Economic Modelling 59 (2016) 23–31

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.elsevier.com/locate/ecmod

Testing threshold cointegration in Wagner's Law: The role of military spending

Maddalena Cavicchioli[a,*], Barbara Pistoresi[b,1]

[a]Department of Economics "Marco Biagi", Viale Berengario 51, 41121 Modena, Italy
[b]Department of Economics "Marco Biagi" and ReCent, Viale Berengario 51, 41121 Modena, Italy

ARTICLE INFO

Article history:
Received 25 February 2016
Received in revised form 13 June 2016
Accepted 22 June 2016
Available online 9 July 2016

JEL classification:
C01
C05
C34
H1
H5

Keywords:
Time series
Nonlinearity
Threshold Vector Error Correction
Public spending
Economic growth
Wagner's Law

ABSTRACT

This paper analyses historical data since the mid-19th century to find support for Wagner's Law in the Italian economy. Unlike previous studies, we accommodate possible nonlinear asymmetric effects of government spending and GDP towards their long-run equilibrium. The results reveal a threshold cointegrating relationship between the two variables with significantly different error correction adjustments in normal and extreme regimes. A long-run tendency for the public sector to grow relative to GDP from 1862 to 2009 is observed only when nonlinearities generated by temporary higher military spending during wars are take into account.

1. Introduction

The provision of public goods and services to citizens using the central state's fiscal capacity implies that government expenditure underwent constant growth during the 20th century, despite institutional and cultural differences (Tanzi and Schuknecht, 2000). A simple explanation for the long-run determination of public spending was proposed by Wagner (1883) and is known as Wagner's Law (henceforth WL). WL states that a positive relation exists between level of economic development and scope of government. State expansion is driven by a growing demand for defence, public investment in infrastructures, education and wealth, but also for the regulation and enforceability of contracts which arises as a society becomes more complex.

These processes are not rigorously derived from a context of individual utility maximization. Some exceptions exist in WL literature. In a public choice framework, Meltzer and Richard (1981), Persson and Tabellini (1990) and Lindert (1994, 2004a,b)propose an economic foundation for WL, in which WL emerges as a game between government and electorate. Governments tailor expenditure policies towards satisfying the median voter and this behaviour induces a relationship between public spending and national income. An alternative theoretical foundation of WL emerges as a Principal-Agent problem. As pointed out by Oxley (1994), bureaucrats are rational utility maximizers that derive utility from power and prestige and expand the size of their bureaus at the expense of efficiency. While the microeconomic foundations of WL are rarely discussed, a large number of studies focus on an empirical assessment of WL from different perspectives and applying different techniques. For recent overviews see Durevall and Henrekson (2011), Kuckuck (2014)and Narayan et al. (2008, 2012).

Generally speaking, an empirical strategy to investigate the relationship between public spending and economic growth involves the detection of causal links in a long-run perspective. Most examples

* Corresponding author.
E-mail addresses: maddalena.cavicchioli@unimore.it (M. Cavicchioli), barbara.pistoresi@unimore.it (B. Pistoresi).
1 Tel.: +39 059 2056881.

http://dx.doi.org/10.1016/j.econmod.2016.06.011
0264-9993/

in the literature start from an analysis of a bivariate error-correction regression model when a long-run relationship is observed between the variables of interest. Causality is then discussed applying a Granger bivariate causal structure. Such linear long-run relationships have been called into questioned in various ways. Only a few studies model structural breaks to evaluate shifts in the long-run; others test for positive or negative deviations from the trend in the short-run, implying asymmetric adjustment in the long-run (see Durevall and Henrekson (2011)).

The present study aims to detect any endogenous nonlinearities within an analysis of the long-run relationship between the variables of interest. The specific purpose is to contribute to the analysis of WL within the Italian economy based on historical time series from 1862 to 2009. Italy is an interesting case-study because it was a late-comer to industrialization that caught up in the late 19th century and then exhibited an excellent economic performance that enabled it to join the G7 group in the 1970s. This long time span is considered to be an appropriate framework for empirically assessing WL since it captures the evolution of government expenditure in response to the country's social and economic progress. Over such a long period, Italy underwent a number of economic and socio-political changes that represent potential sources for nonlinearities in the data including WWI and WWII, the Great Depression, and the socio-political turmoil of the post-war period. Such events might be the causes of different asymmetric responses in government spending to variations in national income. If this is the case, failure to take these data features into account could induce biased empirical results and misleading conclusion.

The main contributions of the paper are the following. Firstly, unlike previous studies, nonlinear cointegration is considered in order to analyse WL. The methodology of Hansen and Seo (2002) is applied to incorporate the possibility of threshold effects in the cointegrating relationship, nesting linear cointegration and allowing for the potential existence of one or more regimes. Secondly, support for WL in the Italian case over the period 1862–2009 can be identified only when the strong asymmetric responses of government spending during WWI and WWII are taken into consideration. Robustness checks recognize nonlinear behaviour of government spending driven by temporary higher military expenditure. Hence, the presence of asymmetric adjustments in the response of government spending may explain why the bulk of empirical evidence concerning WL is inconclusive. Finally, our paper also differs from existing studies of Italy because it relies on up-to-date series of national income and public spending provided on the occasion of the 150th anniversary of Italy's unification. The research department of the Bank of Italy, together with academics from other institutions, presented a reconstruction of new Italian national historical accounts, now in Baffigi (2015). These new series were the basis for the recently published *Oxford Handbook of the Italian Economy since Unification* (Toniolo, 2013), a volume including eighty-five pages of quantitative data on the Italian economy since 1861. Italy's State General Accounting Department also published a special issue on total government expenditure and on its specific economic and functional items (actual payments in fiscal years) since Italy's political unification (RGS, 2011). Broadly speaking, our data differ from existing literature (e.g., Magazzino (2012) and Kuckuck (2014)) which uses either shorter time spans or not-revisited historical data or different public accounting methods (actual payments vs accrued expenses). Note that in the sequel, we intend to detail the differences between the new series on national income and public spending and those used by recent papers testing WL for Italy.

The paper is organised as follows. Section 2 provides an overview of the empirical literature regarding WL. The econometric framework in relation to linear and nonlinear cointegration is described in Section 3. Section 4 describes the data in detail and comments on some stylized facts. Section 5 presents the empirical results. Robustness checks are presented in Section 6. Section 7 concludes and offers suggestions for future research.

