate time series, can be written as:

$$y_t = \mu_t + \varepsilon_t$$
$$\mu_t = \mu_{t-1} + \eta_t$$

where vector error process ε_t is *i.i.d.* covariance matrix Σ_ε while the vector η has covariance matrix Σ_η. We assume that Σ_ε is a positive definite matrix $n \times n$. If $\Sigma_\eta = 0$, the multivariate random walk becomes a constant level, and y is a (level) stationary series.

The null hypothesis is then $H_0 : \Sigma_\eta = 0$ which test for whether there is any non-stationarity in the system. Under the condition that there are N independent random walks in the set of time series, Σ_η is of full rank (N). The test involves roots of the matrix equation:

$$\left|\Sigma_\eta - q_i \Sigma_\varepsilon\right| = 0, i = 1,..., N$$

where $Q = diag(q_1, \ldots, q_N)$, $P\Sigma_\eta P' = Q$ and P is defined from factorization of $P\Sigma_\varepsilon P' - I$. The null hypothesis sets all these roots equal to zero; without any beliefs about relative magnitudes the alternative hypothesis assumes $q_1 = \ldots = q_N = q > 0$ corresponding to the homogeneous model $\Sigma_\eta = q\Sigma_\varepsilon$. The test has critical values which depend only on the number of series and the hypothesized number of common trends.

4 Results

4.1 Data

Health care expenditure components are identified following Hansen and King (1996). All data are annual. Per capita health care expenditure, per capita GDP and relative price of health care expenditure are estimated in GDP purchasing power parity, while the dependency rate and the fraction of public spending in health care expenditure are ratios. The relative price of health care expenditure is calculated as the ratio of health services price index to the GDP deflator. The dependency rate is defined as sum of population aged 0-14 and population 65 and over as percentage of population aged 15-64. The fraction of public financing used for health care expenditure is computed as the ratio of public expenditure on health care (as percentage of GDP) to government final consumption expenditure (as percentage of GDP). As most of the literature, all variables are expressed in natural logarithm.

Data on health care expenditure, GDP, relative price, public expenditure on health and dependency rate are taken from the OECD Health Data 2003 database, while the source for data on government consumption expenditure is World Development Indicators 2003 database.

4.2 Unit root and stationarity tests

Table 1 reports country-by-country ADF *t*-statistics and standardized panel data unit roots tests based

on ADF regressions with intercept and time trend -LLC and IPS- and a non-parametric test performed by MW. On the basis of individual tests, we cannot reject the null hypothesis of unit roots. However, knowing the low power of the time series test in presence of near unit roots series, we perform the three panel tests previously presented. When comparing LLC and IPS tests, we should bear in mind that, even though these tests are based on the same null hypothesis, the alternative is different: while the LLC test is based on homogeneous parameters, the IPS test is based on heterogeneous ones. As a consequence, the LLC test is based on pooled regression and the other on the combination of different independent tests. On the other hand, MW and IPS tests are directly comparable, as both tests are a combination of different independent tests.
The MW test strongly supports non-stationarity, since the statistic is always well below the critical values. On the other hand, the LLC test (t^*_β) does not provide a clear preference for the null hypothesis of unit roots «because the LLC test has to use the panel estimation method which is not valid if there is no need for pooling.» See Maddala and Wu 1999, p.637. This test suggests that all variables are non-stationary, except the dependency rate. The power increases sensibly when we implement the IPS test ($Z_{\bar{t}}$), but still the value of the statistic for health care price (-1.643) is very close to the critical threshold (-1.645).
In order to control for possible multicollinearity, we switch the null and the alternative hypotheses of unit roots tests and we assume that the series are specified as trend stationary under the null, and as difference stationary under the alternative. Table 2 presents results for country-by-country KPSS (1992) tests and standardized panel data stationary Hadri tests with intercept and time trend. On the basis of KPSS test, the hypothesis of trend stationarity can be rejected for health care expenditure, GDP and dependency rate for all countries, and for health price and public financing for all countries, except Iceland and Italy, respectively. These results are coherent with the non-rejection of null by ADF and are strongly supported by the implementation of the Hadri test. Therefore our variables are difference stationary $I(1)$ processes. These results are coherent with those found by Hansen and King (1996), Blomqvist and Carter (1997), Gerdtham and Löthgren (2000) and Roberts (2000).

