Health Care Spending and Economic Growth: Armey-Rahn Curve in a Panel of European Economies

Trofimov, Ivan D.

Kolej Yayasan Saad (KYS) Business School

2 September 2020

Health Care Spending and Economic Growth: Armey-Rahn Curve in a Panel of European Economies

Ivan D. Trofimov*

Abstract

The paper examines the Armey-Rahn hypothesis of the inverted U-shaped relationship between public health care expenditure and economic growth in the European economies over the 1995-2018 period. To this end, the aggregate production function (in levels or logarithms) augmented by spending and economic openness terms is estimated. The fixed-effects panel regression, the panel quantile regression with fixed effects, and the panel ARDL models are used for empirical analysis. The paper unequivocally indicates the existence of the Armey-Rahn curve and the negative effects of health care spending on the output (per capita) beyond the optimal spending level. Irrespective of the functional form of the model or the definition of dependent or independent variables, the optimal level was estimated to be smaller than the actual average health care spending level for the period (5.99% of GDP), indicating the over-provision of public health care. Under-provision of public health care was documented for the transition economies in Eastern Europe (that were characterised by comparatively small size of GDP, low per capita output and higher optimal spending levels, economic transition challenges, and lagging health care spending, in addition to indivisibilities of the public health investment).

JEL Classification: C33, H51, N34
Keywords: Armey-Rahn curve, health care spending, growth, government size, Europe

Introduction

The cost escalation of health care and the expansion of health expenditure beyond the optimal level is a phenomenon that has been observed across the economies and examined extensively in health economics. The literature attributes this development to a number of interrelated factors: the technological change in the medical field (that results in the introduction of new and more costly medical technologies); the demographic changes (population ageing in the developed world and the ongoing demographic boom in certain developing countries, necessitating, albeit for different reasons, the expansion of health care services); the demands and perceptions on the part of users of medical services (the view of health care as basic and universal right, and hence the requirement of universal or expanded coverage, wider range and better quality of medical services); political-economic and bureaucratic pressures (associated with the operation of a complex network of government bodies, service providers, insurance and pharmaceutical companies, stakeholders and vested interests of various sorts); the objective medical challenges (the chronic diseases in the developed world, and new infectious diseases, such as novel corona-virus); to name a few (Finkelstein, 2007; Smith et al, 2009; Jenkner, Leive, 2010: 2).[i]

The cost and expenditure challenges are well manifested in the developed economies, where, according to OECD survey (OECD, 2015), public funds account for around three-quarters of health spending and could increase from around 6% of GDP in 2015 to almost 9% of GDP in 2030, and as much as 14% by 2060 unless governments contains cost growth.

The definitive answer as to the over-, under- or optimal provision of the health care expenditure (and government expenditure in general) requires consideration of possible non-linearities in the expenditure-economic growth relationship. In this respect, the Armey-Rahn hypothesis of the positive growth or output effects of government expenditure up to a certain point and conversely its negative effects beyond this point (represented by the 'Armey-Rahn curve' and the 'Scully point' on this curve) serves as a useful tool for the empirical research and a measure to be used for effective cost containment policies (Rahn, Fox, 1996; Scully, 2000).

* Kolej Yayasan Saad (KYS) Business School, Melaka, Malaysia. E-mail: ivan.trofimov1@gmail.com

This paper considers an empirical issue of the possible presence of Armey-Rahn curve in health spending in 28 European economies over the period of 1995-2018, specifically to establish 1). whether the relationship between health spending and the level and growth rate of GDP is non-linear (as postulated by the Armey-Rahn hypothesis); and 2) whether health spending is over- or under-provided relative to its optimal level.

To this end we introduce the following novelties. Firstly, we include health expenditure (as proportion of GDP) and its square term as regressors in the aggregate production function for the economy together with capital and labour inputs and other relevant variables, an approach that was previously adopted for the analysis of the effects of other expenditure types (Goel et al, 2008, for the analysis of economic impacts of R&D spending in the US; Zhang and Li, 2008, for the identification of the optimal size of rural fiscal expenditure in China). Secondly, we experiment with alternative specifications: basic that includes solely health expenditure, and augmented that includes labour, capital, and a measure of economic openness. Thirdly, given the mixed order of integration of variables, the likely lags in the effect of expenditure on economic outcomes, and unobserved country heterogeneity that can cause health expenditure having different effects along the conditional distribution of GDP (or its growth rate), we employ three econometric techniques: panel OLS model with fixed or random effects, panel quantile model with fixed effects, as well as panel autoregressive distributed lags (ARDL) model.

The remainder of the paper is structured as follows. Section 2 provides a brief review of literature related to Wagner law, its limitations, Armey-Rahn curve hypothesis as a way to address these limitations, as well as empirical research on the topic. Section 3 presents model specification, data sources as well as econometric methodology. Section 4 submits to empirical testing of the Armey curve, determination of the optimal size of the public expenditure and discussion of policy implications, while the last section provides concluding remarks and discusses the possible directions for the future research.

Literature review

We review two complementary streams of literature: 1). Studies that examine the theoretical foundations of 'Armey-Rahn curve' hypothesis as well as the empirical studies that attempt to provide empirical evidence to support this hypothesis; and 2). The research on the relationships between health care expenditure and economic variables (including GDP and economic growth as dependent variables).

Armey-Rahn curve

Wagner Law, formulated by A. Wagner in 1893, and empirically tested on numerous occasions and in diverse settings postulates the growth of government spending in excess of growth in economic output, both when absolute or relative magnitudes are concerned. If Y_i is the level of economic output, G_i is the level of government expenditure in country i, the Wagner Law relationship between the two variables would be represented as $G_i = mY_i^{\beta}$, where $m > 0$, $\beta > 1$ and $\beta = (Y_i / G_i)(dG_i / dY_i)$, i.e. elasticity of government spending with respect to output is greater than one (Balatsky, 2012: 41). A number of mechanisms behind Wagner law were identified: the expansion of administrative and protection activities by the state in lieu of such activities conducted by the private sector; the growth of cultural and welfare functions of the state; the functioning of government-owned natural monopolies; state expansion driven by bureaucratic pressures and preferences; lobbying and voting that lead to higher (re-)distribution through the state (Niskanen, 1971; Persson, Tabellini, 199).

The infinite operation of Wagner Law, however, is impossible on theoretical and logical grounds, as it would imply disproportionate growth of the public finance relative to the total economy, the growing appropriation of income by the government, and eventually, socialisation of the economy, the process that never reaches apex in modern market, mixed or even centrally-planned economies (Balatsky, 2012: 41-46). The operation of Wagner Law would also be constrained by the political and economic context of a given economy, with positive elasticity of government spending with respect to output likely observed in the less developed economies or in economies with smaller degree of state intervention, and negative elasticity in the more developed or the ones with more interventionist state (Yavas, 1998; Tanzi, Schuknecht, 1997; Durevall, Henrekson, 2011). Likewise, the negative effects of public sector expansion on the economy would become prominent at some point, e.g. expansion of public sector that increases tax burden and provides disincentive to invest; the decrease in total productivity due to proliferation of less efficient public enterprises; the comprehensive planning and management of the economy by the government, instead of law and order provision and guaranteeing of property rights; production of private goods, as opposed of provision of public goods and externalities' correction (Baumol, 1993; Vedder, Gallaway, 1998; Magazzino, Forte, 2010: 4).

The U-shaped Armey curve reflects these processes that offset the initially positive effects of growing public expenditure: the elasticity of government spending (as percentage of output) with respect to output growth rate is postulated to be negative to the right of inflection point on Armey curve ('Scully point'), against Wagner law premises, and positive to the left of it, in line with Wagner law. Mathematically, if $g = G/Y$ is a government expenditure as percentage of GDP, and $\theta = \Delta Y/Y$ is output growth rate, the 'Scully point' is identified when $\partial\theta/\partial g = 0$ and $\partial^2\theta/\partial g^2 < 0$ (Balatsky, 2012: 47).

The empirical validation of Armey curve was conducted according to three alternative methods.[ii] Firstly, the (Granger-) causality or path analysis would investigate feed-forward and feedback relationships between output and public expenditure, with causality running from GDP growth to government spending growth (and positive relationship between the two variables) corresponding to Wagner law, while the reverse causality running from public expenditure to GDP growth indicating indirectly the Armey curve relationship (positive relationship superseded by the negative). The studies by Wu et al (2010), and Facchini, Melki (2011) fall into this methodological category. The second approach ('direct' estimation of Armey curve) would represent output (in levels or in growth terms) as a function of the level or growth rate of expenditure, augmented by squared expenditure (to capture non-linearities in growth-expenditure relationship) and control variables to incorporate the influence of other factors on output. The presence of Armey curve relationship and the optimal size of the government expenditure would then be determined from the signs and values of coefficients of the expenditure terms. The studies by Facchini, Melki (2011), Herath (2012), among others, followed this approach. Thirdly, the presence of Armey curve can be established following the estimation of non-linear Cobb-Douglas production function in logs with output as a function of total factor productivity, government expenditure/GDP and tax revenue/GDP ratios, as proposed by Scully (1994). Under the balanced budget assumption and following differentiation of a production function with respect to expenditure, the optimal tax rate and government size are estimated and the presence of Armey curve is established. This method was followed by Chao and Grubel (1998), Schoeman and van Heerden (2009), El Husseiny (2018), among others.

