

IS THERE TOO MUCH GOVERNMENT IN DEVELOPED COUNTRIES? A TIME-SERIES ANALYSIS OF 24 OECD-ECONOMIESⁱ

Sebastian COLL¹

DOI : 10.1515/JHEEC-2015-0001

The economic consequences of public sector oversizing have been analyzed in hundreds of scientific contributions. The empirical branch of this literature tries to measure these effects through the use of three main tools: cross-sectional analysis of countries' samples, time-series analysis of individual economies, and panel-data analysis. Most authors conclude that the current size of the public sector is too big in their analyzed economies, thus leading to reduced output. For different reasons, however, the results of both cross-sectional and time-series analyses are open to criticism. This article aims to check the validity of the converging conclusions obtained by time-series analyses. To this end, a simultaneous-equation model is designed, which evades the critiques addressed to prior contributions. Application of this model to 24 OECD countries during the 1975-to-2007 period suggests that at least most of these have indeed let their public sector grow beyond the level that is optimum for their economic performance.

Keywords: Armey Curve, Size of Public Sector, Public Choice, Public Finance, Economic Growth, Simultaneous Equations, Time Series.

JEL Classification: H5; O4; O5; C3

¹ Universidad de Cantabria, Departamento de Economía. Avenida de los Castros s.n., 39005 Santander (Spain). (colls@unican.es), and Universidad Internacional Menéndez Pelayo. Isaac Peral 23, 28040 Madrid (Spain). (scoll@uimp.es).

1. Introduction

The rapid increase in the size of the public sector has been one of the most salient economic features of 20th century. According to Tanzi and Schuknecht (2000:6-7), public expenditures as a share of GDP more than tripled during the century in a representative sample of industrialized countries. The most rapid increase has corresponded to the heading “transfers and subsidies”, indicative of the re-distributive action of government, whose share in GDP has increased from about 1 per cent towards the end of 19th century to about 21 per cent one century later, according to the same source. As a consequence, the public sector has increasingly become a re-distributor of wealth created by the private sector of the economy.

From this it follows that the growth of the public sector should eventually come to a halt, imposed by the limited extent of the privately created wealth available for re-distribution. Prior to meet that limit, the distorsionary effects of taxes and other sources of public revenue on resource allocation should likely lead to a reduction of the social product, as compared to its potential value. This reflection has led to the notion of an optimum size of the public sector from the point of view of economic growth, an idea that was formulated by a number of authors towards the end of the 20th century. According to Armey (1995), the relationship between GDP and the government share follows the pattern of an inverted U. About the same time, Barro, Rahn and Scully put forward their hypotheses about another inverted-U pattern in the relationship between the *rate of growth* of GDP and the government share.

These patterns would result from the interplay of two opposing forces. On the one hand, public expenditure on those goods which are subject to market failure (i.e., on classical public services) contributes to the generation of private output. This beneficial effect is, however, decreasing on the margin due either to diminishing marginal returns of classical public services, or to displacement of public expenditure to other, not so beneficial targets. In turn, the public revenues necessary for the backing of public expenditures have an opportunity cost in terms of foregone private production and investment. As public revenues increase, the forgone private investment [consumption] includes more and more profitable [valuable] allocations, with the result that the opportunity cost of public expenditure increases on the margin. When we detract a cost that increases more than proportionally with the size of the public sector from an output that increases less than proportionally with that size, the resulting net effect should follow the inverted-U pattern referred to above. Note, however, that the current practice in national accountancy, consisting of valuing the product of the public sector through its cost, tends to more or less mask this pattern.

The hypotheses put forward by Armey, Barro, Rahn and Scully have been submitted to test many times. In the beginning, the most widely used technique was regression analysis of cross-sections of countries. Nonetheless, as indicated in the next section, the government share is but one determinant of GDP or of its rate of growth. As there are

many other determinants whose behavior changes from country to country, cross-sectional analysis needs to control for too many variables. By contrast, to the extent that some of these determinants tend to remain more or less stable within each country, time series analysis seems to provide, in principle, a more suitable approach. On the other hand, the relationship between GDP (or its rate of growth) and the size of the public sector doesn't remain unidirectional in the course of time, a fact that poses a difficulty to this type of analysis. In more recent time panel-data analysis has gained partisans, in the belief that it reduces the seriousness of the problems affecting the other two approaches. This belief is not without ground, so that the present state of the art heavily rests on the results of panel-data analyses.

Be it as it may, this contribution follows the time-series approach, in an attempt to overcome the endogeneity issue by means of the formulation of a simultaneous equation model. Yet, this approach, in combination with the nature of the data, allows to estimating only the short-term effect of the government share on GDP. Since this effect is necessarily less visible than the long-term effect, our test is rather strict. The model devised here is applied in succession to 24 advanced economies during the period that runs from 1975 to 2007. In about half those countries the results strongly suggest that the net growth effect of the increase in public expenditure is indeed negative. In the rest of cases the evidence is less compelling, although it points in the same direction. All in all, it can be said that our results are in line with those of prior time series analysis, as well as with the mainstream of all other empirical approaches.

The remaining of this contribution is organized as follows. Section 1 reviews the literature. Section 2 develops the model. In section 3 the benchmark estimates are presented. Section 4 extends the model to include more variables, and checks the robustness of the estimates obtained for each country. Section 5 concludes.

2. A review of the literature

An exhaustive review of the literature on the relationship between GDP growth or levels and the government share is not possible in the framework of this article. And perhaps it is an impossible endeavor altogether, as the academic contributions dealing with this topic can be counted by hundreds nowadays².

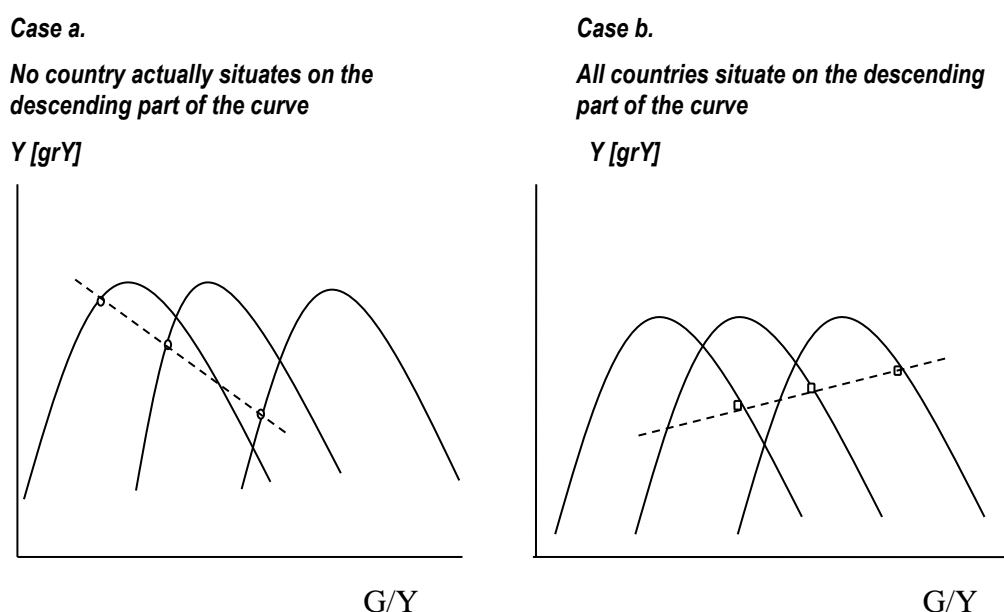
Until 2000, most articles in the field used cross-sectional analysis. With the passing of time some pioneer panel-data analyses also appeared, although these tended initially to cover only a few years. Therefore, most variance continued to stem from inter-country differences.

One problem of cross-country analysis (or, for the case, of panel data analysis using only a few years) is that of ensuring the *cæ. par.* condition. Any variable, other than the

² Review articles were published by Temple (1999), Tanzi and Shuknecht (2000), Poot (2000), Easterly and Levine (2001), Leach (2003), and Nijkamp and Poot (2004). More recently, Facchini and Melki (2012) have offered a good deal of references to individual titles. Berg and Henrekson (2011) focus on recent articles only.

government share, that has an effect on GDP growth and presents different values in different countries causes the exact location of the Armey or Scully curves to differ on the plane among countries. Cross-sectional analysis, however, estimates *one single* curve for all of them. At best, this oversimplifies the results, and, in the worst scenario, it may bias them. For this reason, some authors have expressed serious doubts that this type of analysis cannot help to clarifying the point at issue. This is, for instance, the position taken by Slemrod (1998:102), and this author's contention is shared by Arcand and Dagenais (1995) and by Mueller (2003:550). As indicated by the latter, and illustrated in figure 1, the absence of a common curve for the countries included in the cross section could possibly lead us to conclude that the relationship between income and public expenditure is positive where it is in fact negative in every country, or the opposite.

Figure 1: Two Possible Cases for a "Sample" of Three Countries, Each Presenting Its Own Armey [Scully] Curve



Note: The dotted lines indicate the functions of Y [grY] to G/Y as estimated via cross sectional analysis.

Source: Own elaboration.