4.3 Cointegration tests

Under the result that variables are $I(1)$ processes, the following equation

$$HCE_{it} = \alpha_0 + \alpha_1 t + \beta_1 Y_{it} + \beta_2 P_{it} + \beta_3 DR_{it} + \beta_4 PF_{it} + u_{it}$$

is a long-run relationship if it is a cointegrating regression, where *HCE* is per capita health care expenditure, *Y* is per capita GDP, *P* is the relative price of health care expenditure, *DR* is the dependency rate, *PF* is the fraction of public financing used for health care expenditure, *u* is the error term and *t* is a linear deterministic time trend used as proxy for technical change and common to all countries.

Table 3 presents result from two-steps EG procedure and panel tests (both IPS and LLC). Results from the EG test imply that we cannot reject the null of no-cointegration and therefore that the variables are no-cointegrated for all countries. However, looking at results from Johansen procedure, the hypothesis of $h = 0$ linear combinations of levels of health expenditure is not supported: for each country the null hypothesis of no-vector of cointegration is always rejected. Therefore we conclude that individual analysis of no-cointegration is contradictory. On the other hand, both IPS and (the weaker) LLC panel tests clearly reject the hypothesis of no-cointegration at 5% level, since both the statistic t_{ρ}^{*} (-5.576) and the statistic $Z_{\bar{t}_{\rho}}$ (-6.886) are much lower than the 5% (-1.645) and 10% (-1.280) critical levels. This result is also supported by the LLL test for the hypothesis $h = 0$. Contrary to country-by-country analysis, all panel tests of no-cointegration lead to a rejection of the hypothesis of no-cointegration.

Table 4 presents result from Johansen test and panel LLL test. In particular, the individual trace test confirms the existence of four vectors of cointegration for Austria, Canada, Denmark, Finland, Luxembourg and Norway, three for Australia, Belgium, Germany, Iceland, Italy, Netherlands and Switzerland, and two for the remaining countries.[6] This result is confirmed by the LLL test: for the null $h = 4$, the statistic $\Upsilon_{\overline{L_A^*-L_0^*}}$ (1.368) is well below the 5% critical value (1.645).

Comparing results from both classes of tests, we can conclude that variables are cointegrated. However, as Roberts (2000) suggests, a pooled regression does not represent a cointegrating relation since it introduces a common vector among the countries. Therefore, we check for robustness of our results applying the Nyblom and Harvey (2000) test on common trend among the countries. In particular, we consider the special case in which the rank of the covariance matrix K equals zero: that is, there are no common trends among the variables. In this context, the Nyblom and Harvey (NH) test can be used as a test for cointegration, since common trend implies cointegration, and vice versa. Thus, a failure to reject the null hypothesis of zero common trends is also an indication that the variables do not form a cointegrated combination. From Table 5 it is evident that the variables have a common trend since we always reject the null of no-common trend. Therefore variables form a cointegrating relation. Moreover, compared to KPSS results, the rejection of null of no-common trend ($\Sigma_{\eta} = 0$) is in line with the rejection of null of trend stationarity ($\sigma_u^2 = 0$).

5 Conclusions

Exploring the determinants of health care expenditure is a quite challenging task and still an open

---

[6] It is worth noting that the Johansen test is a sequence of tests, which allow applying the test to null hypothesis of both no-cointegration and cointegration. The null hypothesis of rank $h = 0$ (no-cointegrating relationship) is first tested and, if rejected, subsequent null hypotheses ($H_0 : h = 1$, $H_0 : h = 2$, and so on) are tested until a null can no longer be rejected.

issue. From the survey of the literature, it is clear that research interest has moved from the necessity to find an answer to a question («Is health care expenditure a luxury good?») to the necessity to solve an issue of cointegration between health and income. For this reason, the purpose of the paper is to re-examine the question of stationarity and cointegration with a methodological comparison of stationarity and cointegration tests, applied to a modification of model (2) in Roberts (2000).

