

Public sector efficiency: leveling the playing field between OECD countries

Antonis Adam · Manthos Delis · Pantelis Kammas

Received: 19 February 2008 / Accepted: 18 December 2009 / Published online: 1 January 2010
© Springer Science+Business Media, LLC 2009

Abstract In this paper we seek a robust methodology for the measurement of the relative public sector efficiency of 19 OECD countries over the period 1980–2000. We estimate relative efficiency scores for five disaggregated accounts of public spending as well as for aggregate public spending. Then, we use a semi-parametric econometric method to isolate the impact of government inefficiency from the inefficiency arising from the socioeconomic environment and luck. To verify the validity of our index, we use it to examine a number of well-established relationships in the public choice literature, which have only been tested using local government data.

Keywords Public sector efficiency · Stochastic DEA · Political determinants of efficiency

JEL Classification C14 · H11 · H50

1 Introduction

It has long been recognized that the efficient function of the public sector is a requisite for a country's competent economic performance. The purpose of this paper is to provide a consistent methodology for the measurement of relative public sector efficiency (PSE) at the country level and to show that the indicators constructed can be valuable for the public choice literature in general, and the analysis of the political sources of inefficiency in particular.

The public choice literature has examined various forms of public sector inefficiency and one of the most important originates in the seminal work of Migué and Bélanger (1974) and Niskanen (1975). According to these early studies the bureau, acting within a principal-agent framework, attempts to maximize the difference between total revenues received and the minimum feasible cost of production, thus generating the so-called organizational inefficiency (or alternatively budgetary slack). This type of public sector inefficiency may

A. Adam · M. Delis · P. Kammas (✉)
University of Ioannina, Ioannina, Greece
e-mail: kammas@cc.uoi.gr

be attributed to various non-productive activities that may include (but are not limited to) personnel expansion à la Williamson (1964), lower effort (Wyckoff 1990), excessive risk aversion (Peltzman 1973) and corruption (for a comprehensive review, see Wintrobe 1997). Later theoretical contributions have analyzed various relationships between bureaucratic efficiency and elements of the institutional and the political environment (Breton and Wintrobe 1975; Moene 1986; Falch 2001).

To our knowledge, on the empirical front, the predictions of this literature have been tested solely on the basis of local government data (e.g., Hayes et al. 1998; Geys 2006; Balaguer-Coll et al. 2007; Barankay and Lockwood 2007; Borge et al. 2008).¹ However, while local governments normally face quite similar socio-economic environments (at least in small, geographically concentrated nations), the macroeconomic environments may differ substantially between countries; and this renders the validity of existing country-level efficiency indicators that do not account for these differences questionable. In other words, the basic shortcoming of the existing approaches to country-level efficiency measurement has been their inability to distinguish the inefficiency attributable to bad public sector managerial practices from the inefficiency arising from differences in socioeconomic environments or factors attributable to luck.

In this paper we try to remedy this shortcoming by estimating relative PSE scores purified from the impact of the socio-economic factors and luck, thus “leveling the playing field” across different countries. This is achieved by a three-stage analysis similar to that employed by Fried et al. (2002) and Glass et al. (2006).² Efficiency scores are constructed on the basis of “Classification of expenditure according to purpose” (i.e. the functional classification of public spending). That is, we employ five alternative accounts of public spending and we estimate efficiency scores for each of them. These accounts correspond to spending on: (i) education, (ii) health, (iii) social security and welfare, (iv) general public services and (v) economic affairs. In addition, we construct two general indicators of public sector efficiency that reflect benchmark tasks of government, namely economic performance and economic stability.

Our main results can be summarized as follows. The three-stage analysis indicates that exogenous socio-economic factors and factors attributed to luck, have less of an influence (compared to sound government managerial practices) on the efficient translation of government activities into services of interest to citizens. Subsequently, to examine the value of the new indices and to relate them to the public choice literature, we test a number of hypotheses that have been analyzed before on the basis of local data. The findings, in line with studies such as Barankay and Lockwood (2007) and Borge et al. (2008), suggest that countries with right-wing and strong governments, high voter participation rates and decentralized fiscal systems, are expected to have higher PSE.

The structure of the rest of the paper emerges along the following lines: Section 2 describes the methodology and the data employed for the efficiency estimates. Section 3

¹Only recently a number of studies cultivated an effort towards the computation of relative PSE indicators at the country level (Gupta and Verhoeven 2001; Sijpe and Rayp 2007; Afonso et al. 2005; Angelopoulos et al. 2008).

²In this sense, our paper is closely related to a large branch of public choice literature that emphasizes the importance of accounting for exogenous socio-economic factors on the translation of government activities into services of interest to citizens (see e.g., Bradford et al. 1969; Duncombe et al. 1997). In the same spirit, Tanzi and Schuknecht (2000: 76) note that “Governments cannot be assumed to be responsible for all the differences in the countries’ performance as represented by these indicators. In fact, differences between countries and changes over time often have a lot to do with technical progress or cultural differences between countries”.

presents the PSE estimates. Section 4 examines the value of the new indices by examining the political determinants of PSE, thus relating the present study to a broad swath of the public choice literature. Finally Sect. 5 concludes.