The presence and significance of the Armey curve was established in a large number of studies, however, there is less agreement regarding the optimal size of government. Across the economies, the optimal size of the government (total government expenditure) varied substantially: in a range of 37%-42% of GDP in a panel of 12 European economies (Pevcin, 2004); 41.2% of GDP in 23 OECD economies in 1999 (De Witte, Moesen, 2010); 35.4-43.5% of GDP in a sample of 27 EU economies during 1970-2009 (Magazzino, Forte, 2010); 25% of GDP for 28 EU economies over the same period (Chobanov, Mladenova, 2009); 16.2% and 16.9% in a sample of 21 low income and 11 low-middle income economies (Hajamini, Ali Falahi, 2014); 23% across a large sample of 118 economies (Karras, 1996); and 40% in a panel of low income economies (Davies, 2009). For individual economies, the estimates of optimal government size were as follows: 44% in Israel and 35% in Tunisia, based on non-military public consumption during 1968-1997 period (Handoussa, Reiffers, 2003); 27% in Sri Lanka during 1959-2009 period (Herath, 2012); 30.5%-31.2% in Egypt during 1981-2015 period (El Husseiny, 2018); 28.0% in Russia and 38.7% in Sweden during 1990-2007 period (Balatsky, 2012); 23%

in Italy in 1986-1998 (Magazzino, 2008); 21% in Bulgaria in 1990-2004 (Mavrov, 2007); 22.8%, 7.3% and 14.9% in Taiwan based on total government, investment and consumption government expenditure respectively (Chen, Lee, 2005); 29%-30% in France, with a peak in Armey curve reached in the aftermath of WWII (Facchini, Melki, 2011); around 20% in the US (Peden, 1991; Scully, 1994; Vedder, Gallaway, 1998); 23% in New Zealand (Scully, 1994); and a low 10.8-15.9% across South Korea, Singapore, Taiwan, Thailand and Malaysia (Chiou-Wei et al, 2010).

As far as the relationship between the specific types of public expenditure and the economic output (growth) are concerned, the empirical literature has been scant. Two studies of the Wagner Law and Armey curve at the sub-federal level in the US are notable (Vedeer and Gallaway, 1998, and Miller, 2008). Both use disaggregated expenditure data and focus on the 1947-1997 and 1976-2005 periods respectively. Both studies confirmed the existence of Armey curve in certain (but not all) states and likewise suggested the existence of state-specific and expenditure-specific optimum expenditure levels, thereby preventing from making policy or theoretical generalisations based on Armey curve. The former study indicated optimum levels that were few percentage points higher than the latter study. Another study by Zhang and Li (2008), while not looking explicitly at Armey curve (or non-linearities in expenditure-output pair) considered the optimum levels of rural fiscal expenditure in China during the 1980-2005 period: setting the aggregate production function for the total economy, and estimating the marginal product and output elasticity of the rural expenditure, the authors determine the optimal expenditure level at 13.2% of rural GDP, well above the average expenditure levels for the period.

Health care expenditure and GDP

As far as the relationship between public health expenditure (and its optimal level) and economic outcomes is concerned, the literature focused on the following issues. Firstly, the debate has been underway as to the nature of health, i.e. whether it is a necessity or luxury good, the latter suggesting that growth of spending on health should exceed the overall economic and income growth, and in turn giving rise to the discussion of the overspending and cost containment (Gerdtham et al, 1992; Getzen, 2000). Secondly, the empirical studies examined the determinants of health spending, confirming the important influence of the country's GDP level on the the level of spending (both cointegration and causality between the two, Clemente et al, 2004), but also highlighting the importance of demographic, institutional, fiscal and macroeconomic factors (Hitiris, Posnett, 1992; Zwiefel et al, 1999; Di Matteo, 2003). Thirdly, the literature examined the relationship between private and public health spending (Yazici, Kaestner, 1998; Ying, Chang, 2019), and generally gave support to the crowding out hypothesis (private spending displaced by the public)

The studies that specifically examined the optimal level of health expenditure have been scant. Chang and Ying (2006) considered health expenditure levels in the 15 OECD countries during the 1980-1998 period and estimated the Solow growth model with health expenditure, representing investment in human capital, as one of the regressors in the aggregate production function (alongside labour and physical capital). The optimal level of expenditure was conceptualised as a steady-state, where the growth rate of per unit of labour health (human) capital is zero, but the level of health (human) grows at the same rate as population. The model included control variables to account for demographic state of the respective country, nature of the health care system, and economic factors. The authors demonstrate an increasing trend in both optimal and actual health expenditure, and overspending in all OECD economies in question. The gap between optimal and actual spending attenuated over time in most economies (excepting Spain and Greece), representing the effectiveness of the cost containment measures. Other findings were consistent with previous literature and theory (negative relationship between mortality, population growth, environmental pollution and unemployment on one hand and spending on the other; and conversely, the positive effects of the existing health capital stock on health spending).

A more recent study by Wang (2015) adopts Grossman model of demand for health, where health is defined as a durable capital stock that produces an output of healthy time (Grossman, 1972: 246), and estimates an aggregate production function for OECD economies over 1990-2009 period with health expenditure as regressor and GDP per capita as dependent variable. The generalised method of moments estimation indicates the optimal level of total health expenditure (including private spending) at 7.55% of GDP compared to the actual spending share of 5.48%, suggesting the under-provision of health care (the result that is counter-intuitive in light of previous debate in health economics).

Methodology

Model

In a general form, the relevant equation specifications for the empirical analysis are derived from the Cobb-Douglas production function augmented by government expenditure and exports or openness measures (Anaman, 2004; Kustepeli, 2005: 4; Hok et al, 2014: 36-37):

$$GDP_{it} = \exp\left(\beta_0 + \beta_1 HX_{it} + \beta_2 HX_{it}^2\right)\left(L_{it}\right)^{\beta_3}\left(K_{it}\right)^{\beta_4}\left(OP_{it}\right)^{\beta_5} \quad (1)$$

,where HX_{it} and HX_{it}^2 are public health care expenditure and its squared term, L_{it} and K_{it} are labour and capital inputs, OP_{it} is the openness indicator for country i at period t, and $\exp$ is exponential operator.

In logarithmic form the function is represented as follows and is hence amenable for empirical analysis:

$$\ln GDP_{it} = \beta_0 + \beta_1 \ln HX_{it} + \beta_2 \ln HX_{it}^2 + \beta_3 \ln L_{it} + \beta_4 \ln K_{it} + \beta_5 OP_{it} + \varepsilon_{it} \quad (2)$$

,where all variables are defined as above, and ε_t is a normally distributed residual term. The Armey-Rahn hypothesis is deemed confirmed when quadratic relationship between expenditure and GDP is deemed confirmed, i.e. when $\beta_1 > 0$ and $\beta_2 < 0$.

This specification (*Specification 1*) has been adopted in the empirical studies by Facchini and Melki (2011) and Cetin (2017). The second specification (*Specification 2),* used by Vedder, Gallaway (1994), Miller (2008), and Hok et al (2014) as an alternative estimation method, regresses the logarithm of the real GDP against the health care spending variables without logarithms (i.e. variables in levels). As part of robustness check, we also tried additional three specifications (*Specifications 3, 4 and 5*): with GDP growth as dependent variable; with lags of the dependent and independent variables to account for dynamic effects; and with I(1) order variables in the growth rates representation.

The optimal levels of health care expenditure are determined for Specifications 1 and 2 as:

$$HX^* = \exp\left(\beta_1 / |2\beta_2|\right) \quad (3)$$

and

$$HX^* = -\beta_1 / 2\beta_2 \quad (4)$$

Data

The panel dataset used in this study covered the period from 1995 to 2018 and included 28 countries in the European Union (EU) and the European Free Trade Association (EFTA) area: Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, and the UK. Thus, it consists of the original members of the EU and the new members (formerly socialist economies), economies with different levels of

GDP per capita and standard of living, albeit generally having publicly funded health care or with public health care co-existing alongside private one (with the exception of Switzerland).

The general government health expenditure by function data (in millions of Euro or as percent of GDP) was extracted from the Eurostat database, with respective government functions following *Classification of the Function of the Government (COFOG), Version 1999* (OECD, 2011: 194-5). In COFOG, the public expenditure on health is defined to include (at the second level of the classification) expenditure on medical products, appliances and equipment, outpatient, hospital and public health services, on health-related research and development (R&D), , as well as other health expenditure not elsewhere classified. The GDP at market prices and the chain-linked volume measures of GDP (with the base of the chain-linked volume index set at 2010) were likewise obtained from the Eurostat.

The capital stock data was taken from the Annual Macroeconomic Database of the European Commission (AMECO): net capital stock for the total economy at 2015 prices in Euros (OKND variable code). The values were recalculated to 2010 constant prices to achieve consistency with other variables, using price deflator for the gross fixed capital formation in the total economy (PIGT variable code). Employment figures for the total economy (with original source in the countries' national accounts) were also obtained from AMECO (NETN variable code). Trade openness was estimated from the GDP, exports and imports at constant prices as a ratio of the sum of exports and imports to GDP expressed in percentage terms.