We prove that new tests implemented by Maddala and Wu (1999), Larsson *et al.* (2001) and Nyblom and Harvey (2000) produce better results in favour of stationarity and cointegration between health care expenditure and its components. Notwithstanding, we prove that results on cointegration are strongly dependent on available tests and therefore that, as far as these tests are developed continuously, results are not definitive and are affected consequently. Thus, it will be rather difficult to reach a general agreement and clear cut conclusion about health care expenditure determinants and their cointegration. Our suggestion for future research in the field of health expenditure and its determinants is to always test and verify stationarity and cointegration with the most powerful tests.

References

- Banerjee, A. (1999), Panel data unit roots tests and cointegration: an overview, Oxford Bulletin of Economics and Statistics, Special Issue, pp.607-629.
- Blomqvist, G. and R.A.L. Carter (1997), Is health care really a luxury?, Journal of Health Economics 16, pp.207-229.
- Dickey D.A. and W.A. Fuller (1981), Likelihood-ratio statistics for autoregressive time-series with a unit roots, Econometrica 49, pp.1057-1072.
- Engle, R.F. and C.W.J. Granger (1987), Cointegration and error correction: representation, estimation and testing, Econometrica 50, pp.251-276.
- Gerdtham, U.-G. (1992), Pooling international health expenditure data, Health Economics 1, pp.217-231.
- Gerdtham, U.-G., B. Jönsson, M. MacFarlan and H. Oxley (1998), The determinants of the health expenditure in the OECD countries in Zweifel, P. (eds.), Health, The Medical Profession, and Regulation, Kluwer Academic Publishers.
- Gerdtham, U.-G. and M. Löthgren (2000), On stationary and cointegration of international health expenditure and GDP, Journal of Health Economics 19, pp.461-475.
- Gerdtham, U.-G., J. Sögaard, F. Andersson and B. Jönsson (1992), A pooled cross-section analysis of the health expenditure of the OECD countries in Zweifel, P. and Frech, H. (eds.), Health Economics Worldwide, Kluwer Academic Publishers.
- Gutierrez, L. (2003), On the power of panel cointegration tests: a Monte Carlo comparison, Economics Letters 80, pp.105-111.
- Hadri, K. (2000), Testing for stationarity in heterogeneous panel data, Econometrics Journal 3, pp.148-161.
- Hadri, K. and R. Larsson (2005), Testing for stationarity in heterogeneous panel data where the time dimension is finite, Econometrics Journal 8, pp.55-69.
- Hamilton, J. (1994), Time Series Analysis, Princeton University Press.
- Hansen, P. and A. King (1996), The determinants of health care expenditure: a cointegration approach, Journal of Health Economics 15, pp.127-137.
- Hansen, P. and A. King (1998), Health care expenditure and GDP: panel data unit roots test results - comment, Journal of Health Economics 17, pp.377-381.
- Hitiris, T., (1997), Health care expenditure and integration in the countries of the European Union, Applied Economics 29, pp.1-6.
- Hitiris, T. and J. Posnett (1992), The determinants and effects of health expenditure in developed countries, Journal of Health Economics 11, pp.173-181.
- Im, K.S., M.H. Pesaran and Y. Shin (1997), Testing for unit roots in heterogeneous panels,

Technical Report, Department of Applied Economics, University of Cambridge.

