Testing the long-run relationship between health expenditures and GDP in the presence of structural change: the case of Spain

Vicente Esteve* and José L. Martinez-Zahonero

Departamento de Economía Aplicada II, Universidad de Valencia, Avda. dels Tarongers, s/n, 46022 Valencia, Spain

This article examines the long-run relationship between per capita US$ PPP health expenditures (HE) and per capita US$ PPP national income (GDP), using Spanish data over the period 1960 to 2001. We extend previous analyses by addressing the question of whether this relationship is stable over time, allowing for structural changes at an unknown date. Our empirical results are consistent with the existence of a long-run relationship between both variables, with two structural changes in 1971 and 1991. On the other hand, health would have been characterized as a luxury commodity, even though increasingly less over time.

I. Introduction

The long-run relationship between per capita national income (GDP) and per capita health expenditures (HE) is the subject of a large literature in health economics. The emergence of the literature on unit roots and cointegration, as well as panel unit root tests, has provided an important impulse to the empirical testing of this relationship. A number of recent contributions aimed to testing this relationship using cointegration techniques include Hansen and King (1996), Blomqvist and Carter (1997), and Gerdtham and Löthgren (2000). In turn, some studies that use panel unit root tests include Blomqvist and Carter (1997), McCoskey and Selden (1998) and Gerdtham and Löthgren (2000), where no structural changes are considered; whereas Jewell et al. (2003) explicitly test for the presence of structural breaks.

When examining the long-run relationship between HE and GDP, it is important to determine whether or not this relationship is stable over time, or it exhibits some structural break, allowing the instability to occur at an unknown date. This can be justified since most of the recent empirical literature uses a rather long sample period (i.e from the first 1960s to the last of 1990s), where the OECD economies have experienced several structural changes in health care systems, preferences and technology (OECD, 1994); together with several expansionary and contractionary periods. The lack of control for structural breaks, in the series may be reflected in the parameters of the estimated models that, when used for inference or forecasting, can induce misleading results. In general, structural breaks are a problem for the analysis of economic series, since they are usually affected by either exogenous shocks or changes in policy regimes. As a consequence, the assumption of stability in the long-run relationship between HE and GDP would seem too restrictive, so that not allowing for structural breaks would be an important potential shortcoming of the past research using cointegration techniques.

*Corresponding author. E-mail: vicente.esteve@uv.es
Our empirical results are consistent with the existence of a cointegrating relationship between HE and GDP. More specifically, we estimate a long-run income elasticity of health spending $> 1$ (i.e. health would be a luxury commodity). However, there is strong evidence on the presence of a structural change in the cointegration relationship in 1971, once the 1967 Social Security System Act was fully enforced. There also appears strong evidence on the presence of a structural change in the cointegration relationship in 1991, once the 1986 General Health Care Act was fully enforced. We find up to three different regimes over the period 1960 to 2001, and in each regime the long-run income elasticity of health spending estimated is $> 1$, although it falls over time.

The rest of the article is organized as follows. Section II describes the data, and presents the estimation results. Finally, Section III summarizes the main conclusions.

II. Data and Empirical Results

In this section, we estimate the long-run relationship between GDP and HE in Spain in two steps. First, we first test for unit roots in order to determine the order of integration of the two series. Second, we estimate the cointegration relationship between the variables, and we discuss the possibility of structural changes in this relationship.

All data are from OECD (2003) Health Data File. We use annual data for the Spanish economy along the period 1960 to 2001. Both HE and GDP are measured in per capita, USS purchasing power parity and transformed in natural logarithms.

Unit root tests

We begin by examining the time-series properties of series. We use a modified version of the Dickey and Fuller (1979, 1981) test (DF) and a modified version of the Phillips and Perron (1988) tests (PP) proposed by Ng and Perron (2001) for the null of a unit root.