The reservations expressed by these authors appear to be confirmed by a certain divergence among the results of this literature, depending on particular specifications, countries' samples or time periods chosen by each author. This notwithstanding, it is fair to add that most results favor the hypothesis of an inverse relationship between the government share and the level or the rate of growth of GDP. I have collected a sample of almost one hundred cross-sectional analyses carried on between the 1980s and the early 2000s³. In this, 72.7% of the reported coefficients exhibit minus signs vs. 27.3% of positive ones. The corresponding proportions for OECD or industrial countries alone are

³ The list is available on request, and includes as well those panel-data analyses which use very few time periods.

74.4% and 25.6% respectively. This small difference among the two samples suggests that the customary assumption according to which a small public sector amounts to provision of genuine public goods only is perhaps violated in practice. The proportions don't show significant changes when we restrict our examination to the results of the "best practice" contributions, defining this concept as the use of the production-function approach. However, the proportion of negative coefficients increases when one looks at the significant ones only. When doing this, the share of coefficients exhibiting minus signs increases to 88.6% for the whole sample, and to 93.8% when we consider developed countries only. But the fact remains that 11.4% of significant results in the former case and 6.2% in the latter exhibit the "wrong" sign according to the hypotheses submitted to test. Moreover, less than one estimate out of three (27.3%) succeeds in yielding both significant and negative coefficients for the variable of interest in the whole sample, and the corresponding figure for the subsample of developed countries is just 36.6%.

Nijkamp and Poot (2004), and Facchini and Melki (2012) have analyzed comparable samples of contributions, although they don't distinguish between cross-sectional and other types of analyses. Their distributions of negative and positive coefficients, as well as of significant and insignificant ones are, however, similar to those reported above. This similarity is higher in the case of Facchini and Melki (2012), whose criteria to classify results as favorable, unfavorable or ambiguous is closer to mine. Given the problems that surround cross-sectional analysis on this topic, a possible conclusion to be drawn is that the positive signs that are sometimes obtained correspond to defective specifications or data samples. However, other interpretations are also legitimate.

Since one of the potential sources of divergence is the set of control variables included in each estimation attempt, it is interesting to check how much influence these differences may exert. This idea has led to the use of robustness tests consisting of systematically altering the set of control variables until exhausting their possible combinations. The robustness test currently applied to the relationship between the government share and economic performance is the so-called *extreme bound analysis* (EBA)⁴. In the earliest and most cited application of this test to the proposed determinants of economic growth, Levine and Renelt (1992) discarded as non-robust all the proposed ones with the sole exception of the investment rate, and another similar exercise carried out by Levine and Zerbos (1993) came to similar results. An improved version of EBA devised by Sala i Martin (1997) found that several among the determinants proposed by growth theorists turned to be robustly related to the latter; however, no fiscal variable entered this group.

⁴ In this type of analysis, those explanatory variables whose inclusion seems to be well grounded compose an I-set, which is included in all equations to be estimated. The rest of independent variables are included by turns, in groups of two or three at a time. The test thus runs a number of regressions which can become really high. Attention is focused on the lowest and highest 'βs' among those estimated for any given variable. To the highest one is added a term, equal to two times its own standard deviation (as given by the computer output), and from the lowest one two times its standard deviation are deduced. If the two resulting figures lie to the same side of zero, the variable in question is deemed to be robustly related to the dependent one; else, their relationship is termed as non-robust.

These results seem to run counter the hypotheses of Armey, Barro, Rahn and Scully, but several reservations can be expressed. First, EBA attributes the same value to all specifications imaginable, provided that they contain the I-set of variables and irrespective of any other consideration. It is evident, however, that any regression that includes among its arguments variables that represent different links in the same causal chain could yield, by this very fact, insignificant and unreliable coefficients for these variables. The same may happen when we include, together with a given variable, any other which is co-determined or otherwise correlated with it⁵. Thus, one should be prepared to obtain for each variable a number of coefficients that exhibit insignificant and/or wrong signs, even when it is significantly related to the dependent one in properly specified models. The (excessive) precaution of adding two standard deviations to the *extreme bounds* of the coefficient distribution, the ensuing comparison of the signs of the resulting figures and the verdict of no robust relationship applied to any variable that fails to produce equal signs for the two compose a test apt to yield a lot of false negatives. Therefore, EBA analysis doesn't provide, in my opinion, a good way to deal with the problem of control variables' selection. But, in the absence of a better solution to this problem to which cross-sectional analysis is particularly sensitive, it is impossible for authors to agree on a definite conclusion.

This outcome invites to lend special attention to time-series analysis, focused on individual countries. The underlying idea is that, as each country presents its own inverted-U curve which is depending on the behavior of other variables, and as the latter tend to change less over time than they do across the space, the issue of control variables may be less of a problem. This, of course, doesn't mean that all variables other than the government share have to remain constant. In addition, in time-series analysis we have to deal with the problem of reverse causation. Where the dependent variable is income per person, one cause for this problem is the working of the Wagner's Law, the long-run tendency of the government share to increase with GDP per capita. Next, no matter whether the dependent variable is GDP or its rate of growth, another motive for reverse causation stems from the tendency of public expenditure to remain more stable than is private production during the business cycle. Due to this, the government share tends to present short-term fluctuations contrary to those of GDP and (foremost) to those of its growth rate.

As far as I know, no less than twenty academic contributions have tried to check the sign of the relationship between GDP or its growth rate and the size of government, by means

⁵ Regarding the control variables to be included, one question to bear in mind is that the size of the public sector may exert its purportedly negative effect on growth via two different mechanisms. First, if the marginal productivity of the production factors is lower in the public sector, then their diversion from the private one should have an adverse effect on growth. Second, the increased tax burden may impose a disincentive to savings, investment and, perhaps, effort. Since the production function approach takes investment as given, it ignores the latter effect, and may downplay the estimation of the negative impact of an increase in the size of the public sector. On the other hand, dispensing with investment when one is trying to explain the growth rate or the level of GDP doesn't seem to be advisable. There is no easy way out in this dilemma, however, when one is using single-equation models.

of time-series analysis applied to individual countries⁶. These countries are always developed ones, and, in them, the estimated relationship between GDP or its rate of growth and the government share always exhibits a negative sign⁷. Moreover, a majority of authors have been able to calculate the optimum size of the government share (from the point of view of economic growth) in the countries of their choice. Their estimates vary from 17 to 29 per cent of GDP in the USA, from 22.5 to 34 per cent in the former British dominions, and from 35 to 46 per cent in European countries. In spite of this country diversity, which can be easily attributed to the influence of other variables non-accounted for in the regressions, there is agreement in that most developed countries have gone far beyond the optimum in the size of their public sectors.

This general conclusion, however, is not totally beyond doubt, to the extent that this literature is subject to two possible criticisms. The first has to do with the problem of reverse causation, whose treatment in the papers raises some doubts. Certain contributions include the unemployment rate (or other similar variable) among the controls as an attempt to account for the business cycle⁸. Others report to have run Granger causality tests⁹. In other cases, the reader is told that the variables have been previously submitted to filters or to some other treatment which supposedly eliminates the problem¹⁰. It is possible, however, that these devices don't meet their goal, and that, where the dependent variable is the level of GDP, the problem is aggravated as a result of Wagner's Law. Finally, some papers don't seem to take any precautions against these potential dangers.

The second reservation regarding this literature refers to model specification in general. In a few cases one finds no control variables, or only dummies and/or ARMA and ARIMA terms instead. The issue of the indirect effects via investment is never addressed, either. A case in point is that of the specification used by Scully (1994, 1995, 1996, 2003, 2008) and by Chao and Grubel (1998) after him. This model uses a curious production function whose arguments are the outputs of both the public and private sectors during the

⁶ The list comprises: Peden and Bradley (1989), Peden (1991), Carlstrom and Gokhale (1991), Raymond (1992), Scully (1994, 1995, 1996, 2003 and 2008), Rahn and Fox (1996), Vedder and Gallaway (1998), Chao and Grubel (1998), Branson and Lowell (2001), Dar and AmirKhalkhali (2002), Afonso, Schuknecht and Tanzi (2003), Pevcin (2004), Angelopoulos and Philippopoulos (2007), Hill (2008), Magazzino (2008), Magazzino and Forte (2010), and Facchini and Melki (2012).

⁷ It is fair to add that, in a prior contribution due to Ram (1986), this author obtained a positive sign for 100 out of 115 countries analyzed. However, his explanatory variable was not the government share, but the rate of growth of public expenditures as such. The behavior of this variable is surely driven by the force of the Wagner's Law, and hence it can't be said that Ram's results contradict the conclusion stated above. The few negative signs obtained in Ram's estimations corresponded mainly to the most advanced economies. In two later contributions Grossman (1987 and 1988) also employed the absolute size of the public sector as a regressor, instead of its size relative to GDP. Yet, his results, referred to the USA during the 1929-to-1982 period, situated the optimal size of government for this country at 20 per cent of GDP, a figure that has been largely surpassed since the 1940s.

⁸ This is done by Grossman (1987, 1988), Peden and Bradley (1989), Peden (1991), Vedder and Gallaway (1998), and Facchini and Melki (2012).

⁹ Thus, Magazzino and Forte (2010).