2. Wagner's Law

The long-run relationship between the size of the public sector and economic growth remains an important stylized fact accepted in the literature of public economics. Wagner (1883) offers a simple explanation for this: the growth of government expenditure is a consequence of the expansion of the state driven by a country's social and economic development. The urbanization and greater division of labour that accompany industrialization require, for example, more government regulation and higher expenditure on contractual enforcement and law and order. Other causes are the growing need to finance large-scale investments of benefit to the general public (i.e. infrastructures) and the supposed superior income-elasticity of publicly provided goods and services, such as education, welfare, but also national security or defence.

In line with Wagner's conception of a developing society, North (1985, p. 392) stresses the role of technological progress: " technological changes have led to an enormous increase in specialisation and division of labour, and therefore a radical change in relative prices which fundamentally altered the traditional structure of the polity, the family, and economic organisation. The variety of interest groups that emerged from this expanded division of labour led to political pluralism. The demand for new institutional forms of organisation to replace functions previously undertaken by the family and traditional economic organisation could not be completely realized by voluntary organisations because of moral hazard, adverse selection, and the demand for public goods".

Most of the empirical literature focuses on developed or developing economies over relatively short time spans, generally starting from the 1960s. The majority compare the results for industrialized and emerging economies in order to confirm the relationship between level of development and WL, although there are significant differences between a modern state in the 19th century and recent developing economies, in terms of culture, institutions and the conception of the state's role.

By contrast, the analysis of WL in a long-run perspective, for a single country or countries with similar social, economic and political conditions has attracted much less attention. Few studies analyse very long time spans and generally reject WL. Henrekson (1993) and Bohl (1996) find no support for WL in Sweden from 1861 to 1990 or in the United Kingdom from 1870 to 1995, respectively; Ghate and Zak (2002) do not find any empirical evidence in the United States from 1929 to 2000; Durevall and Henrekson (2011) find direct evidence in favour of WL only for Sweden and the United Kingdom for a time period from around 1860 to 1970. There are, however, shorter time spans during which WL holds. For example, Oxley (1994) for the United Kingdom, Thornton (1999) for Denmark, Germany, Italy, Norway, Sweden, and the United Kingdom, and Durevall and Henrekson (2011) for Sweden and the United Kingdom confirm the validity of WL in the 50 year period preceding World War I. Recently, Kuckuck (2014) examines UK, Denmark, Sweden, Finland and Italy, finding that a long-run equilibrium between public spending and economic growth does exist but WL is seen to be more valid during the early stages of development.

In order to test WL, the literature assumes different functional forms linking public spending and national income. The present paper applies the following specification

$$g_t = \alpha + \theta\, y_t \tag{1}$$

where g is the logarithm of total government expenditure in nominal terms as a share of nominal GDP, and y is the logarithm of real per capita GDP. The above formulation is probably the most common in the literature and the majority of other models are simple reformulations of it (see Durevall and Henrekson (2011)). Contrary to the alternative specification that considers total government expenditure and GDP, it has the advantage, by using real per capita GDP, of better assessing a nation's prosperity and the spending capacity of its citizens. Moreover, the use in Eq. (1) of total government expenditure as a share of GDP instead of real total government expenditure allows for the possibility that differences in productivity growth in government and private sector production lead to an increase in government spending due to " Baumol's disease".[2] Eq. (1) models the evolution of the demand for public goods and services in the long-run through the coefficient θ which according to WL should be greater than zero. In this case, government expenditure increases faster than GDP, i.e., government expenditure is income-elastic, or in other words, it is a superior good. Cointegration and Granger causation offer an econometric framework to estimate Eq. (1). WL requires that if a long-run relationship exists, i.e. y and g are cointegrated, y must also Granger-cause g and not viceversa. In this case, GDP evolution contains additional information about the long run determination of government spending, and not viceversa (i.e. GDP is weakly exogenous).

It should be noted that, while a strict interpretation of Wagner's statement indicates a positive correlation between government expenditure and GDP, this does not imply a strong causality between one variable and the other. The idea is that economic development is associated with an increased role of the government, but Wagner does not provide an articulated model of a growth process in which cause and effect are clearly delineated (on this point, see Peacock and Scott (2000)). Although this view of WL is compatible with weak exogeneity of GDP, the empirical literature widely uses the Granger causation test to derive causal macroeconomic implication for the relationship between the two variables. These policy implications are outlined in Magazzino (2012) where four types of causation are considered: Wagnerian causation (i.e. GDP Granger causes public spending), Keynesian causation (i.e. public spending Granger causes GDP), feedback (bi-direction causality) and neutrality (no causality exists).

In the subsequent paper we will refer to the concept of Granger causation as weak exogeneity, providing evidence that WL holds when y contains relevant information to understand the evolution of g and not viceversa.

3. Linear and threshold cointegration

To examine the relationship between government spending and growth, the usual research strategy starts from a vector error-correction model (VECM) in a bivariate framework. Let $\mathbf{x}_t$ be a p-dimensional $I(1)$ time series which is cointegrated with one $p \times 1$ cointegrating vector $\boldsymbol{\beta}$, with T observations and ℓ maximum lag length. A linear VECM of order $\ell + 1$ can be expressed in compact form as

$$\Delta\mathbf{x}_t = \mathbf{A}'\mathbf{X}_{t-1}(\boldsymbol{\beta}) + \mathbf{u}_t \tag{2}$$

with

$$\mathbf{X}_{t-1}(\boldsymbol{\beta}) = \left\{ \begin{array}{l} 1 \\ w_{t-1}(\boldsymbol{\beta}) \\ \Delta\mathbf{x}_{t-1} \\ \Delta\mathbf{x}_{t-2} \\ \vdots \\ \Delta\mathbf{x}_{t-\ell} \end{array} \right. \tag{3}$$

where $w_t(\boldsymbol{\beta}) = \boldsymbol{\beta}'\mathbf{x}_t$ denotes the $I(0)$ error correction term. The regressor $\mathbf{X}_{t-1}(\boldsymbol{\beta})$ is $k \times 1$ and $\mathbf{A}$ is $k \times p$, with $k = p\ell + 2$. The error $\mathbf{u}_t$ is assumed to be a vector martingale difference sequence with finite covariance matrix $\boldsymbol{\Sigma} = E(\mathbf{u}_t\mathbf{u}_t')$. In our case, with the previous notation, the vector of interest will be $\mathbf{x}_t = (g_t y_t)'$ with $p = 2$ and one cointegrating vector. One element of $\boldsymbol{\beta}$ can be set equal to unity to achieve identification. Therefore, the second row of $\mathbf{A}$ coincides with the vector containing speed of adjustment coefficients and $w_t(\boldsymbol{\beta})$ with the cointegrating relationship described in Eq. (1). The error correction model offers a test of causality in terms of weak exogeneity of the dependent variable. If two variables are cointegrated, at least one error correction term is expected to be significantly non-zero (see, for instance, Islam (2001) for an application).