Econometric method

Panel unit root tests were first conducted to identify possible unit root in the series (Im-Pesaran-Shin/IPS, Levin-Lin-Chu/LLC, Breitung, ADF-Fisher χ^2, PP-Fisher χ^2 and cross-sectionally augmented IPS / CIPS tests)[iii], as well as the cross-sectional dependence tests (Breusch-Pagan, 1980; Pesaran, 2004) to identify correlation across space (cross-sectional dependence). The absence of series integrated of order one, I(1), indicates the appropriateness of standard panel least square techniques (pooled OLS, fixed or random effect models); in contrast, their presence would suggest using panel VECM, panel VAR in first differences, or, in the case of mixed orders of integration, panel ARDL models. As shown in the next section, there was strong evidence of stationarity in levels for all series in question, particularly for variables in levels and when the test that address cross-sectional dependency is used.

We waive the assumption of the absence of systematic differences among the economies, and use fixed- and random-effects panel models (instead of pooled OLS). The fixed-effect model assumes different intercept terms (β_i) for individual panels and homogeneous coefficients of the regressors, as follows:

$$\ln GDP_{it} = \beta_i + \beta X_{it} + \varepsilon_{it} \tag{5}$$

,where X_{it} is the vector of independent variables, $i = 1,\ldots,N$, $t = 1,\ldots,T$, β_i is unobserved and time-invariant individual effect, and ε_{it} is an error term.

The random-effect model includes individual effects as random variables:

$$\ln GDP_{it} = \beta_0 + \beta X_{it} + \nu_{it} \tag{6}$$

,where $\nu_{it} = \varepsilon_{it} + \beta_i$

Hausman test is used to select fixed-effects and random-effects models, and Driscoll-Kray standard errors are obtained, if there is evidence of serial autocorrelation and cross-sectional dependence, as well as heteroskedasticity (Driscoll, Kraay, 1998).

We also apply panel quantile regression (Koenker, Bassett, 1978; Koenker, 2004) that addresses the simplifications of the standard econometric techniques (e.g. least squares method) and estimates the conditional median or quantiles of the dependent variable (as opposed to conditional mean of the dependent variable in the least-squares models). The quantile regressions are robust to outliers, heteroskedasticity and skewness in the dependent variable (Li et al, 2020: 10), describe the entire conditional distribution of the dependent variable, allow for non-identical distribution of the error terms across the conditional distribution, and varying slope parameters at different quantiles (Buchinsky, 1994). Its use is justified by the fact that the influence of health expenditure on GDP and its growth may vary depending on the type of health care financing (private versus public), the levels of health care expenditure and other heterogeneity in the relationship among the variables.

We use method of moments quantile regression (MM-QR) for panel data method recently proposed by Machado and Santos Silva (2019), that includes the fixed effects and that is applicable to non-linear functional forms of the models. It overcomes the limitation of the standard quantile regression that does not account for unobserved heterogeneity across the individual panels. It allows the individual effects to affect the entire distribution rather than shifting means as in the standard quantile regressions, and identifies the conditional heterogeneous covariance effects of independent variables (Ike et al, 2020: 5).

Quantiles are estimated from the estimates of the conditional mean and conditional scale functions (Li et al, 2020: 10). With GDP_{it} in levels or logarithms representing dependent variable for country i at time period t, whose distribution is conditional on X_{it}, a vector of covariates (HX_{it}, HX_{it}^2, L_{it}, K_{it}, and OP_{it} of country i in period t), the conditional quantiles $Q_{GDP}(\tau|Z)$ for a location-scale model is given as (Ike et al, 2020: 5; Li et al, 2020: 10):

$$GDP_{it} = \alpha_i + X_{it}^{'}\beta + (\delta_i + Z_{it}^{'}\gamma)U_{it} \tag{7}$$

,where $P\{\sigma(\delta_i + Z_{it}^{'}\gamma) > 0\} = 1$, parameters $(\alpha_i, \delta_i), i = 1, \ldots, n$ represent fixed effects, X_{it} is i.i.d. for any i and is time-independent, U_{it} is i.i.d. for any i or period t, orthogonal to X_{it} and satisfying moment conditions $E(U) = 0$ and $E(|U|) = 1$. Z_{it} is a vector of differentiable transformations of the components of X.

The panel quantile function for quantile τ is therefore defined as:

$$Q_{GDP}(\tau|X_{it}) = (\alpha_i + \delta_i q(\tau)) + X_{it}^{'}\beta + Z_{it}^{'}\gamma q(\tau) \tag{8}$$

,where $Q_{GDP}(\tau|X_{it})$ is a quantile distribution of the GDP_{it} conditional on the location of independent variables. Quantile τ fixed effect for panel i is given as $\alpha_i(\tau) \equiv \alpha_i + \delta_i q(\tau)$. The sample quantile $q(\tau)$ is estimated from optimization problem (Haylock, 2020: 16):

$$\min \sum_i \sum_t \rho_\tau \left(\hat{R}_{it} - (\hat{\delta} + Z_{it}^{i}\hat{\gamma})q\right) \tag{9}$$

,where $\hat{R}_{it} = GDP_{it} - \hat{\alpha}_i - X_{it}^{'}\hat{\beta}$, $\hat{R}_{it}$ being the residual from OLS regression, $\hat{\alpha}_i$ being the location shift parameter, and $\hat{\gamma}$ being the scale parameter that represents effect heterogeneity and that is obtained from regression of of the time-demeaned absolute value of residuals on variables Z.

As a last step, we employed panel autoregressive distributed lags (panel ARDL) model that is suitable for the data with a mixed order of integration (as demonstrated in the following section), and that allows estimation of the long- and short-run coefficients within a single framework. In contrast to the mean group estimator (Pesaran and Smith, 1995) that averages separate estimates for each group in the panel and assumes intercepts, slope coefficients and error variances to vary across the groups, we used the pooled mean group (PMG) estimator that relaxes such assumption (Pesaran et al, 1999), and imposes homogeneity restriction on the long-run parameters, while allowing short-run specifications to differ across the countries. Due to the short-run parameter heterogeneity, PMG allows more reliable and efficient estimates of the long-run coefficients as well as faster convergence to the long-run equilibrium compared to the mean group estimator (Pesaran et al, 1999; Gemmell et al, 2014: 10).

The ARDL model with (p,q) lag orders is given as:

$$GDP_{it} = \sum_{j=1}^{p} \phi_{ij} GDP_{i,t-j} + \sum_{j=0}^{q} \delta_{ij}^{1} HX_{i,t-j} + \sum_{j=0}^{q} \delta_{ij}^{2} HX_{i,t-j}^{2} + \sum_{j=0}^{q} \delta_{ij}^{3} L_{i,t-j} + \sum_{j=0}^{q} \delta_{ij}^{4} K_{i,t-j} + \tag{10}$$

$$+ \sum_{j=0}^{q+1} \delta_{ij}^{5} OP_{i,t-j} + \mu_i + \varepsilon_{it}$$

,where $j = 1,2,\ldots,N$, $t = 1,2,\ldots,T$, δ_{it} are $k \times 1$ coefficient vectors, and μ_i is the group-specific effect. Other specifications are estimated equivalently.

The error-correction representation of Equation (10) is given as:

$$\Delta GDP_{it} = \alpha_i \left(GDP_{it-1} - \beta_{oi} - \beta_{1i} HX_{i1} - \beta_2 HX_{it}^2 - \beta_3 L_{it} - \beta_4 K_{it} - \beta_5 OP_{it} \right) + \tag{11}$$

$$+ \sum_{j=1}^{p-1} \theta_{ij} \Delta GDP_{it-j} + \sum_{j=0}^{q-1} \psi_{ij}^{1} \Delta HX_{it-j} + \sum_{j=0}^{q} \psi_{ij}^{2} \Delta HX_{it-j}^{2} + \sum_{j=0}^{q} \psi_{ij}^{3} L_{it-j} + \sum_{j=0}^{q} \psi_{ij}^{4} K_{it-j} + \sum_{j=0}^{q-1} \psi_{ij}^{5} OP_{it-j} + \varepsilon_{it}$$

,where $\alpha_i = -(1 - \sum_{j=1}^{p} \phi_{ij})$, $\beta_{0i} = \frac{\mu_i}{-\alpha_i}$, $\beta_{1i} = \frac{\sum_{j=0}^{q} \delta_{ij}^{1}}{-\alpha_i}$, $\beta_{2i} = \frac{\sum_{j=0}^{q} \delta_{ij}^{2}}{-\alpha_i}$, $\beta_{3i} = \frac{\sum_{j=0}^{q+1} \delta_{ij}^{3}}{-\alpha_i}$.

Empirical results

Baseline model

As a first step, we examined the descriptive statistics of the series in logarithms (Table 1). All variables had positive mean and median. The logarithm of the real capital stock had the highest standard deviation, while the standard deviation of the logarithms of openness and the health care spending to GDP ratio was the lowest. The null hypothesis of the normal distribution of the data was rejected for the logarithms of real GDP, capital stock, openness, and the (squared) health care spending as proportion of GDP at 1% significance level (and for the logarithm of employment at 5% level). All variables were approximately symmetric around the mean with the exception of health care spending as proportion of GDP variables that were skewed to the left. The real GDP, employment and real capital stock were platykurtic, while other variables were leptokurtic. For the estimation we also use health care spending variable without logarithmic transformation, as well as the growth rates of real GDP, capital stock, employment and openness. These variables likewise had positive mean and median and the growth rates had the highest standard deviation. The growth rate of real capital stock was positively skewed to the right, the growth rate of employment and the health care spending as proportion of GDP were skewed to the left, while other variables were symmetric. The respective growth rate variables had large positive excess kurtosis, while health care spending variables were platykurtic. The null hypothesis of the normal distribution was rejected for all variables.