- Im, K.S., M.H. Pesaran and Y. Shin (2003), Testing for unit roots in heterogeneous panels, Journal of Econometrics 115, pp.53-74.
- Johansen, S. (1991), Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, Econometrica 59, pp.1551-1580.
- Kao, C. (1999), Spurious regression and residual-based tests for cointegration in panel data, Journal of Econometrics 90, pp.1-44.
- Kwiatkowski, D., P.C.B. Phillps, P. Schmitd, P. and Y. Shin (1992), Testing the null hypothesis of stationary against the alternative of a unit roots, Journal of Econometrics 54, pp.159-178.
- Larsson, R., J. Lyhagen and M. Lothgren (2001), Likelihood-based cointegration tests in heterogeneous panels, Econometric Journal 4, pp.109-142.
- Levin, A. and C.-F. Lin (1993), Unit root tests in panel data: New results, Mimeo, University of California, San Diego.
- Levin, A., C.-F. Lin and C.-S.J. Chu (2002), Unit roots tests in panel data: asymptotic and finite sample properties, Journal of Econometrics 108, pp.1-24.
- Maddala, G.S. and S. Wu (1999), A comparative study of unit roots tests with panel data and a new simple test, Oxford Bulletin of Economics and Statistics, Special Issue, pp.631-652.
- McCoskey, S.K. and T.M. Selden (1998), Health care expenditure and GDP: panel data unit roots test results, Journal of Health Economics 17, pp.369-376.
- Newhouse, J.P. (1977), Medical care expenditure: a cross-national survey, Journal of Human Resources 12, pp.115-125.
- Nyblom, J. and A. Harvey (2000), Test of common stochastic trends, Econometric Theory 16, pp.176-199.
- OECD (2003), OECD Health Data, A comparative analysis of 30 countries, OCSE Paris.
- Osterwald-Lenum, M. (1992), A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics, Oxford Bulletin of Economics and Statistics 54, pp.461-472.
- Pedroni, P. (1995), Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis, Indiana University Working Papers in Economics No. 95-013.
- Pedroni, P. (1997), Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis, New results, Indiana University Working Papers in Economics.
- Pedroni, P. (1999), Critical values for cointegration tests in heterogeneous panels with multiple regressors, Oxford Bulletin of Economics and Statistics, Special Issue, pp.653-670.

- Pedroni, P. (2004), Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis, Econometric Theory 20, pp.597-625.
- Pesaran, M.H., R. Smith (1995), Estimating long-run relationships from dynamic heterogeneous panels, Journal of Econometrics 68, pp.79-113.
- Phillips, P.C.B. and S. Ouliaris (1990), Asymptotic properties of residual based tests for cointegration, Econometrica 58, pp.168-193.
- Phillips, P.C.B. and P. Perron (1988), Testing for unit roots in time series regression, Biometrika 75, pp.335-346.
- Roberts, J. (2000), Spurious regression problems in the determinants of health care expenditure: a comment on Hitiris, Applied Economics Letters 7, pp.279-283.
- World Bank (2003), World Development Indicators, World Bank, Washington, D.C., USA.

Table 1: *Estimated lag orders, country-by-country ADF t-statistics and standardized panel data unit root LLC, IPS and MW tests based on ADF regressions with intercept and time trend for 20 OECD countries 1971–2000.*

Countries	Health care expenditure		GDP		Health care price		Dependency rate		Fraction of public financing	
	Lag order	ADF test	Lag order	ADF test	Lag order	ADF test	Lag order	ADF test	Lag order	ADF test
Australia	1	-3.572	1	-1.508	3	-3.353*	1	-3.329	1	-2.687
Austria	1	-2.891	1	-2.046	1	-2.741	2	-1.594**	1	-2.856
Belgium	1	-2.983	1	-1.679	1	-1.941	1	-3.194	1	-
Canada	1	-0.908	1	-2.098	1	-3.119	3	-3.146**	1	-2.749
Denmark	1	-1.726	1	-0.998	2	-3.180**	1	-2.796	1	-0.836
Finland	1	-0.791	1	-1.941	1	-2.282	1	-1.884	1	-1.816
France	1	-1.578	1	-1.989	1	-2.011	1	-2.966	1	-
Germany	1	-2.985	1	-1.23	1	-0.787	1	-1.793	1	-2.194
Iceland	1	-1.248	1	-1.975	1	-2.847	1	-2.245	1	-1.288
Ireland	1	-2.644	1	-2.076	1	-3.276	1	-1.592	1	-0.393
Italy	1	-1.032	1	-1.222	1	-1.348	1	-3.552	1	-2.494
Japan	1	-1.614	1	0.128	1	-2.295	1	-1.493	1	-1.739
Luxembourg	1	-1.052	1	-2.085	1	-2.147	1	-1.697	1	-2.935
Netherlands	1	-1.781	1	-2.006	1	-2.721	2	-1.728**	1	-2.233
New Zealand	3	-2.843**	1	-1.855	1	-1.930	1	-1.421	1	-2.110
Norway	1	-2.221	1	-2.247	1	-2.467	2	-3.570*	1	-2.555
Spain	1	-1.534	1	-1.930	1	-2.108	1	-2.450	2	-2.290**
Switzerland	1	-1.332	1	-0.514	4	-3.531*	2	-1.833**	1	-2.978
UK	1	-2.387	1	-1.355	1	-2.439	1	-3.187	1	-1.361
USA	1	-1.017	1	-1.349	1	-2.310	1	-2.904	1	-2.577
Panel tests										
t^*_β		-1.2619		-1.4387		-1.5479		-3.7694		-0.4348
$Z_{\bar{t}}$		1.581		1.547		-1.643		2.730		-0.397
$-2\sum_{i=1}^{N}\ln(\pi_i)$		31.520		13.778		51.318		53.112		31.899