The majority of the conventional unit root tests (DF and PP types) generally suffer from three problems. First, Dejong et al. (1992) showed that many tests have low power when the root of the autoregressive polynomial is close to but less than unit. Second, Schwert (1989) and Perron and Ng (1996) showed that the majority of the tests suffer from severe size distortions when the moving-average polynomial of the first differenced
series has a large negative autoregressive root. Third, the implementation of unit root tests often necessitates the selection of an autoregressive truncation lag, \( k \). However, simulations have shown a strong association between \( k \) and the severity of size distortions and/or the extend of power loss, as discussed in Ng and Perron (1995). Recently, Ng and Perron (2001) set out to resolve the three problems. They proposed a class of modified tests, \( M_{\text{GLS}}^{\text{GLS}} \), originally developed by Stock (1999) as \( M \) tests, with GLS detrending of the data as proposed in Elliot et al. (1996), and using the MAIC (Modified Akaike Information Criteria). Also, Ng and Perron (2001) have proposed similar procedure to correct the distortions and/or the extent of power loss, as discussed in Ng and Perron (1995). Recently, Ng and Perron (2001) set out to resolve the three problems. They proposed a class of modified tests, \( M_{\text{GLS}}^{\text{GLS}} \), originally developed by Stock (1999) as \( M \) tests, with GLS detrending of the data as proposed in Elliot et al. (1996), and using the MAIC (Modified Akaike Information Criteria). Also, Ng and Perron (2001) proposed similar procedure to correct the problems of the standard Augmented Dickey–Fuller (ADF) test, ADFGLS.

Table 1 shows the results of the \( M_{\text{GLS}}^{\text{GLS}} \), \( M_{t}^{\text{GLS}} \), MSBGLS, and ADFGLS tests. All test statistics examine formally the null hypothesis that a series is \( I(1) \) against the alternative that it is \( I(0) \). As shown in the table, the null hypothesis of non-stationarity for the two series in levels cannot be rejected, independently of the test. Accordingly, both series would be \( I(1) \).

### The long-run relationship and income elasticity with structural breaks

Once analysed the order of integration of the series, we will estimate the long-run or cointegration relationship between GDP and HE. Given the relatively small sample size, we will estimate and test the coefficients of cointegration equation by means of the Dynamic Ordinary Least Squares (DOLS) method of Stock and Watson (1993), and following the methodology proposed by Shin (1994). This estimation method provides a robust correction to the possible presence of endogeneity in the explanatory variables, as well as of serial correlation in the error terms of the OLS estimation. Also, in order to overcome the problem of the low power of the classical cointegration tests under the presence of persistent roots in the residuals of the cointegration regression, Shin (1994) suggested a new test where the null hypothesis is that of cointegration. This test is similar to the implementation of the KPSS test\(^3\) in two stages, to the case of a cointegration relationship among a group of variables. The first step would consist of estimating a long-run dynamic equation including leads and lags of all the explanatory variables, the so-called DOLS regression; in our case:

\[
HE_t = \alpha_0 + \alpha_1 t + \beta \cdot GDP_t + \epsilon_t
\]

And the second step of the Shin test is based on the calculation of two LM statistics from the DOLS residuals, \( C_t \) and \( C_r \), to test for deterministic (when \( \alpha_1 = 0 \)) and stochastic (when \( \alpha_1 \neq 0 \)) cointegration, respectively.

The coefficients from the DOLS regression and the results of the Shin test are reported in Table 2. Since the concept of deterministic cointegration is stronger than the concept of stochastic cointegration, we first test for the presence of stochastic cointegration and then test for the presence of deterministic cointegration sequentially. First, the null of stochastic cointegration is not rejected at the 1% level. Next, we check for the presence of deterministic cointegration using the demeaned specification. Now, the null of deterministic cointegration is not rejected at the 1% level. We may conclude that there is strong evidence of deterministic cointegration

---

1 See Ng and Perron (2001) and Perron and Ng (1996) for a detailed description of these tests.
2 In the case of MSBGLS test, the unit root hypothesis is rejected in favor of stationarity when the estimated value is smaller than some appropriate critical value.
3 Kwiatkowski et al. (1992) statistics for testing the null hypothesis of stationary against the alternative of a unit root.
between HE and GDP, with a long-run coefficient estimated between HE and GDP of 1.52, i.e., a long-run income elasticity of health care spending greater than 1.