¹⁰ As in Carlstrom and Gokhale (1991), or in Branson and Lowell (2001). The estimates of Dar and AmirKhalkhali (2002) for individual countries apparently stem from a later transformation of the results of their panel-data analysis.

previous year. The rationale underlying this particular choice is never stated, apart from signaling that the optimum government share can be readily calculated from the regression coefficients of these two variables; on the contrary, the validity of the function is implicitly assumed to be a matter of fact. Scully's papers deal with the USA, with the exception of Scully (1996) which refers to New Zealand, and economists of the latter country seem to have been particularly attentive to possible flaws in the model. Thus, Chapple (1997) pointed at an inconsistency between the theoretical model and that estimated in practice, and Sieper (1997) and Kennedy (2000) criticized the model itself on the grounds that it makes the implicit assumption that all investment of a given year is consumed in the next one. Later on, Hill (2008) has demonstrated that the results obtained with the Scully model are sensitive to the correction indicated by Chapple as well as to slight changes in the estimation period. Thus, whereas Scully's estimates for the optimal government share in the US closely cluster between 19.3 and 22 per cent of GDP, Hill's stretch from 9 to 29 per cent.

The preceding comments aren't intended to dismiss the importance of the cumulated evidence in favor of the hypotheses of Armey, Barro, Rahn and Scully, as achieved by time-series analyses and the bulk of cross-sectional ones. On the contrary, I believe that the basic convergence of results obtained from such different approaches can hardly be attributed to mere chance. Thus, my own contribution to the time-series approach, as presented in the ensuing sections, should be seen as no more than a test of these results, after taking strong precautions facing the issues of endogeneity and the indirect effects via investment. Prior to that, however, a reference to the most recent panel-data analyses has to be made.

Lastly, a number of authors have used panel-data analysis for long periods of time, focusing on developed countries only and using measures for public sector size that include *all* public expenditures. In a recent review, Bergh and Henrekson (2011) mention seven such articles, published in international academic journals with peer reviewing¹¹. With the exceptions of Agell *et al.* (2006) and Colombier *et al.* (2009), all other studies indicate that the growth effect of public-sector size is indeed negative for developed countries. Moreover, by adding new control variables Bergh and Karlsson (2010) and Bergh and Öhrn (2011) have managed to reverse the conclusions of the two dissenting papers. Thus, it can be said that the last wave of studies reasserts the negative role of too a big government share.

An innovative feature in the article of Bergh and Karlsson (2010) is that it uses a new type of robustness analysis, the *Bayesian averaging of classical estimates* (BACE) developed by Doppelhoffer, Miller and Sala i Martin (2004)¹². By applying this type of

¹¹ Their list includes Fölster and Henrekson (2001), Dar and Amir Khalkhali (2002), Agell *et al.* (2006), Romero Ávila and Strauch (2008), Colombier (2009), Afonso and Furceri (2010), and Bergh and Karlsson (2010). Were it enlarged to include non-published papers, it could also mention, for instance, Heitger (2001), Chobanov and Mladenova (2009), Josheski *et al.* (2011), Bergh and Öhrn (2011), or Afonso and Tovar Jalles (2011).

¹² This method calculates a coefficient for each explanatory variable as the mean of all its estimates, as provided by the different regressions. Contrary to EBA, BACE weights these coefficients by the goodness of the fitting of the

analysis to a panel of rich countries during the 1970-to-1995 and 1970-to-2005 periods, and using 21 variables (in all of their 7-digit possible combinations), Bergh and Karlsson (2010) conclude that the ratio of tax revenue to GDP and (for the 1970-to-2005 period) the ratio of government expenditure to GDP are indeed robust determinants of economic growth.

At the end of the day, it can thus be said that the results of panel-data analyses confirm the results obtained by time-series analyses as well as by most cross-sectional ones.

3. The model and data

The main contribution of this paper consists of a model designed to confront the problems of endogeneity and the dual role of the government share (direct plus indirect, via investment) on income determination, in a framework of time-series analysis. This choice, in turn, has several consequences:

- A time-series regression of the *rate of growth* of income on the *level* of the government share (i.e. the checking of Scully's hypothesis by means of time series regression) is not possible in most countries, due to the fact that the two series exhibit different orders of integration.
- Thus, the surviving options are either using levels for the two variables, or taking growth rates or logarithms for both. Presented with these options, I have preferred to work with growth rates, which minimize the cointegration problems. Moreover, relating growth rate to growth rate reproduces the standard formula of the (short-term) elasticity of the dependent variable to changes in the explanatory one, a fact that facilitates the economic interpretation of the coefficients to be obtained.
- The use of growth rates instead of levels or logarithms also presents the possible advantage of eliminating the reverse-causation problem due to the working of the Wagner's Law, given that the latter acts mainly in the long run. However, this doesn't eliminate the short-term reverse causation implied by the business cycle. Among the different methods available to deal with this problem, the best suited to the case seems to be the construction of a simultaneous equation model.
- Finally, one consequence of the decision of working with growth rates is the resignation to estimate the Armey curve as such, since the quadratic form of the relationship doesn't apply to growth rates. What I use here is a linear model which yields an estimate of the elasticity of GDP to changes in the public sector size. This specification imposes the restriction of constant elasticity, which doesn't fit too well with the shape of Armey curves: even in those countries for

corresponding regressions. This criterion is also used in order to determine the *ex post* probability of each specification, and (out of this) of each potentially explanatory variable. Also in opposition to EBA, which starts by separating variables in two arrays, BACE considers all of them to be equally probable candidates *ex ante* to influence the dependent variable. At the end of the day, all regressors whose probability *ex post* exceeds that *ex ante* are considered to be robust determinants of the dependent variable.

which all or most observations are located on one and the same side of the Armey curve, it is dubious that the corresponding half of the latter presents a constant elasticity.

Once these comments have been made, we can proceed to present the model, which assumes an aggregated production function of the Cobb Douglas type:

$$Y = A \cdot K^{\beta} \cdot L^{1-\beta} \quad (1)$$

When dividing the two sides of this expression by L, we get the intensive or per-worker from of the function:

$$Y/L = A \cdot (K/L)^{\beta} \quad (2)$$

Taking growth rates (indicated here by means of lower case letters), the preceding expression gives:

$$y-l = a + \beta(k-l) \quad (3)$$

where, by want of the necessary information, we are forced to take 'a' as a constant.

To the latter expression we may now add other possible determinants of 'y/l', as well as an error term. To start with, we can write:

$$y-l = a + \beta(k-l) + \gamma(g-y) + \varepsilon \quad (4)$$

where 'g-y' indicates the growth rate of public expenditures relative to GDP. Of course, other arguments may be added to the right-hand side of this expression.

Since we need to economize on symbols, we will rewrite the last expression as:

$$y-l = a0 + a1(k-l) + a2(g-y) + \varepsilon \quad (5)$$

It is apparent in this expression that, since 'y' enters the calculation of the rate of growth of the government share in GDP (g-y), we have a problem of reverse causality. For this reason we will formulate a second equation, in which 'g-y' is a function of 'y-l', as well as of a vector of other determinants (R):

$$g-y = b0 + b1(k-l) + b2(R) + \eta \quad (6)$$

Another point deserving attention in (5) is the assumption that, in the relationship between product and capital, the causation runs from the latter to the former, as stated in all growth models. However, many models in macroeconomics assume that the investment rate is a function of (among other variables) the rate of growth of the economy. This assumption is most clearly found in the Keynesian theory of the accelerator, which takes investment to be dependent on the increase in output, but macroeconomics also stresses the close dependence of consumption on income, and the refined version of this hypothesis considers consumption to be a function of *permanent* income. If we put apart the permanent and non-permanent components of income, and given the identity $I+C = Y$, it follows from this that investment closely depends from the transitory component of income, which is tantamount to say from income fluctuations. The assumption that investment depends on the rate of growth of the economy is also in accordance with common sense, which suggests that the capacity-utilization rate and, perhaps, the rate of depreciation of the capital stock, also depend on the cycle. With all this, it seems safe to assume that the rate of growth of the stock of capital per worker is a function of 'y-l'. Moreover, if there is some truth in the idea that an increase in the size of the public sector affects investment, the rate of growth of the capital stock per worker

should be also a function of 'g-y'. Allowing for the existence of a full host of other determinants of investment (S), we should then write:

$$k-l = c_0 + c_1 (y-l) + c_2 (g-y) + c_3 (S) + e \quad (7)$$

With this, the expressions (5), (6) and (7) become the structural equations of a system of simultaneous equations, in which 'y-l', 'g-y' and 'k-l' are endogenous variables. The exogenous or predetermined variables of the system are all those comprised in the vectors 'R' and 'S', plus those pertaining to a third vector (T) that we could add to (5) if we consider it convenient.