As shown in Table 2 in the Appendix, the panel average public health care expenditure stood at 5.99% of GDP, with a number of economies having much higher spending (the top five countries with the highest spending as percentage of GDP being Austria, the UK, Denmark, Czech Republic and France). The economies with below-average figure include Bulgaria, Cyprus, Estonia, Hungary, Latvia, Lithuania, Luxembourg, Poland, Romania and Spain, while Switzerland was a clear outlier with the level of public health care spending of 1.87% of GDP.

Table 1. Descriptive statistics

Statistic	Mean	Median	Max	Min	SD	Skewness	Kurtosis	J-B prob.	Obs
Log(GDP)	12.072	12.155	14.905	8.964	1.494	-0.038	2.181	0.000	672
Log(K)	6.232	6.397	9.065	2.959	1.638	-0.140	2.101	0.000	672
Log(L)	8.262	8.294	10.708	5.128	1.290	-0.180	2.708	0.049	672
Log(OP)	1.944	1.930	2.612	1.403	0.207	0.452	3.335	0.000	672
HX	5.987	6.442	9.447	1.574	1.680	-0.693	2.717	0.000	672
HX^2	38.667	41.494	89.248	2.478	18.431	-0.159	2.076	0.000	672
Log(HX)	1.737	1.863	2.246	0.454	0.351	-1.488	5.051	0.000	672
Log(HX^2)	3.142	3.470	5.043	0.206	1.051	-0.899	3.102	0.000	672
GDPGR	2.594	2.639	25.163	-14.814	3.404	-0.529	9.869	0.000	644
KGR	0.534	0.345	79.559	-33.482	5.352	4.728	79.889	0.000	644
LGR	0.735	0.929	6.586	-13.911	2.103	-1.570	10.285	0.000	644
OPGR	3.052	2.773	23.158	-17.678	5.393	-0.044	5.740	0.000	644

Note. The values of GDPGR, KGR, LGR and OPGR are expressed as percentages.

To account for cross-sectional dependence in panels, we conducted Breusch-Pagan (1980) and Pesaran (2004) tests of cross-sectional dependence, given that T and N are not substantially different in the panel (De Hoyos, Sarafidis, 2006: 483-4). The null hypothesis of the absence of cross-sectional dependence (in both tests) is rejected for all variables (Table 3).

Table 3. Cross-sectional dependence tests' results

Variable	Breusch-Pagan LM		Pesaran CD	
	Stat.	p-value	Stat.	p-value
Log(GDP)	7544.669	(0.000)	85.585	(0.000)
Log(K)	2516.427	(0.000)	6.880	(0.000)
Log(L)	4159.721	(0.000)	38.350	(0.000)
Log(OP)	7274.399	(0.000)	82.754	(0.000)
HX	1972.337	(0.000)	17.252	(0.000)
HX^2	1945.365	(0.000)	17.079	(0.000)
Log(HX)	1983.205	(0.000)	17.344	(0.000)
Log(HX^2)	1977.960	(0.000)	17.281	(0.000)
GDPGR	3176.188	(0.000)	52.519	(0.000)
KGR	901.553	(0.000)	11.326	(0.000)
LGR	1429.180	(0.000)	30.572	(0.000)
OPGR	2357.202	(0.000)	43.093	(0.000)

Note. Cross-section means were removed during computation of correlations. p-values are indicated in the parentheses.

The panel unit root tests (Table 4 in the Appendix) were implemented for the levels and the first differences of the series, including constant (Model C) or constant plus trend (Model CT) deterministic components. The real GDP, capital stock, employment, openness and health care to GDP ratio were also represented in logarithmic form. The logarithm of health care spending to GDP ratio (in levels) was stationary in both Models C and CT (the similar results are also obtained for the variable without logarithms). The squared term of the health care spending to GDP ratio was likewise stationary in both models (with or without logarithms). The logarithm of the real GDP (in levels) contained unit root according to IPS and ADF-Fisher χ^2 tests in Model C and ADF-Fisher χ^2 and PP-Fisher χ^2 tests in Model CT. The logarithm of the level of real capital stock was (trend) stationary in Models C and CT according to all tests, except for the Breitung test. The logarithm of employment (in levels) followed unit root behaviour under all tests in Model C, and under PP-Fisher χ^2 test in Model CT, while the first difference of the variable was (trend) stationary. The tests' results for the logarithm of openness are conflicting: in Model C, unit root is identified by the IPS and ADF-Fisher χ^2 tests, while in Model CT, the unit root is indicated only by Breitung test. The first differences of all variables in question were (trend) stationary. The tests performed on the levels of the growth rates of GDP, capital stock, employment, and openness all indicate stationarity around the mean. The Pesaran CIPS (2007) test was performed given that panels have common factor structure. For the variables in levels, the (trend) stationarity was indicated in all instances, except for the logarithm of real capital stock and openness, and health care spending variables (with or without logarithmic transformation) in Model CT. The first differences of the variables were all (trend) stationary. Thus, we conclude that the variables were a combination of I(0) and I(1) orders and none of the variables were integrated of order two.

Table 5 present the estimates of the panel regression with the two-way fixed effects, where the health care as percentage of GDP (and the square term of the variable) are included in either logarithms or levels (*Specifications 1 and 2*). Three equations for each specification are estimated: trivariate, with health care spending variables as regressors; multivariate with capital and labour inputs; and multivariate with capital, labour, and openness.

Table 5. Panel regression with fixed effects estimates

Variable	Specification 1			Specification 2		
	Equation 1	Equation 2	Equation 3	Equation 4	Equation 5	Equation 6
HX	0.448	0.449	0.498	0.093	0.078	0.082
	(0.005)	(0.008)	(0.003)	(0.006)	(0.059)	(0.080)
HX^2	-0.164	-0.170	-0.167	-0.010	-0.009	-0.008
	(0.001)	(0.000)	(0.001)	(0.001)	(0.009)	(0.034)
L		0.197	0.417		0.213	0.422
		(0.097)	(0.012)		(0.063)	(0.010)
K		0.135	0.141		0.115	0.123
		(0.078)	(0.032)		(0.125)	(0.060)
OP			0.723			0.700
			(0.001)			(0.002)
Constant	12.153	9.605	6.209	12.227	9.717	6.489
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
Period-effect	2147.380	1373.460	158.160	2222.910	1421.760	167.810
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Serial correlation	798.870	305.337	83.492	906.658	333.649	84.714
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Heteroskedasticity	11920.450	20899.510	16049.110	9343.950	14422.060	12996.040
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal size (%)	3.930	3.747	4.466	4.855	4.511	5.122

Note. p-values are indicated in the parentheses.

The Hausman test is used to select between fixed and random effects, while the joint significance of variables test is performed to justify the inclusion of the period effect in addition to the cross-section effect. Wooldridge (2002: 282-3) and the modified Wald (Greene, 2000: 598) tests are used to detect serial correlation and groupwise heteroskedasticity. The robust Driskoll-Kraay standard errors are reported instead of conventional standard errors.

Columns 1 to 3 include equations with public health care variables in logarithms (with or without additional regressors). All variables are significant at the 5% level of significance, except for the logarithms of real capital stock and employment in the second equation (column) that are significant at the 10% level. The expected signs of coefficients are positive for all variables except the squared term of the public health spending (as proportion of GDP), indicating the positive effects of capital and labour inputs as well as openness on GDP and the presence of Armey-Rahn curve. The optimal level of public health expenditure to GDP ratio ranges from 3.74% to 4.47% of GDP, suggesting that during the study period the public health expenditure was overprovided, when the whole panel is considered. The diagnostic tests justified the use of a model with two-way fixed effects and

Driskoll-Kraay standard errors. Columns 4 to 6 contain *Specification 2* estimates, where public health care spending variables enter the equation without logarithmic transformation. The results are similar in terms of signs and significance of the coefficients (except for the insignificance of the capital stock variable in one of the equations) and the outcomes of the diagnostic tests. The optimal level of the spending is somewhat higher, with public health care to GDP ratio ranging from 4.51% to 5.12% of GDP.

To account for the effect of health care variables on GDP throughout the conditional distribution, we considered the panel quantile model with fixed effects (Table 6). We provide the estimates for the most encompassing equation in each specification. In *Specification 1*, all variables (except the squared term of the health care spending to GDP ratio) were positive in each quantile, as expected. The logarithms of employment and openness were significant (at the 5% level) in each quantile, while the logarithm of the real capital stock was significant in quantiles 3 to 9 (at the 5% level). The significance of both health care spending variables was established in quantiles 2 to 9, thereby confirming the presence of the Armey-Rahn curve relationship. Comparing the estimates with the average spending level for the period (5.99% of GDP) we conclude that the public health expenditure was overprovided for quantiles 1 to 4 economies and underprovided for higher quantiles (quantiles 5 to 9). *Specification 2* yields similar results: the employment and openness variables were significant in each quantile, the capital stock was significant in quantiles 5 to 9, and the expected signs of coefficients were obtained. At the 10% significance level, the Armey-Rahn curve was identified only in quantiles 5 to 7. The health care expenditure was overprovided in quantiles 5 and 6, and underprovided in quantile 7 (albeit the level of provision was very close to optimal level).