Table 2. Estimation of long-run relationships: Stock-Watson-Shin cointegration tests

<table>
<thead>
<tr>
<th>Parameter estimates</th>
<th>Deterministic cointegration</th>
<th>Stochastic cointegration</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>$-7.96 , (-13.7)$</td>
<td>$-8.23 , (-3.60)$</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>--</td>
<td>$-0.003 , (-0.12)$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>$1.52 , (30.0)$</td>
<td>$1.56 , (4.45)$</td>
</tr>
</tbody>
</table>

Notes: *t*-statistics in brackets. Standard Errors are adjusted for long-run variance. The long-run variance of the cointegrating regression residual is estimated using the Barlett window which is approximately equal to $l = 6 \times \text{INT} \left( T^{1/3} \right)$, as proposed in Newey and West (1987).

Critical values:

<table>
<thead>
<tr>
<th></th>
<th>10%</th>
<th>5%</th>
<th>1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_p$</td>
<td>0.231</td>
<td>0.314</td>
<td>0.533</td>
</tr>
<tr>
<td>$C_r$</td>
<td>0.097</td>
<td>0.121</td>
<td>0.184</td>
</tr>
</tbody>
</table>

Notes: $t$-statistics in brackets. Standard Errors are adjusted for long-run variance. The long-run variance of the cointegrating regression residual is estimated using the Barlett window which is approximately equal to $l = 6 \times \text{INT} \left( T^{1/3} \right)$, as proposed in Newey and West (1987).

Table 3. Bai-Perron tests of multiple structural changes in the cointegration relationship

<table>
<thead>
<tr>
<th>Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_t = { \text{HE}_t }$</td>
</tr>
<tr>
<td>$z_t = {1, \text{GDP}_t }$</td>
</tr>
<tr>
<td>$p = 0$</td>
</tr>
<tr>
<td>$h = 10$</td>
</tr>
<tr>
<td>$M = 2$</td>
</tr>
</tbody>
</table>

Notes: $y_t$, $z_t$, $q$, $p$, $h$, and $M$ denote the dependent variable, the explanatory variable allowed to change, the number of regressors, the number of corrections included in the variance-covariance matrix, the minimum number of observations in each segment, and the maximum number of breaks, respectively. $\gamma$, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. The critical values are taken from Bai and Perron (1998), Tables 1 and 2, and from Bai and Perron (2001), Tables 1 and 2.

Notes: $t$-statistics in parentheses.

Next, we will examine the possibility of instabilities in this relationship, extending the previous analysis in order to address the question of whether the long-run or cointegration relationship estimated between GDP and HE is stable over time, or it exhibits some structural break, allowing the instability to occur at a unknown date. To do so, we use the tests of multiple structural changes proposed by Bai and Perron (1998, 2003) in the context of the OLS recursive estimation applied to linear models.

A key feature of the Bai and Perron procedure is that it allows to test for multiple breaks at unknown dates, so that it successively estimates each break point by using a specific-to-general strategy in order to determine consistently the number of breaks. In particular, Bai and Perron (1998) proposed three test statistics to test for multiple breaks: (i) the sup $F_T(k; q)$ test, a sup $F$-type test of the null hypothesis of no structural break $(m = 0)$ vs. the alternative of a fixed (arbitrary) number of breaks $(m = k)$; (ii) the $U_{\text{max}}$ test, a double maximum test of the null hypothesis of no structural break $(m = 0)$ vs. the alternative of an unknown number of breaks given some upper bound $M(1 \leq m < M)$; and (iii) the sup $F_T(l+1/l)$ test, a sequential test of the null hypothesis of $l$ breaks vs. the alternative of $l+1$ breaks.