From the structural equations we have to formulate those of the reduced form, in order to estimate the model. Since the rate of growth of capital per worker and that of the share of government influence each other (directly the latter to the former, and via the rate of growth of product per worker in the opposite case), the three equations of the reduced model have to include, to the right of the equal sign, all variables that make up the vectors 'R' and 'S'. Namely:

$$y-l = d_0 + d_1 (R) + d_2 (S) + e \quad (8)$$

$$k-l = f_0 + f_1 (R) + f_2 (S) + u \quad (9)$$

$$g-y = h_0 + h_1 (R) + h_2 (S) + u \quad (10)$$

Our next task is to decide which variables should be included in the vectors 'R' and 'S' and, if the case comes, in vector 'T'. This is necessary not only to make the model operational, but also to determine which estimation technique is adequate to our case. Depending on the number of these predetermined variables, if the number of parameters in the reduced-form equations is equal to that in the structural ones, the model is exactly identified, and the adequate estimation method is indirect least squares. If, on the contrary, the reduced-form equations contain more parameters than do the structural equations, the model is over-identified, and the appropriate technique is two-stage least squares. With more than two variables in each of the vectors 'R' and 'S', our system becomes over-identified, and the appropriate method turns to be 2SLS. By the way, it is worth noting that 2SLS also provides a way out the annoying dilemma of the inclusion of investment in the growth equation: in the second-stage growth regression, the instrument for investment represents only that part of the increase in the capital stock that is free from the influence of changes in the size of the public sector. Therefore, the regression coefficient of the instrument for the government share should catch the full impact of the latter on growth, consisting of its direct effect (via resources' productivity) plus the indirect one (via how many resources are allocated).

The selection of the predetermined variables is a task to be performed with great care, since inclusion of any one that is not really exogenous would render the whole exercise useless. Here we are facing a tradeoff. The more numerous the variables included in the sets 'R', 'S' and 'T', the better is the fitting of the first-stage regressions and, at first sight, the higher should be the confidence we put on the estimated instruments to be used in the regressions of the second stage. On the other hand, however, any mistake incurred into in this domain could ruin the whole construction. When reviewing the literatures on the determinants of income growth, investment, and public sector growth, a number of

clear candidates to be included in the corresponding sets appear; others should be undoubtedly rejected, and some cases remain dubious.

Starting with the determinants of the growth of the public sector, two among these look non-suspect. Namely: the so called Baumol effect (i.e. the tendency of service prices to increase over those of material goods overtime) and the degree of openness of the economy. Other determinants of the size of public sector include population ageing and the weakening of the traditional social and family networks that served in the past to facing the risks of illness, unemployment and so on. In a short run analysis as is this one, long-term effects running from GDP level to the levels of these variables aren't motives for worry, but short term effects from economic growth to increases in our variables are. Population ageing is influenced by income through increased life expectation, but only in the long run, so that there seems to be no problem in including the proportion of elderly among the determinants of changes in the size of public sector. This is not the case, however, of the proportion of children, which is affected in the short run by income growth through changes in the birth rate. The same applies to any indicators we can imagine for the destruction of the traditional safety networks (from changes in family size to those in the internal composition of employment, passing through the urbanization rate or that of monetization of the economy), which are influenced by income growth and not only in the long term but also (albeit with the opposite sign in some cases) in the short run. Thus, we have to be content with limiting the R-set to the ratio of service prices to those of material goods (or its rate of growth), to the usual indicators for the degree of openness of the economy, and to the share of the elderly within the population. To these we could add the color of the incumbent government, which probably has also an influence on public expenditures and taxes.

The list of determinants of investment (that is, the S-set) presents more dubious cases. Genuinely exogenous to the system seems to be the behavior of output in other countries, which may affect investment in ours in several manners. Less clear is the case of the interest rate. Were it formed in really free markets, this variable would be influenced by changes in the demand for investment as much as the other way round. However, the interest rate is established by the monetary authorities of each country (or by supranational authorities), with an eye put on the behavior of prices and the other in the interest rates elsewhere¹³. Thus, there are no apparent problems regarding its inclusion. The behavior of the relative prices of investment vs. consumption goods could be another determinant of investment, but this time both the reasoning and the causality tests indicate that investment tends to influence this ratio more often than the opposite. The changes in the openness of the economy may or may not influence investment, but this doesn't pose a problem as that variable has been already included in the group of predetermined ones, due to its presumed influence on the size of the public sector.

¹³ The rate of growth of the economy is, according to Taylor (1993), a likely determinant of central bank policy. However, a Granger causality test run for each of the countries that form the sample indicates that causality runs from interest rates to an increase in capital per worker in ten of our countries; only one instance (that of Greece) of causality seemingly acting in the opposite direction has been detected.

Another possible candidate is the inflation rate of the economy which, according to many authors, may deter capital accumulation.

Regarding those determinants of the rate of growth of output per worker (the dependent variable in our growth equation) other than the increase in capital per worker and changes in the size of the public sector, it would have been desirable to include other usual arguments of the aggregate production function, as may be the stock of human capital per worker. However, the extant series for this variable have been calculated through several steps and, in each of these, heroic assumptions have been made. The resulting series thus exhibit a smooth outline, apt to represent the trend but insensitive to short term fluctuations. More often than not, their authors use only one decimal point (the variable is usually measured in school years per person), so that the series tend to remain stable for years, jumping to the next stair from time to time. In either case, these series aren't suitable for our work with growth rates.

All regressors used in the first-stage regressions have to be used also in one or more of the second-stage ones. However, the second-stage regressions may include other variables that we have refrained to use in the first one on the grounds that they cannot be considered as truly exogenous. Variables that we could include in the growth equation of the second stage only are the already mentioned changes in the proportion of children within the population, the internal composition of the labor force by sector and status, or the rural/urban distribution of the population, all of which tend to accompany economic growth. Most probably, these variables aren't real determinants of growth, but here the target is to check the robustness of the influence of the government share on growth. Given this objective, the presumed association of these variables with economic growth converts them in good controls, as they make the test more stringent. A final comment relating many variables of the system is that, on theoretical grounds, it is unclear whether we should use their levels or, rather, their rates of growth in the equations where they appear. Hence, where possible I have run different regressions using either their levels or their growth rates.

In the two following sections I perform two different experiments. In section 4 I use a rather reduced set of predetermined variables to estimate the first-stage regressions, and no other exogenous variables in the second-stage ones. That set consists only of the behavior of the growth rate in the rest of the countries that form the sample, the interest rate, the ratio of service prices to those of material goods, and the openness of the economy, all of which are available for the (almost) complete sample of countries. In section 5, the list of predetermined variables is reinforced with others that aren't available for all countries. With the same purpose, in that section I employ other available indicators, different from the ones used in section 4, for the same variables as used in it, as well as new control variables to be included in the second-stage regressions. Obviously, the use of all series is preceded by a test in order to ensure that all of them are stationary; where not, the series in question have been discarded.

Going now to the data sources, the three endogenous variables of our system (i.e. the rates of growth of product per worker, of the government share in the social product

and of the stock of capital per worker) have been calculated out of data taken from the Penn World Tables (PWT 6.3). Output per worker is real output, expressed in PPPs, divided by total employment (rgdpl2te in the source). The government share is that corresponding to public consumption, calculated at current prices in order not to interfere with the mechanism identified by Baumol (cg in the source). As for the growth rate of the capital-stock per worker, its construction has required some more effort. One possible indicator for this variable is the investment rate. Since the numerator of its formula indicates the gross increase in the capital stock, and its denominator (Y) is a function of (among other things) the number of persons working in the economy, the investment rate could provide an acceptable proxy for the increase in capital per worker. However, better indicators for this variable can be constructed from the Penn World Tables. In order to calculate the rate of growth of the *stock* of capital, be it in absolute terms or per worker, we need first a series on the capital stock as such. This is hard to get. However, the stock of capital existing in any given year is no more than a summation of past investment, in which the bulk corresponds to the decades immediately prior to the year in question. Given that PWT provide data on investment starting in 1950, the construction of an acceptable proxy for the capital stock since (say) 1975 onwards becomes thus possible. To get a more refined figure we can apply to the simple sum a depreciation rate in the customary form. Through this process, I have calculated two alternative proxies for the capital stock, with and without depreciation. In the cases of those countries and periods for which I have found more careful estimates for this variable, constructed by other researchers or teams of researchers, a comparison of my series with theirs have been possible. The correlation coefficients between them are very high and, surprisingly, higher for my gross measure. For instance, in the case of Spain the 'r' between my gross measure of the capital stock and the best series available is 0.9966.

For the rest of variables, where they were not reported in PWT I have resorted to the customary sources: OECD, IMF, World Bank and, in some cases, to national sources as well. Where several alternative series were at hand, as is the case of the interest rates, all series that I have been able to find have been used in section 5. The color of the government has been assigned out of information provided by the webpage Worldstatesmen.org. To this end I have examined only the party affiliation or political background of the chief of the executive. Only in two cases (those of Switzerland and Mexico) the color of government became too problematic to elucidate, obviously for different reasons in each of these countries.

Finally, the selection of countries has been motivated by an interest in looking at what happens where there are large public sectors, as is the usual case for rich countries. Within OECD, and for evident reasons, I have dispensed with all former communist countries. Germany has been also excluded, due in part to the same reason as well as to the disruption of her time series as a consequence of reunification. As for the time span, three different motives led me to restrict it to recent years. First, as stated before, the use of linear models may be ill suited for the study of long periods, in which the

observations could be distributed almost certainly over the two halves of the Armey curve. Second, I wanted to minimize the risk of using series that could exhibit changes in their structure. Other things being equal, the shorter the period, the lower this risk would be. The period stretching from 1975 to 2007 looks long enough for not getting run of degrees of freedom; at the same time, it is not too long, and, running from the aftermath of the first oil shock of 1973 to the eve of the 2008 crisis, it includes no major economic breakthrough. Last, a period neither too long nor too short was convenient, as stated above, for the calculation of a series for the stock of capital.