Table 6. Panel quantile regression with fixed effects estimates

Specification 1									
Variable	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Log(HX)	0.770	0.718	0.676	0.634	0.599	0.566	0.542	0.515	0.492
	(0.103)	(0.050)	(0.019)	(0.005)	(0.002)	(0.002)	(0.007)	(0.025)	(0.064)
Log(HX^2)	-0.192	-0.185	-0.180	-0.174	-0.170	-0.166	-0.162	-0.159	-0.156
	(0.154)	(0.077)	(0.029)	(0.006)	(0.002)	(0.002)	(0.004)	(0.015)	(0.040)
Log(L)	0.657	0.703	0.741	0.778	0.809	0.838	0.860	0.883	0.904
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log(K)	0.128	0.130	0.132	0.134	0.135	0.137	0.138	0.139	0.140
	(0.262)	(0.141)	(0.057)	(0.013)	(0.003)	(0.002)	(0.004)	(0.012)	(0.029)
Log(OP)	1.533	1.474	1.426	1.380	1.340	1.304	1.276	1.246	1.220
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal size	7.420	6.939	6.540	6.159	5.831	5.530	5.302	5.057	4.840

Note. p-values are indicated in the parentheses.

Specification 2									
Variable	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
HX	0.146	0.131	0.118	0.106	0.097	0.087	0.079	0.071	0.063
	(0.338)	(0.275)	(0.206)	(0.128)	(0.071)	(0.038)	(0.045)	(0.115)	(0.252)
HX^2	-0.010	-0.010	-0.009	-0.008	-0.008	-0.007	-0.007	-0.006	-0.006
	(0.361)	(0.287)	(0.206)	(0.117)	(0.056)	(0.023)	(0.024)	(0.066)	(0.163)
L	0.678	0.720	0.757	0.792	0.819	0.848	0.869	0.893	0.914
	(0.009)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
K	0.113	0.116	0.119	0.122	0.124	0.127	0.128	0.130	0.132
	(0.563)	(0.450)	(0.320)	(0.172)	(0.070)	(0.017)	(0.011)	(0.024)	(0.062)
OP	1.537	1.480	1.432	1.386	1.349	1.311	1.283	1.251	1.222
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal size	N/A	N/A	N/A	N/A	6.275	6.068	5.900	N/A	N/A

Note. p-values are indicated in the parentheses.

Robustness checks

We performed several robustness checks. Firstly, in line with studies by Herath (2012), El Husseiny (2018), and Forte and Magazzino (2018) who focused on the effect of health expenditure / real GDP ratio and its square term, alongside the above variables, on the growth rate of real GDP, we examined the alternative specification (*Specification 3*) with the first difference of the logarithm of real GDP as a dependent variable (Table 7). The positive coefficients were obtained for all variables except real capital stock and the squared term of the health care spending variable. All coefficients are significant (except for health care spending variables in one of the equations), when conventional OLS standard errors are used, whereas the majority are not significant when Driscoll-Kraay robust errors are used. The Armey-Rahn curve relationship is identified only in the former case (equation with conventional errors), with the optimal expenditure level being rather low (3.30% of GDP). This highlights possible misspecification problems in the above studies when the dependent variable in differences (that is likely stationary) is regressed against the variables in levels (that are likely to contain unit root).

Table 7. Panel regression with fixed effects estimates (Specification 3)

Variables	Logarithms of HX and HX^2		Levels of HX and HX^2	
	SE	D-K	SE	D-K
HX	0.092	0.092	0.003	0.003
	(0.052)	(0.335)	(0.680)	(0.861)
HX^2	-0.038	-0.038	-0.001	-0.001
	(0.008)	(0.138)	(0.201)	(0.558)
L	0.037	0.037	0.037	0.037
	(0.053)	(0.336)	(0.052)	(0.333)
K	-0.053	-0.053	-0.056	-0.056
	(0.000)	(0.009)	(0.000)	(0.008)
OP	0.046	0.046	0.042	0.042
	(0.043)	(0.143)	(0.062)	(0.215)
Constant	-0.064	-0.152	0.008	0.297
	(0.723)	(0.646)	(0.962)	(0.277)
Period-effect	407.190		18.210	
	(0.000)		(0.000)	
Serial correl.	48.136		41.266	
	(0.000)		(0.000)	
Heterosk.	1634.400		1667.620	
	(0.000)		(0.000)	
Optimal size (%)	3.305		NA	

Note. p-values are indicated in parentheses. SE and D-K represent conventional and Driscoll-Kraay standard errors.

Secondly, given the mixed order or integration of variables, with strong evidence of stationarity but also certain evidence of unit roots (according to some of the tests), and the likely lags in the effects of expenditure on GDP and its growth, we estimated the panel ARDL model that includes lags of dependent and independent variables (*Specification 4*). The model was estimated with trend components (given that GDP, capital stock and employment experience increase in level over time), with a maximum of two lags (given the short span of the sample), with real GDP as dependent variable, and all variables represented in levels (Table 8). Error correction term is negative and highly significant, serving as indirect indication of convergence to steady state after a temporary shock. In both specifications the Armey-Rahn curve was present, and the optimal levels of health expenditure stood at 4.43% and 4.58% of GDP.

Table 8. Panel ARDL estimates

Estimates	Logarithms of HX and HX^2		Levels of HX and HX^2	
	Coeff	Prob	Coeff	Prob
Long-run				
HX	1.696	(0.000)	0.095	(0.000)
HX^2	-0.570	(0.000)	-0.010	(0.000)
L	0.537	(0.000)	0.521	(0.000)
K	-0.323	(0.000)	-0.422	(0.000)
OP	0.062	(0.072	0.002	(0.000)
Short-run				
ECT	-0.209	(0.000)	-0.277	(0.001)
D(LGDP)			0.136	(0.061)
D(LX)	-0.145	(0.829)	-0.029	(0.593)
D(LHX(-1))	-0.037	(0.968)	-0.029	(0.728)
D(LHX^2)	0.005	(0.979)	0.000	(0.944)
D(LHX^2(-1))	0.015	(0.947)	0.002	(0.828)
D(LL)	0.696	(0.000)	0.586	(0.000)
D(LL(-1))	-0.072	(0.518)	-0.156	(0.190)
D(LK)	-0.018	(0.716)	0.044	(0.357)
D(LK(-1))	0.039	(0.472)	0.076	(0.116)
D(LOP)	0.115	(0.000)	0.001	(0.004)
D(LOP(-1))	0.048	(0.033)	0.000	(0.253)
Constant	1.559	(0.001)	2.573	(0.002)
Trend	0.004	(0.003)	0.004	(0.006)
Optimal size (%)	4.428		4.588	

Note. p-values are indicated in parentheses.

Thirdly, given that all variables, when expressed as growth rates, are stationary and that health care spending variables in levels never contained unit root, we estimated the growth equations, where GDP growth rate is a dependent variable, and the growth rates of other variables alongside the levels of health care spending variables are the regressors (Table 9). The Armey-Rahn curve was likewise present, and the optimal levels of expenditure were established at 3.74% and 4.22% of GDP in *Specification 5* (with or without logarithms of health care spending variables). The variables' statistical significance was adequate with both conventional OLS and Driscoll-Kraay errors. With cross-section fixed effects included, the signs of coefficients that form Armey-Rahn relationship remain intact, but the coefficients become statistically insignificant.

Table 9. Panel regression estimates (GDP growth as dependent variable)

Variable	Logarithms of HX and HX^2		Levels of HX and HX^2	
	SE	D-K	SE	D-K
KGR	-0.068	-0.068	-0.069	-0.069
	(0.000)	(0.031)	(0.000)	(0.027)
LGR	0.845	0.845	0.841	0.841
	(0.000)	(0.000)	(0.000)	(0.000)
HX	9.536	9.536	1.095	1.095
	(0.000)	(0.000)	(0.003)	(0.001)
HX^2	-3.617	-3.617	-0.130	-0.130
	(0.000)	(0.000)	(0.000)	(0.000)
OPGR	0.135	0.135	0.136	0.136
	(0.000)	(0.041)	(0.000)	(0.042)
Constant	-3.601	-3.601	0.069	0.069
	(0.007)	(0.000)	(0.942)	(0.933)
R^2_{adj}	0.430		0.422	
Optimal size (%)	3.738		4.217	

Note. p-values are indicated in parentheses. SE and D-K represent conventional and Driscoll-Kraay standard errors.

Fourthly, in the panel of economies considered, Switzerland stands out as outlier, with significantly low public health care expenditure share. This reflects the organisation and financing of the health system that is quite distinct from other European economies. Based on a crude classification of health systems outlined by Jenkner and Leive (2010: 3), Switzerland relies on private delivery and financing of health services similar to the USA (as opposed to public delivery and financing in, for instant, Sweden and UK, public delivery and private financing in Singapore, or private delivery and public financing in France and Germany). Based on this feature, we provide alternative estimates of the optimal spending levels on a panel that excludes Switzerland. The results of the augmented model, are not fundamentally different from the ones presented above: the optimal level of public health care spending was established at 5.12% and 4.63% of GDP (for the fixed effects regression with the levels and logarithms of the expenditure variables), and hence the expenditure is deemed overprovided.