The results of applying the Bai–Perron tests to the relationship between GDP and HE are shown in Table 3. We apply the procedure with a constant and GDP as regressors (i.e., $z_t = \{1, \text{GDP}_t \}$) and account for potential serial correlation via...
non-parametric adjustments. When implementing the procedure, we allowed up to two breaks. As can be seen, all the sup $F_T(k)$ tests are significant, with $k$ running between one and two, so that at least one break would be present in this relationship. Also, the sup $F_T(2/1)$ is highly significant, and the sequential procedure selects two breaks. Hence, we conclude in favor of the presence of two breaks, whose dates are estimated in 1971 and 1991.

An economic interpretation of these results would be as follows. The year 1971 would witness a first change of regime in the health-care policy, which could be related to the effects of the full implementation of the 1967 Social Security Act. This Act introduces a social insurance system aimed to finance the health care system, which was governed by the social partners and overseen and tightly regulated by the government. An ever greater change in the health policy, however, would have occurred in 1991, following the full implementation of the 1986 General Health Care Act. This year starts the Spanish health care system as an integrated National Health service which is publicly financed and provides nearly universal health care free of charge at the point of use. Notice that, even though both acts are dated some years before, structural changes would only appear once their provisions were fully implemented in practice. In fact, building hospitals, or developing primary and preventive services take some time for their consequences for the whole system to be completely absorbed.

Finally, the estimate of the coefficient measuring the long-run elasticity between HE and GDP, would decrease from 2.02 to 1.27 after 1971, and from 1.27 to 1.02 after 1991. In other words, health would have been characterized as a luxury commodity, even though increasingly less over time.

### III. Conclusions

The objective of this article has been to provide an empirical test of the long-run relationship between HE and GDP using cointegration techniques, for the Spanish economy during the period 1960 to 2001, where the existence of instabilities in the cointegrating or long-run relationships is explicitly tested. To this end, we apply some recent econometric techniques aimed to detect eventual structural changes, allowing the instability to occur at an unknown date.

Our results are consistent with the existence of a cointegrating relationship between HE and GDP, with an estimated long-run income elasticity of health spending of 1.54, so that health would have been characterized as a luxury commodity. Also, we provide evidence against the stability of this relationship, finding two breaks in 1971 to 1972 and 1991 to 1992. The first break was related to the full implementation of the 1967 Social Security System Act, which envisaged the extension of health services, with the introduction of a new public network of modern hospitals and primary and preventive services in the public network, with the subsequent increase in health expenditures. In turn, the second structural change was related to the full implementation of the 1986 General Health Care Act, which envisaged the universal extension of health services.

On the other hand, up to three regimes were detected over the whole sample, with an estimated long-run income elasticity of 2.02 for 1960 to 1971, 1.27 for 1972 to 1991, and 1.02 for 1992–2001. This implies that, although health would have been a luxury commodity over the whole period, it would have been increasingly less along time. Thus, this decrease in income elasticity would imply that the Spanish economy would have assigned larger fixed amounts to health expenditure, whilst dependence on the growth of the economy would have gradually decreased.

Finally, the reduction found in the above elasticity might be related with the gradual increases experienced by coverage, from 50% in 1960 to 100% in 2001. Regarding the government, an increasing coverage of the public health sector can be a cause of public health expenditures growing at a faster rate than GDP. Also, the progressiveness of the tax burden can be a complementary explanation for the reduction in the income elasticity of health spending, since public expenditures (and, in particular, public health expenditures) would grow at a faster rate than GDP.

### Acknowledgements

Vicente Esteve acknowledges the financial support of the Spanish Ministry of Science and Technology, project SEC2002-03651.

---

4 See Bai and Perron (1998) for more details.
References