However, while previous studies have massively employed cointegration analysis, only some have adequately addressed the issue of regime change and, generally speaking, of any asymmetric shifts in the relationship between the variables. As government spending peaks may occur only above certain levels of economic activity, the use of threshold cointegration could be potentially more meaningful to capture the underlying dynamics of the data (see Chevallier (2011) or Subervie (2011) for other applications). In a threshold model, one set of dynamics often describes the usual state of the world, while another set describes the behaviour in less usual periods. Hence threshold cointegration extends the linear case by allowing adjustment to occur after deviation exceeds some critical level. So, while a linear VECM model assumes a constant adjustment rate towards a long-run equilibrium, a threshold cointegration approach instead assumes that error correction occurs depending on the threshold.

The approach adopted here was developed by Hansen and Seo (2002), with a two-regime threshold VECM model proposed as a convenient method to combine nonlinearity and cointegration. The threshold VECM of order $\ell + 1$ extends the model (2)–(3) taking the form

$$\Delta\mathbf{x}_t = \begin{cases} \mathbf{A}_1'\mathbf{X}_{t-1}(\boldsymbol{\beta}) + \mathbf{u}_t & \text{if} \quad w_{t-1}(\boldsymbol{\beta}) \le \gamma \\ \mathbf{A}_2'\mathbf{X}_{t-1}(\boldsymbol{\beta}) + \mathbf{u}_t & \text{if} \quad w_{t-1}(\boldsymbol{\beta}) > \gamma \end{cases} \tag{4}$$

where γ is the threshold parameter and the coefficient matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ govern the dynamics in the two regimes. The threshold effect has content if

$$\pi_0 \le P(w_{t-1} \le \gamma) \le 1 - \pi_0 \tag{5}$$

where $\pi_0 > 0$ is a trimming parameter, which is set as 0.05, otherwise the model simplifies to linear cointegration. Estimation of the above model is performed by Maximum Likelihood (MLE) under the assumption that the errors are *iid* Gaussian.

Within this framework, a preliminary test verifies the presence of threshold cointegration. It tests a null hypothesis of linear cointegration (VECM) versus an alternative of threshold cointegration

[2] " Baumol's disease" involves a rise in wages for jobs, like, for example the government sector, which has experienced no increase in labour productivity in response to rising wages in jobs in private sector production which has experienced improved labour productivity. The rise in wages for jobs without productivity gains is a result of the need to compete for employees against jobs that have experienced improvements and can thus naturally pay higher salaries (Baumol, 1967).

(two-regime VECM). The test statistic is a Lagrange Multiplier (LM) test of the following form:

$$\mathrm{SupLM} = \sup_{\gamma_L \leq \gamma \leq \gamma_U} \mathrm{LM}(\tilde{\boldsymbol{\beta}}, \gamma) \tag{6}$$

where $\tilde{\boldsymbol{\beta}}$ is the null estimate of $\boldsymbol{\beta}$, the search region $[\gamma_L, \gamma_U]$ is set so that γ_L is the π_0 percentile of $\tilde{w}_{t-1}$ and γ_U is the $(1-\pi_0)$ percentile. This imposes constraint (5). The statistics are computed with heteroskedasticity-consistent covariance matrix estimates. However, the SupLM statistic has a nonstandard asymptotic distribution. As discussed in Hansen and Seo (2002), the fixed regressor bootstrap of Hansen (1996, 2000) can be used to calculate asymptotic critical values and p-values together with the residual bootstrap technique. This framework seems to be the most suitable to consider potential nonlinear effects in the long-run relationship between our variables of interest if the testing procedure described above produces evidence for this.

4. Data

The annual time series data for Italy's central government spending and GDP considered in this work come from a new quantitative study conducted in Italy over three decades and radically changing the interpretation of Italian development.

The reconstruction of new Italian historical national accounts obtained the GDP at current equivalents and real per capita - 2010 prices. It was developed by the Bank of Italy together with academics from other institutions. This documentation was presented on occasion of the 150th anniversary of Italy's unification and recently published in Baffigi (2015).[3] The new official series formed the basis for the recent volume *Oxford Handbook of the Italian Economy since Unification* (Toniolo, 2013).

For the same anniversary, Italy's State General Accounting Department published a special issue on specific economic and functional items of public spending (RGS, 2011).[4] Data on total government expenditure and the national defence item are drawn from this source. They are at current prices and terminate in 2009. Spending refers to the total payments disbursed in the year, which were obtained from the final state budget.[5] Population data are from the Ricostruzione della popolazione residente e del bilancio demografico database (Istat, 2012).

Our final sample includes the above mentioned up-to-date series over the period 1862–2009.[6] Our data differ from those of the two most recent papers testing WL for Italy. Magazzino (2012) relies on the Informative Public Base (IBP), a database developed by the Bank of Italy that covers the shorter 1960–2008 time span and refers to the expenditure of the Italian public administration as a whole, including not just the expenditure of Italy's central government but also the expenditure of local government bodies (regional, provincial and municipal administrations). Our series also differ from those of Kuckuck (2014) who uses not-revisited data drawn from Mitchell, B.R. (2007) for the years 1850–1995 and from Eurostat for the years 1996–2010. These data are also provided by Italy's State General Accounting Department, but, unlike ours, they refer to the expenses accrued and not to actual payments during the fiscal year. Our GDP data again differ from those of Kuckuck (2014), who for the years 1850–1995 relies on not-revisited national income series published by Mitchell, B.R. (2007). Mitchell's data are derived from Istat's (1957) first series of Italy's national accounts.[7]