Note: (a) The maximum lag order for the test (8 lags) is by default calculated from the sample size, using a rule provided by Schwert (1989) (from STATA 7.0). (b) For individual tests, 5% and 10% critical values are -3.588 and -3.233, respectively (from STATA 7.0) (c) For LL and IPS panel tests, 5% and 10% critical values are -1.645 and -1.280, respectively (from Hamilton (1994), Appendix B, p.751). (d) For MW test, 5% and 10% critical values of a chi-squared with 40 degrees of freedom are 55.758 and 51.805, respectively (from Hamilton (1994), Appendix B, p.754). (d) For the *p*-values π_i, we use the MacKinnon approximate *p*-values for the *t*-statistic t_{iT} provided by STATA 7.0. (e) * and ** represent 5% and 10% levels of significance.

Table 2: *Country-by-country KPSS tests and standardized panel data stationary Hadri test with intercept and time trend for 20 OECD countries 1971–2000.*

Countries	Health care expenditure	GDP	Health care price	Dependency rate	Fraction of public financing
	KPSS test	KPSS test	KPSS test	KPSS test	KPSS test
Australia	0.2016	0.2157	0.1422***	0.2126	0.1202**
Austria	0.1656	0.2147	0.1250**	0.1956	0.1883***
Belgium	0.2012	0.2136	0.1971	0.2164	-
Canada	0.2169	0.2191	0.1217*	0.2022	0.1164**
Denmark	0.2105	0.2156	0.1465***	0.1551	0.1189**
Finland	0.2192	0.2090	0.1632	0.2010	0.1669
France	0.2177	0.2160	0.2229	0.1997	-
Germany	0.2147	0.2185	0.1201	0.2017	0.1286
Iceland	0.2138	0.2093	**0.0456**	0.2101	0.1913
Ireland	0.1482	0.1431*	0.1686	0.2183	0.1350
Italy	0.2096	0.2198	0.2183	0.1556	**0.0771**
Japan	0.2112	0.2172	0.1390	0.1432*	0.1768
Luxembourg	0.1965	0.1597	0.1994	0.2143	0.1221
Netherlands	0.1906	0.2111	0.1723	0.2176	0.1191**
New Zealand	0.1409	0.2107	0.1923	0.2209	0.1427
Norway	0.2137	0.2065	0.1225	0.1595	0.1276
Spain	0.1520	0.2072	0.1784	0.1460	0.1470***
Switzerland	0.2155	0.2145	0.1654	0.2055	0.1567
UK	0.2187	0.2226	0.1430	0.1937	0.1921
USA	0.2169	0.2215	0.1375*	0.1938	0.1223*
Panel test					
Z_τ	19.515	21.202	18.333	19.614	8.404

Note: (a) Since «the tests have approximately correct size except when T is small and l is large» (from Kwiatkowski *et al.* (1992), p.170), we exclude the case $l = 8$. Following Gerdtham and Löthgren (2000), we set lag length to $l = \text{integer}[4(T/100)^{1/4}]$. (b) For individual tests, 5% and 10% critical values are 0.146 and 0.119, respectively (from Kwiatkowski *et al.* (1992)) (c) For Hadri panel test, 5% and 10% critical values are 1.645 and -1.280, respectively (from Hamilton (1994), Appendix B, p.751). (d) For Hadri panel test, serial dependence in the disturbances is taken into account using a Newey-West estimator of the long run variance. (e) *, ** and *** represent significance when l = 2, 1 0, respectively, while bold indicates that the variable is trend-stationary.