Throughout the preceding presentation it should become evident that the test undertaken in this article is purposely restrictive, and biased against the hypotheses submitted to check:

- The use of growth rates implies that we are estimating short-term elasticities only.
- Certain control variables included in second-stage regressions may improperly detract from the levels of significance of the variable of interest.
- Government expenditures are represented by consumption expenditures, which is not the type of public expenditures that has increased the most during the period.

And, the post-2008 years have been excluded, as the possible oversizing of the public sector during these may have been magnified due to shrinking private production.

4. Presenting the baseline estimates

As indicated in the precedent section, here I present the regression results corresponding to a very basic model, which includes only those predetermined variables that are available for almost all countries. Namely: the growth rates of product per worker in the rest of OECD countries; the interest rate; the ratio of service prices to those of material goods, and the openness of the economy. Since our main interest lies in the growth equation, the results presented here correspond only to those second stage regressions that take the growth rate of GDP per worker as their dependent variable, and the instruments for the rate of growth of capital per worker and the Government share as the explanatory ones. The results of the two first stage regressions, by means of which these two instruments were estimated, are available upon request.

Table 1 thus presents the results of the second-stage regressions of the rate of growth of product per worker on the instruments estimated for the rates of growth of the government share and capital per worker, without taking into account other control variables. The fitting of the regressions is modest in most cases, but this doesn't come as a surprise given the short list of explanatory variables used in the estimations. Since our main interest lies in the growth effect of the government share, I will restrict to this most of my comments.

First, it is remarkable that all coefficients but one (that for Spain) are negative, and even in this case it is very close to zero. This said, one has to add that only in half the cases (13

countries) these coefficients are statistically significant at the 0.1 or lower levels. Moreover, according to the relevant statistics our estimated instruments for (the rate of growth of) the government share are weak in the cases of Finland, Italy, Japan, New Zealand, Portugal, the United Kingdom and the United States¹⁴. Moreover, the corresponding first-stage regressions are only marginally significant at the 0.1 level in the additional cases of Austria, Canada, and the Netherlands. In all of these cases, therefore, one shouldn't put too much confidence in our coefficients for the variable of interest, as estimated in this baseline model.

There is no reason, however, to expect any such bias in the other cases, which include nine significant results. Up to this point, then, our baseline results support the hypothesis of excess government for roughly one third of the countries that form our sample. In seven additional cases the coefficients are also negative but insignificant, and in the rest of countries but one the coefficients are still negative, although suspect.

The persistence of a possible bias, the lack of significance of other coefficients of interest and the sometimes low R^2 s of other regressions (although significant in all cases but two, as indicated by the F-test), invite to see what happens when using other alternative indicators for some of the variables included in these baseline regressions, and, more important, when other control variables are incorporated to the model.

¹⁴ The F-statistics have been calculated according to Angrist and Pischke (2009:218) with their correction later posted at <http://www.mostlyharmlesseconometrics.com/2009/10/multivariate-first-stage-f-not/>.

Table 1: Results of Second-Stage Regressions of the Rate of Growth of Product per Worker on the Instruments Estimated for the Rates of Growth of Capital per Worker and for the Government Share. Baseline Model (1975-2007^a).

Country	Constant	[^] GrKpw	[^] gr G/Y	AR term	Adjusted R ²	Prob(F) ^b
Australia	-4.411***	1.225***	-0.626***		0.49	<0.01
Austria	-4.310**	1.142***	-0.690		0.25	<0.01
Belgium	-5.209*	1.360**	-0.114		0.10	<0.10
Canada	-2.325	0.671	-0.356*	0.373*	0.26	<0.05
Denmark	-5.616	1.355	-0.477**		0.11	<0.01
Finland	2.666	-0.102	-0.361***		0.29	<0.01
France	-12.702**	0.547**	-0.261**		0.20	<0.05
Greece	4.691	-0.745	-0.160		-0.05	<0.05
Iceland	-3.175	1.257*	-0.660**		0.19	<0.01
Ireland	0.236	0.485	-0.416*		0.02	>0.10
Italy	0.799	0.074	-0.481*	0.649***	0.35	<0.01
Japan	5.406	-0.710	-0.471	0.504**	0.22	<0.01
Korea	4.112	-0.091	-0.589***		0.36	<0.01
Luxembourg	5.739	-0.922	-0.149		-0.02	<0.01
Mexico	-3.773	1.088	-1.038		0.02	<0.01
Netherlands	-0.605	0.314	-0.287		-0.06	<0.01
New Zealand	-5.848***	1.377***	-2.489**	-0.625***	0.74	<0.01
Norway	-2.945	1.127**	-0.090		0.14	<0.05
Portugal	-2.907	0.920	-0.196		0.03	>0.10
Spain	-0.711	0.484**	0.021		0.17	<0.01
Sweden	-3.494*	1.081***	-0.294*		0.20	<0.01
Switzerland	-3.604**	1.110***	-0.731***		0.57	<0.01
Turkey	12.008	2.387*	-0.100		0.09	<0.05
U. Kingdom	-3.047	0.857	-0.498*		0.05	<0.10
United States	3.216*	-0.383	-0.033		0.15	<0.01

Notes: ^a The start and ending years may differ slightly from country to country, depending on data availability.

^b As calculated according to Angrist and Pischke (2009:140).

Source: See text.

5. Improving the results and examining their robustness

In this section, the results of two different exercises are presented. The first one, corresponding to table 2, is a repetition of the estimates presented in the preceding section, using alternative indicators for the variables of the baseline model. For one of these, the interest rate, there are usually many series available corresponding to as many types of financial assets. Furthermore, in relation to interest rates both nominal and real interest rates deserve consideration. For the rate of growth of the capital stock we may have up to three different proxies: the gross rate, the net rate, and the investment rate. Finally, regarding all predetermined variables it is unclear on theoretical grounds whether we should use their growth rates or their levels. All of this gives rise to a theoretically high number of combinations that varies from country to country. In practice, however, these numbers are lower either because certain indicators present inadequate orders of integrations, or certain combinations, whose proportion on the total varies among countries, give rise only to weak instruments for the variables to be used in second-stage regressions. The information offered in table 2 refers to second-stage growth regressions only, and specifies the numbers of valid regressions run, the number of times where the coefficient of the growth rate of the government share is negative [positive], how many times it turns to be significant, its average value and its SD. For this table we have discarded all those regressions where the instruments to be used weren't strong enough, as indicated in the results of the corresponding first-stage ones¹⁵.

The second exercise whose results are presented in this section (table 3) consists of repeating the estimations of section 1, including now a higher number of predetermined variables in the first-stage regressions, as well as new controls in the second-stage ones. The former include the growth rate of the share of the elderly within the population, the color of the incumbent government, and the rate of inflation, taken one or two at a time. As control variables for the second-stage growth regressions I have used these plus the rates of growth of the share of the population under 15 years, that of employees and that of agricultural workers in total labor force. Again, these variables have been included one or two at a time. Since the share of the elderly and the color of government are obvious determinants of public expenditures, even consumption ones, and no doubts arise about their exogeneity, it doesn't make sense to include them in the second-stage regressions solely, so that they appear in them only where they have been previously included in the corresponding first-stage regressions. With this, the theoretical number of second-stage growth regressions to be run for each country amounts to 297. Again, however, the actual numbers are lower for all countries, by want of suitable series for certain variables and countries, or because some series exhibit inadequate orders of integration, or because certain first-stage regressions fail to produce good instruments

¹⁵ The criterion for inclusion has been that the F-statistic of the corresponding first-stage regression, calculated in the way described in the precedent section, was significant at least at the 0.1 level.

for the government share. This notwithstanding, the numbers of useful estimates in table 3 is higher than those in table 2 for all countries (except for New Zealand, for which the two are zero). On average, they are higher in proportion of four to one.

Table 2: Short-Term Elasticity of Product per Worker to the Government Share within GDP (Coefficients Estimated for the Government Share in Second-Stage Regressions). Baseline Model, Using Alternative Proxies for the Variables (1975-2007^a).

Countries	Regressions run (Ljung & Box Q significant $\alpha = 0.05$)	With negative sign for $\hat{gr}(G/Y)$	of which significant (at least) at 0.1	With positive sign for $\hat{gr}(G/Y)$	of which significant (at least) at 0.1	Mean of the coefficient for $\hat{gr}(G/Y)$	SD of the coefficient for $\hat{gr}(G/Y)$
Australia	18	All	9	None	—	-0.426	0.127
Austria	35	All	12	None	—	-0.700	0.183
Belgium	8	All	None	None	—	-0.290	0.172
Canada	4	All	3	None	—	-0.323	0.099
Denmark	8	All	5	None	—	-0.379	0.131
Finland	14	All	All	None	—	-0.368	0.021
France	8	All	7	None	—	-0.346	0.101
Greece	5	3	None	2	None	-0.013	0.072
Iceland	4	3	1	None	—	-0.371	0.063
Ireland	8	All	All	None	—	-0.485	0.048
Italy	15	All	14	None	—	-0.783	0.187
Japan	—	—	—	—	—	—	—
Korea	17	All	All	None	—	-0.681	0.198
Luxembourg	9	7	4	2	None	-0.157	0.197
Mexico	16	All	None	None	—	-1.063	0.235
Netherlands	12	All	None	None	—	-0.332	0.048
N. Zealand	—	—	—	—	—	—	—
Norway	44	42	19	2	None	-0.103	0.041
Portugal	6	5	None	1	None	-0.285	0.244
Spain	11	8	None	3	None	-0.006	0.125
Sweden	18	All	11	None	—	-0.370	0.176
Switzerland	57	All	All	None	—	-0.682	0.101
Turkey	8	2	None	6	None	0.063	0.105
U. Kingdom	2	All	1	None	—	-0.278	0.179
U. States	2	1	None	1	None	-0.044	0.150

Note: ^a The exact period under analysis varies from one estimate to another, depending on the starting and ending years of the series employed in each of them.