Fig. 1 depicts the evolution of the ratio of total government expenditure to GDP compared to real per capita GDP on the full sample.[8] It shows that between 1862 and the mid-1890s, total government spending and real per-capita GDP followed very similar trends. During this period the investments in railways were particularly significant and spending on education and culture also increased constantly from Unification to WWI. In the first fifty years following Unification, Italy's total government spending in real terms increased slowly, on average. Between 1862 and 1913 total government spending was around 10% of GDP, but participation in WWI led to a drastic increase in total government spending to 35% of GDP. In the years following the conflict, government spending dropped again sharply, in 1926 settling to the pre-war values compared to the GDP: defence spending dropped, while investments in public works and other economic interventions resumed. Participation in WWII led to an another drastic increase in government spending, from just over 10% to more than 45% of GDP. After defeat in WWII, Italy was prohibited from reconstructing its own independent military forces. This led to a drastic reduction in defence spending, as can been seen in Fig. 2. On the other hand, the economic and social components of expenditure started to increase: infrastructures, welfare, and redistribution by the state. Pressure from the extension of suffrage (universal suffrage was introduced in 1946) and an unprecedented wave of social conflicts from the end of the 1960s, led to a progressive expansion of welfare services to new social categories in Italy until a universal welfare system was introduced in 1978. Moreover, Italy's pro-American stance during the Cold War and the possibility that the Italian Communist Party (the largest Communist party in the Western world) might organise a revolution meant that a large proportion of government expenditure was allocated to national security. In the 1970s this expenditure was further increased to counter political terrorism. Thus, between the end of WWII and 1963 public spending remained well below 25% of GDP while from the mid-1970s state spending began to grow more rapidly than the GDP, reaching a peak of 44% in 1986 (see Fig. 1). During recent years the level of expenditure has been more stable. Between 1980 and 2009 total government expenditure actually declined in proportion to GDP. Starting from 1993, with a view to Italy's joining the single European currency, the imbalance in the national accounts began to be countered and clear results seen in 1995, achieving decisive progress in 1997, when the deficit fell. A preliminary analysis of the series suggests a long-run relationship between total government spending and national income with some wide deviations during WWI and WWII. This supports our idea of taking into account any large but transitory deviations in the general long-run development of the variables.

5. Empirical results

In this Section we explore the data available. First, some preliminary investigations are conducted on the (linear and threshold) stationarity of the series and the existence of a cointegrating

[3] The GDP series is available at the Bank of Italy's website at the URL: https://www.bancaditalia.it/pubblicazioni/collana-storica/pil-storia-italia/index.html.

[4] The public spending series are available at the website of Italy's General Accounting Department at the URL: http://www.rgs.mef.gov.it/VERSIONE-I/Pubblicazioni/Pubblicazioni_Statistiche/La-spesa-dello-stato/.

[5] From 1884 to 1964 Italy's fiscal year ran from July 1st to June 30th. The data were attributed to solar years by adding half of the expenditures disbursed in two consecutive fiscal years and assuming an equidistribution of expenditure over each fiscal year.

[6] Even though the GDP series runs until 2013, our final sample stops in 2009 given that the General Accounting Department does not provide an updated database with the same aggregation criteria as the RGS publication (2011). However, we are confident that four additional observations would not affect the results.

[7] Note that these series were heavily flawed by Cohen and Federico (2001) so that the Bank of Italy launched a project for the through reconstruction of the national accounts on which we rely.

[8] See Tanzi and Schuknecht (2000) for a discussion of the role of government spending in the main industrialized economies during the 20th century and Cohen and Federico (2001) for an in-depth discussion of the Italian case.

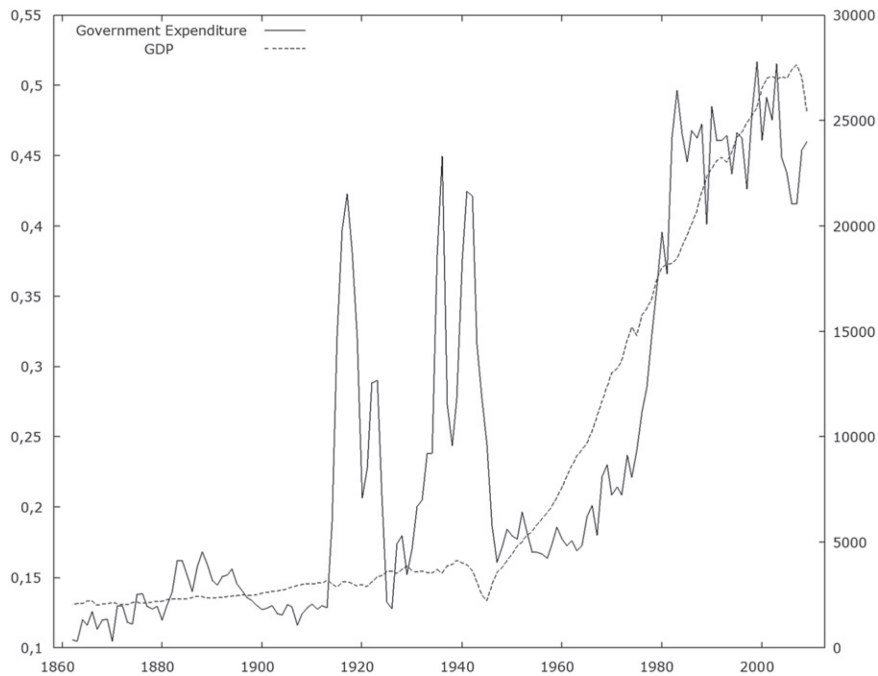

Fig. 1. Total government expenditure as a share of GDP and real per capita GDP (1862–2009).

relationship. After testing nonlinear cointegration, the model is estimated with description of the empirical results.

5.1. Preliminary analysis

The stationarity of the two series of interest were investigated by applying alternative unit root tests. The first test is the standard augmented Dickey-Fuller test which considers a null hypothesis of a unit root against the alternative of stationarity. The second is the modification of the above test proposed by Elliott (1999). The last is the KPSS test proposed by Kwiatkowski et al. (1992) for the null hypothesis of stationarity versus the alternative of non-stationarity. The results are reported in Table 1. Both tests suggest that government expenditure and real per capita GDP are realizations of $I(1)$ processes.

However, it is known that standard unit root tests have almost no power when the alternative is nonlinear (Basci and Caner, 2005). Therefore, further tests are conducted that allow for the joint consideration of nonlinearity (threshold) and nonstationarity (unit roots). The Caner and Hansen (2001) test was applied (for a short description, see Chevallier (2011)). If the null hypothesis H_0 holds, the series is nonstationary with a classical unit root. The alternative can be