Table 3: *Estimated country-by-country EG tests and standardized panel IPS and LLC tests of no-cointegration based on ADF regressions of the residuals for 20 OECD countries 1971–2000.*

Countries	EG test
	t_ρ-statistic
Australia	-3.537
Austria	-4.206[a]
Belgium	-2.639
Canada	-3.017
Denmark	-1.650
Finland	-1.894
France	-3.076
Germany	-4.456*
Iceland	-3.322
Ireland	-2.724
Italy	-1.961
Japan	-3.812
Luxembourg	-2.397
Netherlands	-3.583
New Zealand	-3.268
Norway	-2.084
Spain	-2.699
Switzerland	-4.677*
UK	-2.771
USA	-1.476
Panel tests	
t^*_ρ	-5.576
$Z_{\bar{t}_\rho}$	-6.886

Note: (a) For individual tests, 5% and 10% critical values are -4.740 and -4.460, respectively (from Hamilton (1994), Appendix B, case 3, p.766) (b) For LL and IPS panel tests, 5% and 10% critical values are -1.645 and -1.280, respectively (from Hamilton (1994), Appendix B, p.751). (c) [a] represents significance with $l = 2$, while * represents significance at 5%.

Table 4: *Estimated country-by-country Johansen procedure and panel data LLL test with intercept and time trend for 20 OECD countries 1971–2000.*

Countries	Lag	Trace statistics					Rank
		h=0	h=1	h=2	h=3	h=4	
Australia	2	124.982	73.894	36.508	14.808	2.927	3
Austria	3	137.672	78.070	47.403	21.701	2.497	4
Belgium	3	61.575	36.309	16.200	3.738		3
Canada	3	125.527	70.544	33.369	17.148	1.078	4
Denmark	2	100.992	71.145	43.930	21.398	3.678	4
Finland	2	119.846	74.148	40.735	15.802	3.611	4
France	1	185.774	73.440	10.979	1.225		2
Germany	2	79.335	47.331	27.244	12.165	0.929	3
Iceland	1	251.360	115.563	37.311	0.419	0.050	3
Ireland	1	161.241	55.706	5.893	0.673	0.230	2
Italy	2	102.947	62.750	30.515	14.786	3.476	3
Japan	1	234.375	58.794	9.540	1.612	0.605	2
Luxembourg	3	119.374	72.634	38.821	18.395	0.012	4
Netherlands	2	141.015	81.342	38.731	13.260	0.764	3
New Zealand	1	194.033	96.463	11.428	1.811	0.328	2
Norway	2	99.875	62.138	37.307	15.706	0.005	4
Spain	1	185.407	60.680	17.908	2.076	0.079	2
Switzerland	2	148.230	88.719	43.517	13.615	4.137	3
UK	1	206.499	50.790	6.360	1.225	0.104	2
USA	1	286.807	111.321	10.045	2.318	0.144	2
Panel tests							
$\Upsilon_{L_A^*-L_0^*}$		57.710	29.487	11.000	4.971	1.368	

Note: (a) Critical values for the Trace test are tabulated in Phillips & Ouliaris (1990) and are 68.52, 47.21, 29.68, 15.412 and 3.76 for testing h=0, h=1, h=2, h=3, and h=4 respectively. (b) For LLL tests, 5% and 10% critical values are 1.645 and 1.280, respectively (from Hamilton (1994), Appendix B, p.751).

Table 5: *Panel data test on common trend performed by Nyblom and Harvey with intercept and time trend for 20 OECD countries 1971–2000.*

Nyblom and Harvey tests	Health care expenditure	GDP	Health care price	Dependency rate	Fraction of public financing
$\sum_{k=1}^{\infty}(\lambda_k)^{-2}x_k$	10.517	9.956	10.744	10.076	9.680

Note: (a) 5% and 10% critical values for the NH test are 4.4957 and 4.1794, respectively (from STATA 7.0).