Source: See text.

Table 3: Short-Term Elasticity of Product per Worker to the Government Share within GDP (Coefficients Estimated for the Government Share in Second-Stage Regressions). Estimations Including New Control Variables (1975-2007 ^A).

Countries	Regressions run (Ljung & Box Q significant $\alpha = 0.05$)	With negative sign for $\hat{\alpha}$ gr(G/Y)	of which significant (at least) at 0.1	With positive sign for $\hat{\alpha}$ gr(G/Y)	of which significant (at least) at 0.1	Mean of the coefficient for $\hat{\alpha}$ gr(G/Y)	SD of the coefficient for $\hat{\alpha}$ gr(G/Y)
Australia	32	All	27	None	—	-0.540	0.121
Austria	60	All	19	None	—	-0.886	0.345
Belgium	32	27	2	5	None	-0.137	0.107
Canada	32	All	23	None	—	-0.285	0.087
Denmark	102	All	84	None	—	-0.416	0.058
Finland	44	All	35	None	—	-0.349	0.082
France	168	157	112	11	None	-0.274	0.161
Greece	27	24	None	3	None	-0.114	0.138
Iceland	11	All	8	None	—	-0.849	0.356
Ireland	45	34	18	11	None	-0.335	0.563
Italy	88	75	13	13	None	-0.223	0.218
Japan	60	All	49	None	—	-0.441	0.042
Korea	44	42	34	2	—	-0.467	0.172
Luxembourg	102	All	51	None	—	-0.251	0.112
Mexico	7	5	2	2	—	-0.641	0.978
Netherlands	102	83	None	19	None	-0.187	0.178
New Zealand	—	—	—	—	—	—	—
Norway	76	56	22	20	None	-0.066	0.121
Portugal	22	All	19	None	—	-0.564	0.067
Spain	102	23	None	79	None	0.068	0.091
Sweden	89	All	60	None	—	-0.329	0.028
Switzerland	75	All	71	None	—	-0.721	0.266
Turkey	29	26	None	3	None	-0.064	0.047
U. Kingdom	29	All	23	None	—	-0.588	0.175
United States	51	44	18	7	None	-0.175	0.126

Note: ^a The exact period under analysis varies from one estimate to another, depending on the starting and ending years of the series employed in each of them.

Source: See text.

A detailed comment on the results presented in tables 2 and 3, country by country, is out of question in this article. Thus, the following comments are restricted to the signs of the coefficients of the variable of interest, their values, their significance levels, their means

and their standard deviations within each country. The similarity between tables 2 and 3 in all of these respects is another point to be commented.

Regarding the signs of the coefficients for the government share, as they appear in tables 2 and 3, these are predominantly negative, in the proportions of 18 to one in table 2, and seven to one in table 3. Almost all positive coefficients differ from zero starting in the second or third decimal only. Furthermore, they stem from specifications that include as controls so powerful explanatory variables as are the rate of growth in other countries, and/or the inflation rate. The aggregate figures, however, conceal differences among economies. In ten countries all coefficients estimated are negative, these being Australia, Austria, Canada, Denmark, Finland, Iceland, Sweden, Switzerland and the United Kingdom (plus Japan, for which table 2 provides no information). The coefficients are also overwhelmingly negative in the cases of Belgium, France, Italy, Korea, Luxembourg, Mexico and Portugal, and predominantly negative though not so much in the rest of cases, with the sole exception of Spain. Even in this country the positive coefficients correspond to estimates that include the late 1970s, whereas the positive ones pertain to estimates that, by want of long series for the corresponding variables, start around 1980. This is not so strange, as Spain emerged from an authoritarian regime in 1976 with a very small public sector that quickly increased in subsequent years. A replication of the Spanish regressions starting in the early 1980s instead of 1975 causes all positive coefficients to change to negative. Thus, a likely interpretation is that Spain surpassed her optimum size for the public sector at a later date than did other countries. Something similar, though far less evident, seems to have also happened in the parallel cases of Portugal and Greece. In any event, the coefficients for Spain are close to zero (0.06 or -0.07 on average, depending on the starting year), and those of the other two countries are mainly negative all the time.

The next question is that of the significance of the coefficients obtained. Nowhere in the tables do we find positive coefficients that turn to be significant. But there is no country for which *all* coefficients are both negative and significant in the two tables, either. Four countries are in this situation in table 2 (Namely: Finland, Ireland, Korea and Switzerland), but none in table 3. In the aggregate, in one country (namely: Switzerland) the proportion of negative and significant coefficients is over 90 per cent (97%), and in other four (Denmark, Finland, Japan and Korea) it is over 80 per cent. It is over 70 per cent in other three countries (Australia, Canada and the United Kingdom), and over 60 per cent in four more (France, Iceland, Portugal, and Sweden).

Thus far, then, we can say to have found rather strong indications of a negative relationship between the rates of growth of the government share in more than half the countries, and no similar indication of a positive one for any one. Since in no country all coefficients are both negative and significant, the partisans of the extreme bound analysis are free to interpret from this exercise that the inverse relationship between the rates of growth of output per worker and of the government share is not robust in any country whatsoever. My own interpretation is different, however. In a prior section I have compared the extreme bound analysis to a clinical test that yields a lot of false

positives. Focusing on the trees of the extreme bounds is perhaps the best way to lose sight of the wood of the whole distribution of the coefficients estimated¹⁶.

In the two last columns of tables 2 and 3 we can find the means and SDs of the distributions corresponding to the coefficients estimated for each country, and their comparison is always interesting. As a matter of fact, in those countries for which we have many estimations (which are almost all in table 3) we could made the assumption that the estimated coefficients compose random samples, and calculate the confidence level of their means in the manner that is usual in random sampling. This procedure is also analogous to the one used to calculate the level of significance of any coefficient in any particular regression, with the main difference that it can be estimated directly instead of being inferred from the distribution of the errors.

Generally speaking, the mean of our coefficients is high relative to the SD, in both tables, for Australia, Austria, Canada, Denmark, Finland, France, Iceland, Korea, Japan, Luxembourg, Sweden, Switzerland, and the UK: 13 countries in total. Were we to formally calculate the analogous to the significance levels of the mean coefficients corresponding to these countries, we would place them at the 0.01 (in both tables) for the cases of Australia, Austria, Finland, Korea and Switzerland, and above the 0.05 level for Sweden. Whereas the number of regressions reported in table 2 is too small to allow for such calculations, the “significance levels” arousing from table 3 would be 0.01 for Canada, Japan, Luxembourg, and the United Kingdom, and 0.05 for Denmark, France, and Iceland.

Thus far we have employed three different criterions to evaluate our results. Namely: whether all (or almost all) coefficients exhibit the same negative sign; whether a majority of them are significant, and whether their means compare favorably with their SDs. There are seven countries that accomplish with all three criteria to a large extent, these being Australia, Canada, Denmark, Finland, Japan, Switzerland and the United Kingdom. Other five countries (Iceland, France, Korea, Luxembourg and Sweden) also accomplish with these three criterions, although not as well as the former ones. Generally speaking, then, we can say to have found *strong* indications that an increase in the government share is bad for economic growth in at least twelve out of 24 countries during the period under examination. Regarding the rest of the countries, the conclusion is less firm, although in Austria, Belgium, Ireland, Italy, Mexico, Norway, Portugal, the United States the results clearly point in the same direction. In Greece, the Netherlands, Turkey and Spain after 1980, the indications also point to this direction, albeit timidly.

Is the assessment of a negative effect of the government share equivalent to a confirmation of the Army hypothesis? For this we should compare the ranking of

¹⁶ Furthermore, we have to take into account that, even where all components of any given distribution exhibit the same sign, almost by definition one of the extreme bounds will be not too far from zero and this implies a high probability of its non-being significant. There is another possible reason that may cause one coefficient not to be significant where others in the same distribution are: this is that the error term in *that* precise specification (which provides the denominator for the expression of the t-values) is higher than those of the rest of others. But this usually means that the former specification is less efficient than are the other ones, so that we shouldn't put much stress on that particular result while being oblivious of the rest.

countries that we can establish according to our chosen criteria with the one we can construct from the relative size of the public sector throughout the period under analysis.

The results of this comparison are mixed. It is true that all OECD countries tend to exhibit large public sectors, and that Sweden, Denmark, Finland and the United Kingdom hold the four top posts in this respect. According to our estimates, they have in fact *too large* public sectors. But increases in public expenditures also seem to have had clear negative economic effects in the cases of Australia, Canada, Japan, Korea, Luxembourg or (especially) Switzerland, all of which exhibit small public sectors. On their part, Norway or the Netherlands are close to be in the opposite case.