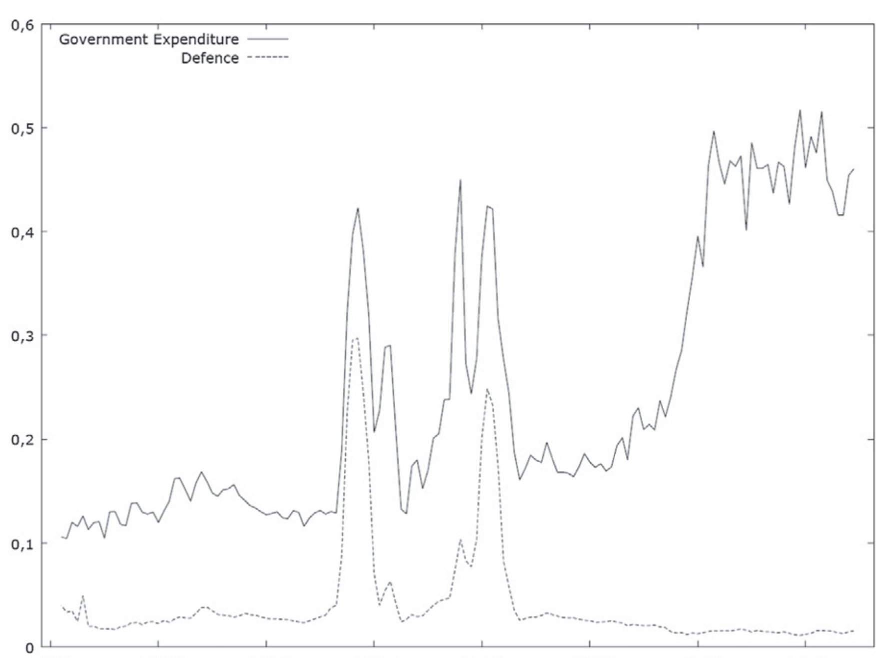

Fig. 2. Total government expenditure as a share of GDP and national defence expenditure as a share of GDP (1862–2009).

Table 1
Unit root and stationary tests for total government expenditure (denoted by g) and real per capita GDP (denoted by y). The symbol Δ denotes the first-difference transformation of the series. The symbol *** indicates significance at 99% confidence level, while the symbol ** indicates significance at 95% confidence level. These symbols refer to the choice of not rejecting the null hypothesis for both the ADF (presence of unit root) and KPSS (stationarity) tests. Numbers in square brackets refer to the selected lag order for each statistic. Finally, t_μ and t_τ correspond to test statistics where the auxiliary regression contains a constant and a constant and a trend, respectively; η_μ and η_τ are test statistics for level and trend stationary, respectively.

Series	ADF t-tests		ADF-GLS tests		KPSS tests	
	t_μ	t_τ	t_μ	t_τ	η_μ	η_τ
g	−1.541 [2]	−2.977 [2]	−0.511 [2]	−3.002 [2]	3.349*** [2]	0.206** [2]
Δg	−9.809*** [1]	−9.775*** [1]	−9.405*** [1]	−9.471*** [1]	0.025 [1]	0.024 [1]
y	0.443 [2]	−1.785 [2]	1.479 [2]	−1.091 [2]	6.831*** [2]	1.046*** [2]
Δy	−7.374*** [1]	−7.461*** [1]	−7.653*** [1]	−7.485*** [1]	0.372 [1]	0.127 [1]

stated in two versions: a first H_1 with a stationary threshold autoregressive pattern and a second H_2 with partially stationary threshold process. The test statistics are a one-sided Wald R_{1T} in the first case and t-ratios t_1 and t_2 in the second case. Limit distributions and critical values are tabulated in Table 3 from Caner and Hansen (2001). Even though the unit root tests have an asymptotic bound distribution, benefit is achieved with a bootstrap procedure in finite sample. Table 2 reports asymptotic and bootstrap p-values for the one-sided Wald R_{1T} and t-ratio t_1, t_2 tests. Regarding the R_{1T} statistics, the one-sided Wald test (unit root vs threshold stationary model) is rejected at 5% level only for the government expenditure series, suggesting a classical unit root for the real GDP series. Moreover, for government expenditure the individual t_1 and t_2 ratios suggest rejection only for the second regime. Hence, it can be concluded that government expenditure follows a partially stationary threshold process while real GDP contains a unit root in the classical sense.

In the following analysis the first difference of both series is considered, and preliminary summary statistics are given in Table 3. Normality tests reject the null of normality in the data, and this could be partially due to temporal dependence in the moments of the series or to the presence of nonlinearities in the data.

Before proceeding with the analysis, a search was conducted for the existence of a cointegrating relationship between the two variables of interest. Note that, as a preliminary condition for cointegration, it was checked that the two time series are integrated by the same order ($I(1)$). Next, the linear Johansen cointegration rank tests were implemented (Johansen, 1988), using 2 lags in the VAR, as suggested by the BIC criterion, and including an unrestricted constant. As shown in Table 4, these preliminary tests easily reject the null hypothesis of no-cointegration, indicating the presence of one cointegrating relation. Therefore, the linear VECM was estimated and the results are shown in Table 5. Here the estimates indicate bidirectionality in the two variables with no evidence of WL for this country. However, this could be due to a misspecification of the model due to threshold-type nonlinearities present in the data, as shown by previous tests. Hence, threshold cointegration techniques are applied in the following.

5.2. Threshold cointegration

Firstly, the presence of threshold effect under the null hypothesis of linear cointegration is explicitly tested for the complete bivariate specification. The fixed regressor bootstrap and residual bootstrap methods were used, and both were simulated using 5000 bootstrap replications. A lag length of $\ell = 1$ in Eq. (3) was selected, based on AIC and BIC criteria. Table 6 reports the test results for the linear versus nonlinear cointegration hypothesis, together with threshold and cointegrating parameter estimates. The resulting LM statistic computed as a function of the threshold parameter estimate γ is plotted in Fig. 3. The fixed regressor bootstrap method rejects the null at 5% confidence level and the residual bootstrap method at

Table 2
Threshold unit root tests R_{1T}, t_1 and t_2 as described in Caner and Hansen (2001). R_{1T} tests H_0 vs H_1, t_1 and t_2 test H_0 vs H_2. The trimming region is set as [0.150.85] and the delay parameter m is estimated by minimizing the Sum of Squared Errors (SSE). Bootstrap p-values are computed from 10,000 replications. The symbol ** denotes rejection of the hypothesis at the 5% level.

Series	R_{1T}	t_1	t_2
Asymptotic p-values			
g	0.012**	0.837	0.008**
y	0.457	0.282	0.952
Bootstrap p-values			
g	0.016**	0.469	0.007**
y	0.320	0.147	0.684

Table 3
Summary statistics on first-differenced series of total government expenditure (g) and real per capita GDP (y). J–B and S–W denote the Jarque-Bera and Shapiro-Wilk tests for the null of normality, respectively. The symbol *** indicates a 1% significance level.