Lastly, we considered specification with GDP per capita as dependent variable. The results of the panel quantile estimation with fixed effects confirm the earlier findings (Table 10). The expenditure variables were significant and has correct sign (in line with Armey-Rahn curve hypothesis) in most quantiles, the openness variable was positive and significant in all quantiles, while employment and real capital stock were positive, but significant only in higher quantiles. In the specification with logarithms of expenditure, the spending was under-provided in Quantiles 2 and 3 and over-provided in higher quantiles, while in the specification with levels, the spending was under-provided in Quantiles 1 to 6. The somewhat higher optimal levels of healthcare spending with GDP per capita as dependent variable (compared to the ones when absolute level of GDP is used) are attributed to presence in the panel of a large number of transition economies with below average levels of GDP per capita and more resources as percentage of GDP devoted to health expenditure (Chang, Ying, 2006: 10).

Table 10. Panel quantile regression with fixed effects estimates (GDP per capita as dependent variable)

Specification 1									
Variable	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Log(HX)	0.793	0.748	0.720	0.684	0.651	0.627	0.610	0.592	0.572
	(0.118)	(0.050)	(0.021)	(0.004)	(0.002)	(0.003)	(0.009)	(0.025)	(0.061)
Log(HX^2)	-0.212	-0.204	-0.199	-0.193	-0.187	-0.183	-0.180	-0.177	-0.173
	(0.144)	(0.062)	(0.026)	(0.005)	(0.002)	(0.003)	(0.007)	(0.019)	(0.047)
Log(L)	0.098	0.141	0.168	0.201	0.233	0.255	0.272	0.289	0.307
	(0.546)	(0.251)	(0.095)	(0.010)	(0.000)	(0.000)	(0.000)	(0.001)	(0.002)
Log(K)	0.042	0.062	0.074	0.089	0.104	0.114	0.121	0.129	0.138
	(0.728)	(0.498)	(0.320)	(0.120)	(0.035)	(0.025)	(0.029)	(0.040)	(0.059)
Log(OP)	1.603	1.554	1.523	1.485	1.448	1.423	1.404	1.384	1.363
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal size	N/A	6.264	6.107	5.904	5.700	5.561	5.454	5.336	5.212

Specification 2									
Variable	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
HX	0.173	0.160	0.151	0.141	0.130	0.122	0.117	0.111	0.105
	(0.067)	(0.028)	(0.011)	(0.003)	(0.001)	(0.003)	(0.009)	(0.028)	(0.071)
HX^2	-0.014	-0.013	-0.012	-0.011	-0.011	-0.010	-0.010	-0.009	-0.009
	(0.058)	(0.022)	(0.007)	(0.001)	(0.000)	(0.001)	(0.004)	(0.015)	(0.043)
L	0.132	0.167	0.191	0.219	0.249	0.271	0.285	0.301	0.317
	(0.408)	(0.175)	(0.058)	(0.006)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
K	0.026	0.044	0.056	0.071	0.086	0.097	0.104	0.112	0.121
	(0.826)	(0.634)	(0.456)	(0.236)	(0.087)	(0.061)	(0.064)	(0.079)	(0.103)
OP	1.612	1.566	1.535	1.498	1.458	1.429	1.411	1.390	1.369
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal size	6.638	6.295	6.247	6.183	6.123	6.036	5.987	5.939	5.876

Note. p-values are indicated in parentheses.

Conclusion

An increase in public health expenditure in the developed economies, accompanied by the rising budget deficits and public debt make cost containment and control over health expenditure a pressing problem. The determination of the optimal level of public health expenditure level (i.e. the level beyond which the negative effects on economic growth and GDP are experienced) thus becomes instrumental in policy debate and decision-making.

The hypothesis of the inverted U-shape relationship between government spending and the output and output growth (Armey-Rahn curve) has recently been subject to extensive research. The novelty and the contribution of this paper is the extension of the hypothesis and its empirical verification in the area of health spending, as a specific component of the aggregate spending. We examined the Armey-Rahn curve relationship in a panel of European economies over the 1995-2018 period and experimented with alternative specifications (in logarithms and levels of variables): the trivariate that included health spending and its squared term as determinants of GDP, and multivariate that included labour, capital and the measure of trade openness as additional regressors. To account for the spending effects across the GDP conditional distribution as well as the dynamic effects and lagged relationships, we respectively used the panel quantile model with fixed effects and the panel ARDL model. Lastly, as part of robustness checks, we estimated models with the growth rates of the variables, the GDP per capita as dependent variable, or for the panel that excluded Switzerland as an outlier economy.

In every specification the health care spending variable was positive and significant, while its square term was negative and significant, thereby confirming the Armey-Rahn hypothesis. In most specifications, the level of optimal public spending on health care was below the actual average spending of 5.99% of GDP, giving support to the excessive health spending thesis in European economies. This result is in line with the findings from the study by Chang and Ying (2006) conducted for a panel of OECD economies. The above-optimal levels of health spending were observed notwithstanding the attempts that have been made in recent decades to control health spending and the general tendency to downsize government spending. The findings of the panel quantile model with (per capita) GDP as dependent variable suggest that public health spending was underprovided in the lower quantiles. The panel included a significant number of transition economies from Eastern Europe that a). are characterised by a smaller size of GDP compared to many of the original EU members and hence larger optimal level of health spending as proportion of GDP; b). health spending that has been lagging behind the one in Western Europe during the transition process and the implementation of economic adjustment policies in the post-communism era, and c). Indivisibility of public goods and health investment that result in high levels of optimal health spending. As stated by Chang and Ying (2006) this result is counterintuitive, as more developed economies are expected to value health outcomes more than less developed ones, and are likely to have higher optimal (desired) levels of health spending.

The findings of this and similar studies need to be interpreted appropriately. Firstly, while the majority of countries in the panel have socialised health care, the well-entrenched policies and above-optimal spending levels, it may be argued that it is not the level of spending that is too high, but the growth rate of GDP (that is a product of many other factors unrelated to the level of spending) that is too low. Secondly, the deficiencies of the 'single figure' method of determining spending optimality need to be noted. As put by Savedoff (2007), the issue of public health care spending is to be discussed concurrently with the issue of public-private spending crowding out. Additionally, the 'single figure' approach may need to be complemented by the discussion of the composition of the health care spending, the efficiency of health spending (with similar levels of spending per capita or spending shares delivering profoundly different health outcomes in different countries), the influence of vested interests on the level of spending (resulting expenditure level that is too high or too low than the optimal one), the national development and planning priorities and objectives, as well as the costs of providing public health and the relevant budget considerations. In this regard, the question of what the optimal health care spending is not purely technical or econometric, but has political-economic and social dimensions. Thirdly, it may be theoretically incorrect to make generalisations and consider a single optimal level for a diverse set of countries that have specific institutional and political-economic settings, the pace of medical innovation, the health care reforms implemented, and other factors. Each country may then have its own unique shape of the Armey-Rahn curve and unique set of factors that shift it. Fourthly, longer time series may be needed to examine the Armey-Wagner curve: given that sample size is restricted to the years 1995-2018 (that fall within a longer period of government downsizing that has started in the early 1980s), there is possibility that the optimal point has been reached prior to that period. Nonetheless, the findings in this paper may be instrumental in the discussions on health policy, in the policy design, and in the appraisal of the level of government expenditure.

Appendix

Table 2. Descriptive statistics for individual cross-sections

Country	GDP		GDP per capita		HX	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Austria	280473	33507	33753.73	3028.5	**7.951**	0.305
Belgium	340715	42517	31809.1	2790.6	7.180	0.445
Bulgaria	34035	7160	4488.437	1175.8	*4.307*	0.937
Cyprus	16573	2920	21441.67	2210.6	*2.849*	0.185
Czechia	142305	26056	13677.12	2350.2	**7.767**	0.401
Denmark	238508	22257	43447.23	2898.6	**7.831**	0.665
Estonia	14231	3623	10557.35	2917.6	*4.933*	0.381
Finland	174332	23843	32814.22	3870.0	7.246	0.314
France	1906537	199274	29936.79	2010.5	**7.690**	0.354
Germany	2518150	238615	30962.86	2871.1	6.478	0.559
Greece	201787	27329	18513.63	2281.1	6.075	0.488
Hungary	95122	14999	9483.666	1640.9	*5.375*	0.401
Ireland	163351	51619	37739.37	8718.3	6.733	0.695
Italy	1572830	73004	26797.59	1165.2	6.832	0.455
Latvia	17162	4514	8008.044	2618.5	*3.683*	0.447
Lithuania	26205	7130	8325.631	2887.9	*5.535*	0.785
Luxembourg	36874	8178	74406.91	8600.7	*4.593*	0.404
Netherlands	601314	71564	36642.49	3323.3	6.296	1.368
Norway	308899	38887	64568.95	4492.4	7.217	0.468
Poland	319686	85402	8336.583	2207.4	*4.383*	0.688
Portugal	169353	12742	16308.1	1051.5	6.991	0.573
Romania	116180	28004	5596.596	1627.5	*4.465*	1.054
Slovakia	58814	15780	10879.16	2881.0	6.269	0.940
Slovenia	32920	5559	16245.05	2534.7	6.714	0.275
Spain	980507	134642	22261.25	1927.8	*5.701*	0.648
Sweden	347144	58220	37308.26	4732.7	6.830	0.269
Switzerland	413477	55910	53660.99	4210.0	***1.871***	0.183
United Kingdom	1793808	242426	29045.62	2785.3	**7.853**	0.712
All	461475	659128	26322.01	17745.7	5.987	1.680

Note. The economies with above or below average health care spending (as proportion of GDP) as indicated in bold and italics respectively. Switzerland as outlier is indicated in both bold and italics.