One possible interpretation is that, when countries increase the size of their public sectors over time, they don't do this in the orderly manner implicitly assumed in Armey's hypothesis, i.e., starting with the most growth-enhancing public activities and going down the scale up to those activities with no positive effect that could counteract the necessary increase in the fiscal burden. This could allow for negative effects in countries with comparatively small public sectors, as well as to the opposite case, due to differences in the internal composition of expenditures and taxes. Other possible explanations could point to the relative efficiency of the national bureaucracies and decision making processes, as determined by the quality of both formal and informal institutions of each country. These points have been stressed by the recent literature in order to explain anomalies not dissimilar from the ones we have found here. Thus, Dar and AmirKhalkhali (2001), as well as Bergh and Henrekson (2011), employ similar arguments to explain the more than acceptable economic record of Scandinavian countries. Afonso and Tovar Jalles (2011) also find that institutional quality may counteract the effect of the government share in countries that rate differently in these two respects, or reinforce it in countries that score similarly in the two of them.

Finally, a different interpretation is that it is not so much the absolute levels of public expenditures, as their changes, that matters for GDP growth. Our method of estimation of the elasticity of GDP to changes in the size of public consumption, through regression of the rate of growth of the former on that of the latter, estimates indeed short-run elasticities, thus allowing for this interpretation. If the explanatory variable doesn't exert a *durable* effect on the dependent one, then the long-run elasticity of the latter to the former would be lower than the short-run elasticity, and perhaps much lower. However, it is not easy to imagine why and how this may happen.

In any case, the results obtained don't seem to be an artifact of inefficient estimation. Care has been taken with problems of reverse causality, specification selection and elimination of weak instruments. As a further precaution, the results of the whole exercise have been validated by one additional test that can be described as follows. Our ultimate reason to use a simultaneous-equations' model and two-stage least squares is that the explanatory variable, G/Y , depends to some extent on changes in the dependent one, which provides the denominator for its expression. This dependence happens because, although the movements of G along the business cycle follow those of Y , they

are smoother than these. To put it in another way, the elasticity of G to Y is positive in all countries, though lower than unity. On the contrary, the elasticity of the denominator of G/Y to Y is one by definition. Now, if we proceed to replacing the Y s in expression G/Y by their *trend* values, then the elasticity of its numerator to GDP turns to be higher than that of its denominator. Thus, when we recalculate the G/Y ratios using a smoothed series for Y (as are the one that results from passing the original series through the Hendrick-Prescott filter), and we use the series of G/Y thus modified as a regressor with which to explain the behavior of the growth rate of product per worker in a single-equation model using OLS, the results turn to be biased towards the positive, instead of towards the negative as it happens when the same regression is run using the original GDP figures. Hence, this exercise provides us with an *upwards-biased* figure for the coefficient that we are trying to estimate. Thus, when this upwards-biased coefficient turns to be negative for any given country, the conclusion to be drawn is that, in that country, the rate of growth of G/Y indeed exerts a strong negative impact on that of GDP. In regressions not shown here I have run this test. The results still yield negative coefficients for the variable of interest in about half the countries for which we have obtained clearly negative coefficients throughout the paper, and about zero in most other countries pertaining to this group. Incidentally, this test also indicates a negative growth effect of increases of the public sector in New Zealand, a country for which I didn't manage to get efficient instruments for the variable in the 2SLS exercise, and is, therefore, absent from tables 2 and 3¹⁷.

5. Concluding remarks

All academic contributions aimed at estimating the growth effects of big public sectors by means of time-series analysis unanimously conclude that these effects are negative, and this is also the conclusion of most panel-data and (to a lesser extent) of most cross-sectional analyses addressed to the same question. Recently, Bergh and Henrekson (2011) have stated a strong case in favor of the negative role of big public sectors, relying exclusively in the results of recent panel-data analyses. Given the unanimity reached through the analysis of time series, however, it seemed to be worthwhile to check whether their conclusions hold after correcting for possible weaknesses in their underlying models or, on the contrary, were caused by those purported flaws.

Three key points are at issue here: those of reverse causation, control-variables' selection, and the indirect effect of the government share through the investment rate. This article tackles with these issues (especially with 1 and 3) through the formulation of a simultaneous-equation model and the use of 2SLS. Due partly to the characteristics of the available data, the choices we have made have resulted into a rather strict test of the conclusions of prior contributions:

¹⁷ Use of the weak instruments obtained for this country also point in the same direction, however.

- The use of growth rates implies that we are estimating short-term elasticities only.
- Certain control variables included in second-stage regressions may bias toward zero the values and the levels of significance of the elasticities to be estimated.
- Government expenditures are represented by public consumption expenditures, which is not the heading that has increased the most during the period.
- And the post-2008 years have been excluded, during which the possible oversizing of the public sector may have been magnified by shrinking private production.

Moreover, the nature of our exercise imposes the restriction of estimating single arch-elasticities throughout the whole period under analysis. Strictly speaking, this restriction may be hardly compatible with the notion of a functional relationship that mimics the shape of an inverted U. It may be admissible for those countries that have remained on the descending branch of their curves during the whole period under analysis, but it is less adequate when applied to countries that may have moved around their tops. This may help to explain why, in certain countries where the vast majority of the estimated coefficients are negative (though close to zero), these fail to be significant at conventional levels.

The exercise performed here provides with estimates of the elasticity of GDP per worker to changes in the relative size of public consumption expenditure in 24 OECD countries. The number of estimates varies from country to country, from a minimum of 23 (for Mexico) to a maximum of 176 (for France). In all, we have made 1758 valid estimations. In none of these a positive and statistically significant elasticity of GDP to the share of government has been obtained. On the contrary, 872 estimates (that is, almost exactly 50 per cent) are both negative and significant at least at the 0.1 level.

In an exam country by country, the elasticities estimated are all negative for 10 out of 24 economies, and overwhelmingly negative in most others. Since there are no countries for which *all* estimates are both negative and significant at conventional levels, skeptics have the right to remain unconvinced. Nonetheless, other indications such as the proportion of negative coefficients estimated for each country, that of significant coefficients, and the degree of coincidence of the estimates provide strong evidence in favor of the conclusion of a negative elasticity in roughly half the countries that form the sample. In the rest of these, the evidence also favors the conclusion of a negative relationship, with the possible exception of only a few cases. A likely conclusion is, thus, that most OECD countries increased their public expenditures during the period under analysis beyond that level that was optimum for their economic growth, and remained above that optimum at least during most years of the period under analysis.

This conclusion serves, in my opinion, as confirmation that the inverse relationships between the two variables estimated by other authors who have performed prior-time series analysis, referred to the same countries, are valid in general terms. The conclusions of time-series analyses thus join those of panel-data analyses and, for the case, those of most cross-sectional analyses published since the 1980s. The elasticities estimated here are generally higher than those reported by most panel-data analyses,

but similar to the ones estimated by Dar and AmirKhalkhali (2001) for roughly the same sample of developed countries.

The results are broadly consistent with Armeij's hypothesis, although we have detected several seemingly deviant cases. These may be due to the action of other variables not controlled for, to a preference of certain small-sized governments for non-productive expenditures, or to our focus in the short run. When regressing rate of growth on rate of growth we are reproducing almost literally the expression of the elasticity, but the use of yearly observations implies that we are estimating short-run elasticities. If an increase in the government share fails to exert a durable effect on GDP, then the long-run elasticity would be lower than the short-term one, and the absolute level of G/Y would be irrelevant for GDP growth whereas its growth rate is. However, it is difficult to see why the long-term effect of public sector size could be lower than is the short-term one.

This question aside, the conclusion of the article is that, with a degree of certainty that varies from country to country, a reduction in their current levels of public consumption expenditures could have enabled most OECD countries to enhance GDP growth during the period under analysis. This conclusion refers to a period in whose final years the countries' GDPs were inflated by the boom of residential construction and production of other durables. This boom was the result of expansive monetary policies. However, once the sound demand for those goods became satiated, the crisis followed in both the real and the financial sector of the economy. If the booming national economies prior to 2008 already had problems with their overgrown public sectors, it is not difficult to imagine what may be happening in the subsequent period.