Statistics	Δg	Δy
Mean	0.010	0.017
Standard deviation	0.139	0.049
Skewness	−0.088	0.106
Kurtosis	3.645	10.118
J–B	81.551***	627.41***
S–W	0.922***	0.840***

Table 4
Johansen cointegration rank tests (CE stands for Cointegrating Equation) based on VAR(2) with unrestricted constant. The symbol *** denotes rejection of the hypothesis at the 0.01 level. Both trace test and max-eigenvalue tests indicate 1 cointegrating equation at 0.01 level.

Series	g_t	y_t	
Trace test			
No. of CE(s)	Eigenvalue	Trace statistic	p-value
None	0.1269	19.830	0.0092***
At most 1	0.00006	0.0091	0.9240
Maximum eigenvalue test			
No. of CE(s)	Eigenvalue	Max-eigen statistic	p-value
None	0.1269	19.821	0.0048***
At most 1	0.00006	0.0091	0.9240

Table 5
Estimates of the linear VECM(1) with unrestricted constant for government expenditure (g) and real per capita GDP (y). Standard errors are shown in parentheses. The symbol *** denotes 99% significance level and the symbol ** 95% significance level.

Variables	Δg_t	Δy_t
Cointegrating vector	1.0000	−0.5124
	(0.0000)	(0.0718)
Intercept	−0.6953***	−0.1909**
	(0.2039)	(0.0741)
w_{t-1}	−0.1186***	−0.0349***
	(0.0342)	(0.0124)

Table 6
Lagrange Multiplier (LM) test for threshold cointegration between government expenditure and real per capita GDP. The number of grid points for threshold and cointegrating vector is equal to 80. For p-values, the number of bootstrap replications is set to 5000.

Estimates	
Threshold parameter estimate (γ)	−5.532
Cointegrating parameter estimate (β)	0.501
Lagrange multiplier threshold test	
sup LM value	20.082
p-Value of fixed regressor bootstrap	0.024
p-Value of residual bootstrap	0.009

Table 7
Estimates of the threshold VECM for government expenditure (g) and real per capita GDP (y). Eicker-White standard errors are shown in parentheses. In Wald test diagnostics, the null hypothesis is equality of the dynamic coefficients and of the coefficients on the error correction terms across the two regimes, respectively. The symbol *** denotes a 99% significance level and the symbol ** a 95% significance level.

Variables	1st regime (87% obs)		2nd regime (13% obs)	
	Δg_t	Δy_t	Δg_t	Δy_t
Intercept	−0.136**	−0.015	−0.834***	−0.162
	(0.055)	(0.014)	(0.219)	(0.089)
w_{t-1}	−0.781**	−0.079	−4.395***	−0.839
	(0.334)	(0.081)	(1.143)	(0.458)
Δg_{t-1}	0.105	0.029	0.631***	0.086
	(0.154)	(0.028)	(0.094)	(0.061)
Δy_{t-1}	−0.350	0.495***	−0.019	0.179
	(0.198)	(0.074)	(0.433)	(0.307)
Wald tests				
Equality dynamic coefs.	19.661	(p-Value: 0.001)		
Equality EC coefs.	18.923	(p-Value: 0.001)		

1%. Thus, the threshold cointegration model seems more appropriate for our data than a linear model. In fact, ignoring asymmetric adjustment may lead to biased inferences and misleading conclusions. Hence the error correction mechanism differs depending on deviations from equilibrium below or above the threshold parameter. The first regime (say, "normal") corresponds to $g_{t-1}-0.501y_{t-1} \leq -5.532$ while the second regime (say, "extreme") corresponds to $g_{t-1}-0.501y_{t-1} > -5.532$. We also observe that 87% of all the observations belong to the first regime and the remaining 13% to the second regime.

Estimation is performed by MLE following the grid-search algorithm proposed by Hansen and Seo (2002) over an 80 × 80 grid of the parameters β and γ. Table 7 reports estimated threshold VECM values and the following results. The first is that during "normal" periods lagged values of real per capita GDP significantly influence the dynamic behaviour of the Italian economy, while in the "extreme" regime government expenditure during the previous period tends to matter most. The second result concerns the coefficients of the error-correction term w_{t-1} in the two regimes. This term suggests causality (at the level of weak exogeneity) running from economic growth to government activity: the negative and statistically significant adjustment parameters in both government equations are evidence that WL holds. In this sense the GDP contains relevant information to predict the long-run path of government spending but not viceversa. Also the magnitude of the response of government expenditure is between 9 ("normal" regime) and 5 ("extreme" regime) times greater than the coefficient in the GDP equation. The diagnostics reported at the bottom of Table 7 reinforce the evidence for nonlinearity given that the null of equality of the dynamic coefficients as well as equality of coefficients in the error-correction term are strongly rejected. Finally, the estimated long-run elasticity between government expenditure and national income is significantly greater than zero (p-value = 0.009) suggesting that government expenditure is income elastic, i.e. a superior good, over the entire sample.

To allow visual interpretation of these results, the error correction mechanism is illustrated in Fig. 4. It can be noted that the strong error-correction effect for the two variables is depicted on the right-hand side of the estimated threshold. On the contrary, it shows a flat near-zero effect for real per capita GDP and a slightly greater effect for government expenditure on the left-hand side. Asymmetry shows a stronger error-correction effect in the "extreme" regime compared to the "normal" regime due to the government spending series.

Furthermore, the timing of the regime shift was investigated. The estimated regime classification according to the threshold VECM is visually presented in Fig. 5. Our estimated model suggests that the economy happened to be in the "extreme" regime between 1915 and 1919 (with some follow-up in 1921–1923) and between 1934 and 1945. These facts are in line with the preliminary assessment

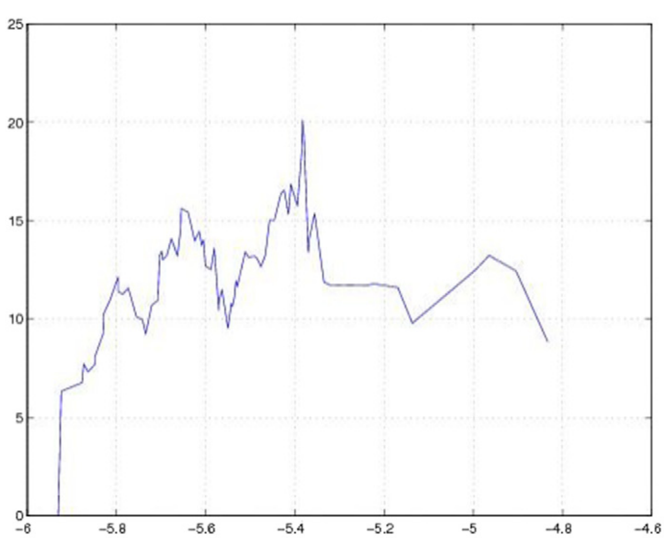

Fig. 3. Lagrange Multiplier (LM) statistic for the bivariate (government expenditure and real per capita GDP) threshold cointegration model as a function of the threshold parameter γ.