Mean HX per quantile (as % of GDP)			
Q1	3.471	Q6	6.811
Q2	4.557	Q7	7.141
Q3	5.088	Q8	7.462
Q4	5.782	Q9	7.858
Q5	6.442		

Table 4. Panel unit root tests’ results

Model in levels (constant)

Test	GDP	K	L	OP	HX	HX^2	Log(HX)	Log(HX^2)	GDPGR	KGR	LGR	OPGR
LLC	-5.006	*-1.533*	**-0.915**	-6.725	-3.289	-3.370	-3.201	-3.139	-9.436	-14.297	-6.215	-12.430
IPS	**0.286**	*-1.434*	**0.773**	**-0.943**	-3.070	-2.987	-3.239	-3.322	-8.362	-14.534	-7.674	-12.296
ADF-Fisher	**51.298**	78.698	**55.088**	**65.923**	89.520	88.182	92.515	94.038	173.643	480.522	160.789	253.348
PP-Fisher	74.926	103.063	**39.902**	120.996	90.819	80.765	107.824	113.512	203.843	550.238	138.045	443.867
Pesaran CIPS	-2.478	-2.089	-2.209	-2.100	-2.119	-2.086	-2.176	-2.126	-2.663	-3.052	-2.469	-3.036

Note. The statistics highlighted in italics or bold indicate non-rejection of the null hypothesis of a unit root at the 5% and 10% level respectively.

Model in levels (constant plus trend)

Test	GDP	K	L	OP	HX	HX^2	Log(HX)	Log(HX^2)
LLC	-4.225	-1.982	-4.197	-3.684	-2.299	-2.420	-1.847	*-1.622*
Breitung	-2.586	**1.908**	-4.763	**-0.356**	-1.853	-1.831	-2.073	-2.086
IPS	-2.157	-2.344	-3.963	-3.045	-2.453	-2.481	-2.473	-2.487
ADF-Fisher	**69.061**	102.711	99.094	93.755	84.243	84.481	84.940	85.354
PP-Fisher	**33.242**	152.258	**24.982**	81.547	*72.248*	**58.520**	87.872	92.159
Pesaran CIPS	-3.124	**-2.406**	-2.765	**-2.338**	**-2.316**	**-2.272**	**-2.363**	**-2.325**

Note. As above.

Model in the first differences (constant)

Test	GDP	K	L	OP	HX	HX^2	Log(HX)	Log(HX^2)	GDPGR	KGR	LGR	OPGR
LLC	-9.441	-12.397	-6.171	-12.271	-12.487	-12.831	-11.376	-10.948	-18.085	-37.841	-14.206	-20.365
IPS	-8.399	-13.549	-7.673	-12.175	-11.936	-12.145	-11.798	-11.758	-18.098	-25.268	-13.702	-21.457
ADF-Fisher	174.201	476.485	160.717	250.580	244.787	249.194	241.867	241.061	376.091	650.060	281.181	448.910
PP-Fisher	205.921	554.214	138.423	442.578	404.001	398.945	428.252	438.904	1086.520	2072.940	770.071	4148.830
Pesaran CIPS	-2.687	-3.032	-2.477	-3.005	-3.034	-3.035	-3.051	-3.039	-3.402	-4.632	-3.087	-4.453

Note. As above.

Model in the first difference (constant plus trend)

Test	GDP	K	L	OP	HX	HX^2	Log(HX)	$Log(HX^2)$
LLC	-7.164	-13.110	-3.778	-10.229	-10.597	-10.916	-9.960	-9.190
Breitung	-8.362	-3.420	-8.197	-10.720	-11.069	-11.767	-10.304	-10.185
IPS	-5.368	-11.615	-4.640	-10.028	-9.229	-9.439	-9.088	-9.060
ADF-Fisher	122.282	430.763	110.290	197.075	183.490	187.227	181.013	180.540
PP-Fisher	155.768	523.994	94.825	417.418	342.657	332.224	370.703	401.582
Pesaran CIPS	-2.629	-3.175	*-2.544*	-3.146	-3.130	-3.124	-3.164	-3.140

Note. As above.

References

Aly, H., Strazicich, M. (2000). Is Government Size Optimal in the Gulf Countries of the Middle East? An Empirical Investigation. International Review of Applied Economics, 14 (4), pp. 475-483.

Anaman, K. A. (2004). Determinants of Economic Growth in Brunei Darussalam. Journal of Asian Economics, 15 (4), pp. 777-796.

Balatsky, E. (2012). Wagner's Law, the Armey-Rahn Curve, and the Paradox of Wealth. Problems of Economic Transition, 54 (12), pp. 40-54.

Baumol, W. J. (1993). Health Care, Education and the Cost Disease: A Looming Crisis for Public Choice. Public Choice, 77 (1), pp. 17-28.

Breitung, J. (2000). The Local Power of Some Unit Root Tests for Panel Data. In Baltagi, B. (Ed.) Nonstationary Panels, Panel Cointegration, and Dynamic Panels, Advances in Econometrics, Vol. 15. Amsterdam: JAI, pp. 161-178.

Breusch, T., Pagan, A. (1980). The Lagrange Multiplier Test and its Application to Model Specification in Econometrics. Review of Economic Studies, 47, pp. 239-253.

Buchinsky, M. (1994). Changes in the US Wage Structure 1963-1987: Application of Quantile Regression. Econometrica, 62 (2), pp. 405-458.

Cetin, M. (2017). Does Government Size Affect Economic Growth in Developing Countries? Evidence from Non-Stationary Panel Data. European Journal of Economic Studies, 6 (2), pp. 85-95.

Chang, K., Ying, Y.-H. (2006). Economic Growth, Human Capital Investment, and Health Expenditure: A Study of OECD Countries. Hitotsubashi Journal of Economics, 47, pp. 1-16.

Chao, J. C. P., Grubel, H. (1998). Optimnal Levels of Spending and Taxation in Canada. In Grubel, H. (Ed.). How to Use Fiscal Surplus - What is the Optimal Size of Government? Vancouver: The Fraser Institute.

Chiou-Wei, S.-Z., Zhu, Z., Kuo, Y.-H. (2010). Government Size and Economic Growth: An Application of the Smooth Transition Regression Model. Applied Economics Letters, 17 (14), pp. 1405-1415.

Chobanov, D., Mladenova, A. (2009). What is the Optimum Size of Government? Sofia: Institute for Market Economics.

Choi, I. (2001). Unit Root Tests for Panel Data. Journal of International Money and Finance, 20 (2), pp. 249-272.

Clemente, J., Marcuello, C., Montanes, A., Pueyo, F. (2004). On the International Stability of Health Care Expenditure Functions: Are Government and Private Functions Similar? Journal of Health Economics, 23, pp. 589-613.

Davies, A. (2009). Human Development and the Optimal Size of Government. The Journal of Socio-Economics, 38 (2), pp. 326-330.

Di Matteo, L. (2003). The Income Elasticity of Health Care Spending. European Journal of Health Economics, 4, pp. 20-29.

Driscoll, J. C., Kraay, A. C. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. Review of Economics and Statistics, 80 (4), pp. 549-560.

Durevall, D., Henrekson, M. (2011). The Futile Quest for a Grand Explanation of Long-Run Government Expenditure. Journal of Public Economics, 95 (7-8), pp. 708-722.

De Hoyos, R., Sarafidis, V. (2006). Testing for Cross-Sectional Dependence in Panel-data Models. The Stata Journal, 6 (4), pp. 482-496.

De Witte, K., Moesen, W. (2010). Sizing the Government. Public Choice, 145 (1/2), pp. 39-55.

El Husseiny, I. (2018). The Optimal Size of Government in Egypt: An Empirical Investigation. The Journal of North African Studies, 24 (2), pp. 271-299.

Facchini, F., Melki, M. (2011). Optimal Government Size and Economic Growth in France (1871-2008): An Explanation by the State and Market Failures. Centre d'Economie de la Sorbonne, CES Working Paper.

Finkelstein, A. (2007). The Aggregate Effects of Health Insurance: Evidence from the Introduction of Medicare. Quarterly Journal of Economics, 122 (1), pp. 1-37.

Forte, F., Magazzino, M. (2018). Wagner's Law, Government Size and Economic Growth: An Empirical Test and Theoretical Explanations for Italy 1861-2008. In Backhaus, J., Chaloupek, G., Frambach, H. (Eds.). Gustav von Schmoller and Adolf Wagner. Legacy and Lessons for Civil Society and the State. Heidelberg: Springer.

Gemmell, N., Kneller, R., Sanz, I. (2014). Does the Composition of Government Expenditure Matter for Long-Run GDP Levels? Victoria University of Wellington, Working Paper No. 10/2014, pp. 1-37.

Gerdtham, U.-G., Sogaard, J., Andersson, F., Jonsson, B. (1992). An Econometric Analysis of Health Care Expenditure: A Crosssection Study of the OECD Countries. Journal of Health Economics, 11, pp. 63-84.

Getzen, T. (2000). Health Care is an Individual Necessity and a National Luxury: Applying Multilevel Decision Models to the Analysis of Health Care Expenditure. Journal of Health Economics, 19, pp. 259-270.