References

- Afonso, A., Furceri, D., 2010. "Government size, composition, volatility and economic growth.", *European Journal of Political Economy* 26: 517-532.
- Afonso, A., Schuknecht, L., Tanzi, V., 2003. "Public sector efficiency: An international comparison.", European Central Bank Working Paper No. 242.
- Afonso, A., Tovar Jalles, J., 2011. "Economic performance and government size.", European Central Bank Working Paper No. 1399.
- Agell, J., Ohlsson, H., Skogman Thoursie, P., 2006 "Growth effects of government expenditure and taxation in rich countries: A comment.", *European Economic Review* 50: 211-219.
- Angelopoulos, K., Philoppopoulos, A., 2007. "The growth effects of fiscal policy in Greece 1960-2000.", *Public Choice* 131: 157-175.
- Angrist, J.D., Pischke, J.-S., 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press. See also the authors' tag posted at www.mostlyharmlesseconometrics.com/2009/10/multivariate-first-stage-f-not/

- Arcand, J.-L., Dagenais, M.-G., 1995. "The empirics of economic growth in a section of countries: Do errors in variables really not matter?", Montréal: Université de Montréal. Faculté des Arts et des Sciences. Département de Sciences Économiques, CAHIER 9546.
- Armey, R., 1995. *The Freedom Revolution*. Washington, D.C.: Regnery Publishing Co.
- Barro, R.J., 1989. "A cross-country study of growth, saving and government." Washington: National Bureau of Economic Research Working Paper No. 2855.
- Barro, R.J., 1990. "Government spending in a simple model of endogenous growth.", *Journal of Political Economy* 98 (5, part 2): S103-S125.
- Barro, R.J., 1991. "Economic growth in a cross section of countries.", *Quarterly Journal of Economics* 1991: 407-443.
- Baumol, W.J., 1972. *Welfare Economics and the Theory of the State*. Cambridge, Mass.: Harvard University Press.
- Bergh, A., Henrekson, M., 2011. "Government size and growth: A survey and interpretation of the evidence. *Journal of Economic Surveys* 25: 872-897.
- Bergh, A., Karlsson, M., 2010. "Government size and growth: Accounting for economic freedom and globalization.", *Public Choice* 142: 195-213.
- Bergh, A., Öhrn, N., 2011. "Growth effects of fiscal policies: A critical appraisal of Colombier's study.", Stockholm: Research Institute of Industrial Economics IFN Working Paper No. 858.
- Branson, J., Lowell, C.A.K., 2001. "A growth maximising tax structure for New Zealand.", *International Tax and Public Finance* 8: 129-146.
- Carlstrom, C.T., Gokhale, J., 1991. "Government consumption, taxation, and economic activity.", *Economic Review – Federal reserve Bank of Cleveland* 27 (3): 18-29.
- Chao, J.C.P., Grubel, H., 1998. "Optimal levels of spending and taxation in Canada.", In H. Grubel (ed.) *How to Use the Fiscal Surplus. What is the Optimal Size of Government?* Vancouver: The Fraser Institute: 53-68.
- Chapple, S., 1997. "One for the X-files: A critical assessment of Professor Scully's 'Taxation and economic growth in New Zealand'. Report to the Periodic Report Group. Wellington: New Zealand Institute of Economic research.
- Chobanov, D., Mladenova, A., 2009. "What is the optimum size of government?" Sofia: Institute for Market Economics.
- Colombier, C., 2009. "Growth effects of fiscal policies: An application of robust-modified M-estimator.", *Applied Economics* 41: 899-912.
- Dar, A., Amir Khalkhali, S., 2002. "Government size, factor accumulation, and economic growth: Evidence from OECD countries.", *Journal of Policy Modeling* 24: 679-692.
- Doppelhoffer, G., Miller, R., Sala i Martin, X., 2004. "Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach.", *American Economic Review* 94: 813-835.
- Easterly, W., Levine, R., 2001. "It's not factor accumulation: Stylized facts and growth models.", *World Bank Economic Review* 15: 177-219.

- Easterly, R., S. Zerbos, 1993. "Looking at the facts. What we know about policy and growth from cross-country analysis.", Washington: The World Bank, Policy Research Working Paper WPS 1115.
- Facchini, F., Melki, M., 2012. "Optimal government size and economic growth in France (1871-2008): An explanation by the state and market failures". Paper presented at the 2nd World Congress of Public Choice Societies. Miami, 8-11 March 2012.
- Fölster, S., Henrekson, M., 2001. "Growth effects of government expenditure and taxation in rich countries.", *European Economic Review* 45: 1501-1520.
- Grossman, P.J., 1987. "The optimal size of government.", *Public Choice* 53:131-147.
- Grossman, P.J., 1988. "Growth in government and economic growth: The Australian experience.", *Australian Economic Papers* 27: 33-45.
- Heitger, B., 2001. "The scope of government and its impact on economic growth in OECD countries.", Kiel: Institute of World Economics, Kiel Working Paper No. 1034.
- Hill, R., 2008. "Optimal taxation and economic growth: A comment.", *Public Choice* 134: 419-427.
- Josheski, D., Lazarov, D., Koteski, C., 2011. "Analysis of the optimal size of the government consumption.", Munich: MPRA Paper No. 32983.
- Kennedy, P.E., 2000. "On measuring the growth-maximizing tax rate.", *Pacific Economic Review* 5:89-91.
- Leach, G., 2003. "The Negative Impact of Taxation on Economic Growth". London: Reform.
- Levine, R., Renelt, P., 1992. "Cross-country studies of growth and policy. Methodological, conceptual and statistical problems.", World Bank Policy Research Working Paper WPS 608.
- Levine, R., Zerbos, S., 1993. "Looking at the facts. What we know about policy and growth from cross-country analysis.", Washington: World Bank Policy Research Working Paper WPS 115.
- Magazzino, C., 2008. "Modeli d'interpretazione della dinamica delle spese pubbliche ed 'curva de Armey': Il caso italiano 1862-2001", *Notizie di Politeia* XXIV n° 92: 45-60.
- Magazzino, C., Forte, F., 2010. "Optimal size of government and economic growth in EU-27.", Munich: MPRA Paper No. 26669.
- Mueller, D.C., 2003. *Public Choice III*. Cambridge: Cambridge University Press.
- Nijkamp, P., Poot, J., 2004. "Meta-analysis of the effect of fiscal policies on long-run growth.", *European Journal of Political Economy* 20: 91-124.
- 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: 153-173.
- Peden, E.A., Bradley, M.D., 1989. "Government size, productivity, and economic growth: The post-war experience.", *Public Choice* 61: 229-245.
- Pevcin, P., 2004. "Does optimal size of government spending exist?", Paper presented to the European Group of Public Administration (EGPA) Annual Conference. Ljubljana, Slovenia, 1-4 September 2004.

- Poot, J., 2000. "A synthesis of empirical research on the impact of government on long-run growth.", *Growth and Change* 31: 516-546.
- Rahn, R., Fox, H., 1996. *What is the Optimum Size of Government*. Denver, CO: Vernon K. Kriebel Foundation.
- Ram, R., 1986. "Government size and economic growth: A new framework and some evidence from cross-country and time-series data.", *American Economic Review* 76: 191-203.
- Raymond Bara, J.L., 1992. "Gasto público y crecimiento. Un análisis de los efectos del tamaño del sector público en España y en la Europa comunitaria.", *Papeles de Economía Española* 52/53: 180-196.
- Romero Ávila, D., Strauch, R., 2008. "Public finances and long-term growth in Europe: Evidence from a panel data analysis.", *European Journal of Political Economy* 24: 172-191.
- Sala i Martin, X., 1997. "I just ran two million regressions.", *American Economic Review 87 Papers and Proceedings*: 178-183.
- Scully, G.W., 1989. "The size of the state, economic growth and the efficient utilization of national resources.", *Public Choice* 63: 149-164.
- Scully, G.W., 1991. "Tax rates, tax revenues and economic growth." Washington and Dallas: National Center for Policy Analysis, Policy Report 98.
- Scully, G.W., 1994. "What is the optimal size of government in the United States?" Washington and Dallas: National Center for Policy Analysis, Policy Report 188.
- Scully, G.W., 1995. "The 'growth tax' in the United States.", *Public Choice* 85: 71-80.
- Scully, G.W., 1996. "Taxation and economic growth in New Zealand.", Wellington: Inland Revenue, Working Papers on Monitoring the Health of the Tax System 14.
- Scully, G.W., 2003. "Optimal taxation, economic growth and income inequality.", *Public Choice* 115: 299-312.
- Scully, G.W., 2008. "Optimal taxation, economic growth and income inequality in the United States.", Washington and Dallas: National Center for Policy Analysis, Policy Report No. 316.
- Sieper, E., 1997. "Review by E. Sieper of Gerald W. Scully, 'Taxation and economic growth in New Zealand.'", Wellington: Revised IRD Working Paper No. 14. Study Commissioned for the Treasury. Government of New Zealand.
- Slemrod, J., 1995. "What do cross-country studies teach about government involvement, prosperity, and economic growth?", *Brookings Papers on Economic Activity* 1995: 373-431.
- Slemrod, J., 1998. "How costly is a large, redistributive public sector?", *Swedish Economic Policy Review* 5: 87-105.
- Tanzi, V., Schuknecht, L., 2000. *Public Spending in the Twentieth Century: A Global Perspective*. Cambridge: Cambridge University Press.
- Taylor, J., 1993. "Discretion versus policy rules in practice.", *Carnegie-Rochester Series on Public Policy* 39 : 195-214.

Temple, J., 1999. "The new growth evidence.", *Journal of Economic Literature* 37: 112-156.

Vedder, R.K., Gallaway, L.E., 1998. "Government size and economic growth. Prepared for the Joint Economic Committee". Washington, DC: Joint Economic Committee.

ⁱ The author thanks Christian Bjørnskov, Carlos Díaz Vela, Kevin Grier, Ignacio Moral, Santiago Lago, Martin Rode and Blanca Sánchez-Robles for useful comments. He is also indebted to Vikram Nehru, from the World Bank, for data supplying. Earlier versions of this article were presented at the 2010 Conference of the Public Choice Society held in Monterey, CA, 11-14 March, and at the 2010 Conference of the European Public Choice Society held in Izmir, Turkey, 8-11 April.