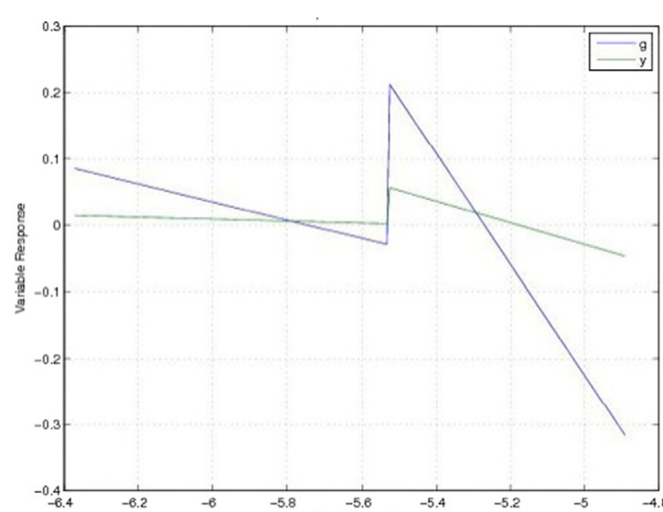

Fig. 4. Variable response to error-correction. Variables are government expenditure (g) and real per capita GDP (y).

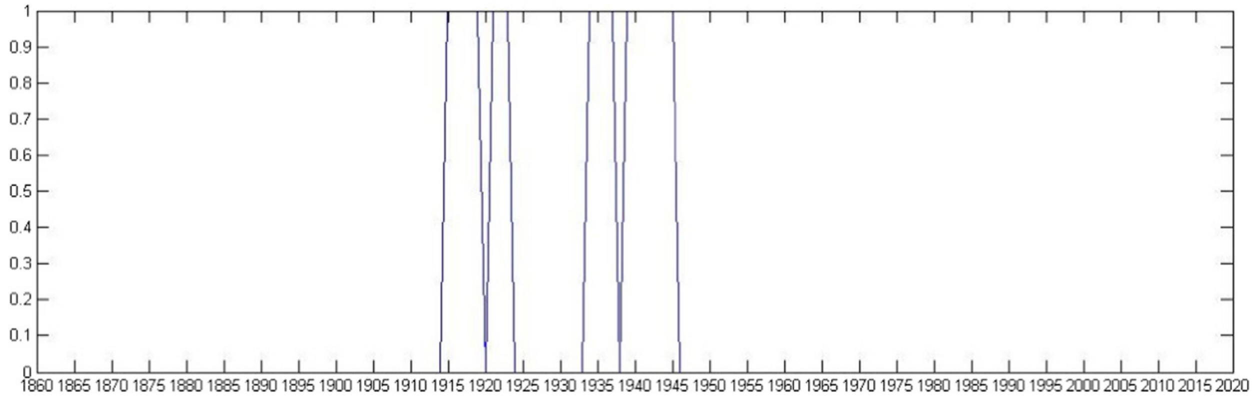

Fig. 5. Timing of the realization of "extreme" regime obtained from the threshold VECM in government expenditure (g) and real per capita GDP (y).

in Fig. 2, highlighting a drastic increase in military spending during these time spans associated with wars. Hence, the source of disequilibrium in the "extreme" regime must be ascribed to large transitory increases in defence spending, not necessarily induced by large variations in income. This generates an abnormal adjustment of government spending to return to the long-run equilibrium.[9] On the contrary, during "normal" periods, variations in public spending generally derive from changes in national income.

In conclusion, when these asymmetric adjustments are taken into account, all these results support WL for the Italian economy from 1862 to 2009.

6. Robustness

Two checks were conducted to confirm the accuracy of our results. Firstly, the linearity of WL for public spending was verified when the "extreme" periods are excluded from the data. Secondly, the linearity of this relationship was tested subtracting military spending from the total expenditure series.

Specifically, the linearity versus nonlinearity LM test was run for the two remaining subsamples after exclusion of periods that induce nonlinearity. Applying the fixed regressor and residual bootstrap to the subsample 1862–1914 (1946–2009, respectively) generated p-values of 0.39 and 0.054 (0.092 and 0.052, respectively), which do not reject the null hypothesis of linearity at a 5% significance level.

Since the model specifically addresses wars, public spending on defence was subtracted from total government expenditure to remove the source of nonlinearity in the data. A new variable was defined as the natural logarithm of the ratio between total government expenditure minus national defence expenditure over nominal GDP. The linearity LM test was run again for the bivariate model and the p-values from the fixed regressor and residual bootstrap were 0.26 and 0.105, respectively. This means that the linearity hypothesis at a 10% significance level is not rejected, specifically indicating military spending as a driving force for government expenditure. This evidence reinforces the conclusion outlined in Section 5. Hence, the source of nonlinearities is large but transitory and it is simply due to abnormal national defence spending.

7. Conclusion

This paper tests the long-run tendency for Italian total government expenditure to grow relative to per capita GDP over the period 1862-2009. Evidence was found for a threshold cointegrating relationship between these variables, which turns out to be consistent with WL, given the different adjustment speeds to the long-run path. Asymmetric error-correction effects identify two different regimes and the WWI and WWII periods perfectly match one of them. The abnormal response of government spending in this "extreme" regime was due to temporary increases in defence spending during the wars. This implies a hyper-adjustment of total government spending to return to the long-run equilibrium. On the basis of these results, we conclude that the model for public spending is basically linear and consistent with an expanding government sector as the economy progresses and the nonlinearities are transitory.

The above result suggests that the Italian governments invested constantly in public expenditure from 1862 to 2009. This may have served to strengthen the Italian state in the face of potential external and internal threats (i.e. the World Wars, the Cold War and socio-political turmoil in the post-war period) and in response to the growing demand of society for public services. However, this might not have been the most effective way to prompt economic growth, in particular if the composition of government expenditure shifted towards less productive spending, in terms of innovation and economic growth. For example, social spending on health and pensions increased enormously in response to social disruption and the demands of the electorate during the 1960s and 1970s becoming very high compared to spending on education and scientific research. As a result, despite the growth of Italy's human capital over the past thirty years, in terms of both secondary and tertiary education the country has not managed to close the long-standing gap separating it from the other OECD economies (Tanzi and Schuknecht, 2000, Visco, 2014). Consequently, Italian governments could manage investments to stimulate long-term growth in this area. For example, the creation of a national innovation system and substantial investments in schools and universities could play a key role towards improving the quality of human capital.