Goel, R. K., Payne, J. E., Ram, R. (2008). R&D Expenditures and U.S. Economic Growth: A Disaggregated Approach. Journal of Policy Modeling, 30 (2), pp. 237-250.

Greene, W. (2000). Econometric Analysis. New York: Prentice-Hall.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. Journal of Political Economy, 80 2), pp. 223-255.

Hajamini, M., Ali Falahi, M. (2014). The Nonlinear Impact of Government Consumption Expenditure on Economic Growth: Evidence from Low and Low-middle Income Countries. Cogent Economics and Finance, 2 (1), pp. 1-15.

Handoussa, H., Reiffers, J.-L. (2003). Accelerating the Implementation of the Necessary Reforms to Modify the Growth Regime. In FEMISE Report 2003 on the Euro-Mediterranean Partnership. Marseilles: Institut de la Mediterranee.

Haylock, M. (2020). Executives' Short-Term and Long-Term Incentives - A Distributional Analysis. University of Tubingen, Working Paper in Business and Economics No. 131.

Herath, S. (2012). Size of Government and Economic Growth: A Nonlinear Analysis. Economic Annals, 57 (194), pp. 7-30.

Hitiris, R., Posnett, J. (1992). The Determinants and Effects of Health Expenditure in Developed Countries. Journal of Health Economics, 11, pp. 173-181.

Hok, L., Jariyapan, P., Buddhawongsa, P., Tansuchat, R. (2014). Optimal Size of Government Spending: Empirical Evidence from Eight Countries in Southeast Asia. The Empirical Econometrics and Quantitative Economics Letters, 3 (4), pp. 31-44.

Ike, G. N., Usman, O., Sarkodie, S. A. (2020). Testing the Role of Oil Production in the Environmental Kuznets Curve of Oil Producing Countries: New Insights from Method of Moments Quantile Regression. Science of the Total Environment, 711, pp. 1-10.

Im, K. S., Pesaran, M. H., Shin, Y. (2003). Testing for Unit Roots in Heterogeneous Panels. Journal of Econometrics, 115 (1), pp. 53-74.

Jenkner, E., Leive, A. (2010). Health Care Spending Issues in Advanced Economies. Washington, DC: International Monetary Fund.

Karras, G. (1996). The Optimal Government Size: Further International Evidence on the Productivity of Government Services. Economic Inquiry, 34 (2), pp. 193-203.

Koenker, R. (2004). Quantile Regression for Longitudinal Data. Journal of Multivariate Analysis, 91 (1), pp. 74-89.

Koenker, R., Bassett Jr., G. (1978). Regression Quantiles. Econometrica, 46 (1), pp. 33-50.

Kustepeli, Y. (2005). The Relationship Between Government Size and Economic Growth: Evidence from a Panel Data Analysis. Dokuz Eylul University, Department of Economics, Discussion Paper Series No. 05/06, pp. 1-10.

Levin, A., Lin, C. F., Chu, C. (2002). Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties. Journal of Econometrics, 108 (1), pp. 1-24.

Li, Z., Wang, Y., Tan, Y., Huang, Z. (2020). Does Corporate Financialization Affect Corporate Environmental Responsibility? An Empirical Study of China. Sustainability, 12 (9), pp. 1-19.

Machado, J. A. F., Santos Silva, J. M. C. (2019). Quantiles via Moments. Journal of Econometrics, 213 (1), pp. 145-173.

Maddala, G. S., Wu, S. (1999). A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test. Oxford Bulletin of Economics and Statistics, 61 (S1), pp. 631-652.

Magazzino, C. (2008). Modeli Interpretation della Dinamia della Spesa Pubblica e 'curva di Armey': Il Caso Italiano 1862-1998. Notizie di Politica, XXIV (92), pp. 45-60.

Magazzino, C., Forte, F. (2010). Optimal Size of Government and Economic Growth in EU-27. MPRA Paper No. 26669.

Mavrov, H. (2007). The Size of Government Expenditure and the Rate of Economic Growth in Bulgaria. Economic Alternatives, 1, pp. 53-63.

Miller, K. A. (2008). An Analysis of State and Local Public Expenditure and Output-Maximizing Spending Optimums. The University of Oklahoma, Unpublished Ph.D. Dissertation.

Niskanen, W. (1971). Bureaucracy and Representative Government. Chicago: Aldine and Atherton.

Organisation for Economic Cooperation and Development / OECD (2011). Government at a Glance, 2011. Paris: OECD Publishing.

Organisation for Economic Cooperation and Development / OECD (2015). Fiscal Sustainability of Health Systems: Bridging Health and Finance Perspectives. Paris: OECD Publishing.

Peden, E. A. (1991). Productivity in the United States and its Relationship to Government Activity: An Analysis of 57 Years, 1929-1986. Public Choice, 69 (2), pp. 153-173.

Persson, T., Tabellini, G. (1990). Macroeconomic Policy, Credibility and Politics. London: Harwood Academic.

Pesaran, M. H. (2004). General Diagnostic Tests for Cross Section Dependence in Panels. IZA Discussion Paper No. 1240, pp. 1-39.

Pesaran, M. H. (2007). A Simple Panel Unit Root Test in the Presence of Cross Section Dependence. Journal of Applied Econometrics, 22 (2), pp. 265-312.

Pesaran, M. H., Smith, R. (1995). Estimating Long-Run Relationships from Dynamic Heterogeneous Panels. Journal of Econometrics, 68 (1), pp. 79-113.

Pesaran, M. H., Shin, Y., Smith, R. (1999). Pooled Mean Group Estimation of Dynamic Heterogeneous Panels. Journal of the American Statistical Association, 94 (446), pp. 621-634.

Pevcin, P. (2004). Size of Budgetary and Non-Budgetary Government: Explaining Cross-Country Differences. University of Ljubljana, School of Economics and Business Working Paper.

Rahn, R., Fox, H. (1996). What is the Optimum Size of Government? Denver: Vernon K. Krieble Foundation.

Savedoff, W. (2007). What Should a Country Spend on Health Care? Health Affairs, 4, pp. 962-970.

Schoeman, N. J., Van Heerden, Y. (2009). Finding the Optimum Level of Taxes in South Africa: A Balanced Budget Approach. The South African Journal of Industrial Engineering, 20 (2), pp. 15-31.

Scully, G. W. (1994). What is the Optimal Size of Government in the United States? Dallas: National Centre for Policy Analysis, Policy Report No. 188.

Scully, G. W. (2000). The Growth-maximizing Tax Rate. Pacific Economic Review, 5 (1), pp. 93-96.

Smith, S., Newhouse, J., Freeland, M. (2009). Income, Insurance, and Technology: Why Does Health Spending Outpace Economic Growth? Health Affairs, 28 (5), pp. 1276-1284.

Tanzi, V., Schuknecht, L. (1997). Reconsidering the Fiscal Role of Government: The International Perspective. American Economic Review, 87 (2), pp. 164-168.

Vedder, R., Gallaway, L. (1998). Government Size and Economic Growth. US Congress, The Paper Prepared for the Joint Economic Committee.

Wang, F. (2015). More Health Expenditure, Better Economic Performance? Empirical Evidence from OECD Countries. Inquiry: The Journal of Health Care Organization, Provision, and Financing, 52, pp. 1-5.

Wooldridge, J. M. (2002). Econometric Analysis of Cross Section and Panel Data. Cambridge: MIT Press.

Wu, S.-Y., Tang, J.-H., Lin, E. S. (2010). The Impact of Government Expenditure on Economic Growth: How Sensitive to the Level of Development? Journal of Policy Modeling, 32 (6), pp. 804-817.

Yavas, A. (1998). Does Too Much Government Investment Retard Economic Development of a Country. Journal of Economic Studies, 25 (4), pp. 296-308.

Yazici, E., Kaestner, R. (1998). Medicaid Expansions and the Crowding Out of Private Health Insurance. NBER Working Paper No. 6527.

Ying, Y.-H., Chang, K. (2019). The Effect of National Health Insurance on Private Health Expenditure in Taiwan: Crowd In or Crowd Out? Journal of Asia Pacific Economy, 25 (2), pp. 1-16.

Zhang, R., Li, Z.-F. (2008). Optimal Scale of the Rural Fiscal Expenditure Based on Economic Growth in China. Proceedings of the 15th International Conference on Management Science and Engineering, Long Beach, USA, 10-12 September 2008.

Zwiefel, P., Felder, S., Meirs, M. (1999). Ageing of Population and Health Care Expenditure: A Red Herring? Health Economics, 8, pp. 485-496.

---

[i] Additional to these are other health expenditure challenges, for instance over- or under-spending in particular areas (as demonstrated by the lack of the emergency capacity and spending during the corona-virus crisis).

[ii] We note a stream of literature that examines marginal product and output elasticity of government expenditure and determines the over-, under-, or optimal provision of government expenditure and the respective optimum points (Karras, 1996; Aly, Strazicich, 2000). These studies, however, do not establish the presence of Armey curve (since over- or under-provision of expenditure may take place alongside the linear relationship between output and expenditure variables).

[iii] Maddala, Wu, 1999; Breitung, 2000; Choi, 2001; Levin et al, 2002; Im et al, 2003; Pesaran, 2007. We note that all of these tests contrast the null hypothesis of the presence of a unit root against the alternative hypothesis of (trend) stationarity (for all panel members in LLC and Breitung tests and for some but not all members in other tests).