For future research the following issues might be addressed. Firstly, in the context of nonlinear models, other forms of nonlinearity could be explored, like for example (and as kindly suggested by a referee) a quadratic effect in the relationship between public expenditure and growth. Secondly, it would be useful to study the nonlinear role of different items of public spending in economic growth, i.e. welfare. Finally, it would be worthwhile investigating the role of military spending from a comparative international perspective based on reliable historical data.

Acknowledgments

The authors are grateful to the Editor-in-Chief Prof Paresh Narayan and the three anonymous referees for their very useful

[9] Note that our results do not confirm the empirical study by Barro (1981) on temporary increases in US defence spending: in his work these shocks also induce marked increases in output.

suggestions and remarks which improved the final version of the paper. The authors gratefully acknowledge financial support from the PRIN-MIUR Grant 2010J3LZEN-003 and the research grant FAR 2014 (awarded by the University of Modena and Reggio Emilia, Italy).

References

Baffigi, A., 2015. Il PIL per La Storia d'Italia, Collana Storica Della Banca d'Italia Statistica. Roma, Marsiglio Editore.
Barro, R.J., 1981. Output effects of government purchases. J. Polit. Econ. 89 (6), 1086–1121.
Basci, E., Caner, M., 2005. Are real exchange rates nonlinear or nonstationary? Evidence from a new threshold unit root test. Studies in Nonlinear Dynamics & Econometrics 9 (4).
Baumol, W.J., 1967. Macroeconomics of unbalanced growth: the anatomy of urban crisis. Am. Econ. Rev. 57 (3), 415–426.
Bohl, M.T., 1996. Some international evidence on Wagner's law. Public Finance/Finances Publiques 5, 185–200.
Caner, M., Hansen, B.E., 2001. Threshold autoregression with a unit root. Econometrica 69 (6), 1555–1596.
Chevallier, J., 2011. Evaluating the carbon-macroeconomy relationship: evidence from threshold vector error-correction and Markov-switching VAR models. Econ. Model. 28 (6), 2634–2656.
Cohen, J., Federico, G., 2001. The Development of the Italian Economy 1820–1960. Cambridge University Press, Cambridge.
Durevall, D., Henrekson, M., 2011. The futile quest for a grand explanation of long-run government expenditure. J. Public Econ. 95 (7), 708–722.
Elliott, G., 1999. Efficient tests for a unit root when the initial observation is drawn from its unconditional distribution. Int. Econ. Rev. 40 (3), 767–784.
Ghate, C., Zak, P.J., 2002. Growth of government and the politics of fiscal policy. Struct. Chang. Econ. Dyn. 13, 435–455.
Hansen, B.E., 1996. Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64, 413–430.
Hansen, B.E., 2000. Testing for structural change in conditional models. J. Econ. 97 (1), 93–115.
Hansen, B.E., Seo, B., 2002. Testing for two-regime threshold cointegration in vector error-correction models. J. Econ. 110 (2), 293–318.
Henrekson, M., 1993. Wagner's law - a spurious relationship? Public Finance / Finances Publiques 48, 406–415.
Islam, A.M., 2001. Wagner's law revisited: cointegration and exogeneity tests for the USA. Appl. Econ. Lett. 8 (8), 509–515.
Istat, 2012. "ricostruzione della popolazione residente e del bilancio demografico".
Johansen, S., 1988. Statistical analysis of cointegration vectors. J. Econ. Dyn. Control. 12 (2), 231–254.
Kuckuck, J., 2014. Testing Wagner's law at different stages of economic development. FinanzArchiv: Public Finance Analysis 70 (1), 128–168.
Kwiatkowski, D., Phillips, P.C., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? J. Econ. 54 (1), 159–178.
Lindert, P.H., 1994. The rise of social spending, 1880–1930. Explor. Econ. Hist. 33 (1), 1–34.
Lindert, P.H., 2004. Growing Public: Volume 1, the Story: Social Spending and Economic Growth Since the Eighteenth Century. Cambridge University Press, Cambridge and New York.
Lindert, P.H., 2004. Growing Public: Volume 2, Further Evidence: Social Spending and Economic Growth Since the Eighteens Century. Cambridge University Press, Cambridge and New York.
Magazzino, C., 2012. Wagner versus Keynes: public spending and national income in Italy. J. Policy Model 34, 890–905.
Meltzer, A.H., Richard, S.F., 1981. A rational theory of the size of the government. J. Polit. Econ. 89, 914–917.
Mitchell, B.R., 2007. International Historical Statistics 1750–2005. 6th ed., vol. 3. Palgrave Macmillan, New York.
Narayan, P.K., Prasad, A., Singh, B., 2008. A test of the Wagner's hypothesis for the Fiji islands. Applied Economics 40 (21), 2793–2801.
Narayan, S., Rath, B.N., Narayan, P.K., 2012. Evidence of Wagner's law from Indian states. Econ. Model. 29 (5), 1548–1557.
North, D.C., 1985. The growth of government in the United States: an economic historian perspective. Journal of Public Economic 28, 383–399.
Oxley, L., 1994. Cointegration, causality and Wagner's law: a test for Britain 1870–1913. Scottish Journal of Political Economy 41, 286–298.
Peacock, A.T., Scott, A., 2000. The curious attraction of Wagner's Law. Public Choice 102 (1-2), 1–17.
Persson, T., Tabellini, G., 1990. Macroeconomics Policy Credibility and Politics. Harwood Academic, London.
RGS, 2011. La Spesa Generale Dello Stato Dall'unitÁ d'Italia. Anni 1862–2009. Ragioneria Generale dello Stato, Roma.
Subervie, J., 2011. Producer price adjustment to commodity price shocks: an application of threshold cointegration. Economic modelling 28 (5), 2239–2246.
Tanzi, V., Schuknecht, L., 2000. Public Spending in the 20th Century. Cambridge University Press, Cambridge.
Thornton, J., 1999. Cointegration, causality and Wagner's law in 19th century Europe. Appl. Econ. Lett. 6, 413–416.
Toniolo, G., 2013. The Oxford Handbook of the Italian Economy since Unification. Oxford University Press, Oxford.
Visco, I., 2014. Investire in Conoscenza. Il Mulino, Bologna.
Wagner, A., 1883. Finanzwissenschaft, 3Rd Edition, Leipzig. Partly Reprinted in Classics in the Theory of Public Finance, Edited by R.A. Musgrave and A.T Peacock (1958). Macmillan, London.