2 Measuring public sector efficiency

2.1 Methodology: leveling the playing field of public sector efficiency

In this study we focus on productive efficiency, which reflects the ability of a public sector (government) to maximize output for a given set of inputs. Governments that succeed are on the frontier, while those that do not optimize the use of their inputs are inefficient to a variable extent. This type of (in)efficiency can be measured in terms of the distance of a given public sector from the best-practice equivalent (which forms the frontier) and the resulting score is a scalar measure ranging between zero (the lowest efficiency score) and one (the best-practice public sector).

In the present study, we measure public sector efficiency along the lines of Fried et al. (2002) and Glass et al. (2006). In particular, we use a three-stage model, which involves a combination of data envelopment analysis (DEA) and stochastic frontiers. The merit of this model is that (i) it accounts for the effect of environmental variables on the estimated efficiency scores and (ii) incorporates statistical noise into DEA. In other words, this model ensures that governments with favorable initial socioeconomic environments or good fortune are not given an undue advantage. A detailed discussion of this methodology is offered in the aforementioned studies; here we discuss the three stages of the estimation procedure only briefly.³

In the first stage we employ simple input-oriented, variable returns to scale DEA to obtain an initial evaluation of PSE. DEA is a deterministic, linear programming technique that provides a piecewise frontier, by enveloping the observed data points, and yields a convex production possibilities set. Therefore, each score obtained reflects the distance between the respective data point (country) and the best practice country that shapes the frontier. However, the efficiency scores derived from this first stage are “gross” initial DEA estimates which include three components; (i) the pure efficiency of governments, (ii) the impact of the socioeconomic environment and (iii) the impact of luck. That is, governments with a favorable socioeconomic environment, or merely lucky governments, are given an undue advantage.

Thus, in the second stage we use the stochastic frontier approach to separate Stage 1 efficiency scores into efficiency attributable to government managerial practices and efficiency that is due to the macroeconomic environment and statistical noise (luck). In particular, we estimate a model of the form:

$$\begin{aligned} \text{eff}_{it} &= f(z_{it}; a) + v_{it} + u_{it} \\ u_{it} &= \beta(t)\mu_i \end{aligned} \quad (1)$$

³Another method proposed in the recent econometrics literature is that of Simar and Wilson (2007), who suggest using a double bootstrapping procedure to analyze the determinants of efficiency scores obtained from linear programming techniques. The main reasons we used the Fried et al. (2002) method are that first we are more interested here in leveling the playing field between countries and second the fact that the method of Simar and Wilson (2007) is carried out for cross-sectional data. Extension of this method to panel data is probably not trivial and beyond the scope of the present study.

where eff represents the vector of efficiency scores obtained in the first stage, z_{it} are country-specific variables that characterize the socioeconomic environment, a are parameter vectors to be estimated, v_{it} represents random statistical noise, u_{it} represents inefficiency, and the function $\beta(t)$ is specified as a set of time dummy variables β_t . The β_t s are treated as the coefficients of the fixed effects μ_i , and once both are estimated (by maximum likelihood techniques), u_{it} is obtained as $u_{it} = \max_i \{\hat{\beta}_t \hat{\mu}_i\} - (\hat{\beta}_t \hat{\mu}_i)$. For the distribution of the error term we assume $v_{it} \sim N(0, \sigma^2)$, but no distributional assumption is made on u_{it} , which is allowed to be correlated with the control variables or with the remainder disturbance.⁴

The inclusion of vector z_{it} is crucial as it allows us to account for: (i) country-specific socioeconomic characteristics and (ii) economies of scale (the set of control variables z_{it} is described in detail in Sect. 2.2). Moreover the introduction of the fixed effect μ_i controls for the time invariant institutional characteristics. These two vectors correspond to what we call “socioeconomic environmental variables”.

The stochastic frontier regression model is very similar to that of Fried et al. (2002). The impact of both the environment and noise on Stage 1 efficiency scores is captured by the deterministic frontier $f(z_{it}; a) + v_{it}$, while u_{it} represents the government inefficiency disentangled from the impact of the socioeconomic environment and luck. This inefficiency u_{it} is based on the ability of public sectors to adapt to the vagaries of the socioeconomic environment and luck but does not account for the ability of some public sectors to conserve input use. That is because the original input data implicitly are used in estimating (1), which would probably be different in light of a different socioeconomic environment or luck scenario.

Therefore, to account for this last problem, the results derived from (1) for a, β, u_{it} and v_{it} are used to penalize public sectors that have been advantaged by their relatively favorable economic environments, their relatively good luck, or both. To this end, we adjust upwards the inputs used by governments relative to the least favorable environment and most unlucky government. Governments’ adjusted inputs are constructed by means of

$$x_{it}^A = x_{it} + [\max_{it} \{z_{it} \hat{\beta}\} - z_{it} \hat{\beta}] + [\max_{it} \{\hat{v}_{it}\} - \hat{v}_{it}] \tag{2}$$

where x_{it}^A and x_{it} are adjusted and observed inputs, respectively. The first adjustment, i.e. the second term on the right side of (2), places all governments into a common socioeconomic environment, which is the least favorable environment encountered in the sample. The second adjustment, i.e. the third term on the right side of (2), places all governments into a common “state of nature”, which is the unluckiest situation encountered in the sample. This implies that potential asymmetric effects of macroeconomic shocks across countries are taken into account and PSE is no longer affected by them.

To implement (2) it is mandatory to separate managerial inefficiency u_{it} from the remainder disturbance v_{it} (luck) in the residuals of (1) in order to obtain estimates of v_{it} for each government. As outlined in Fried et al. (2002), the conditional estimators for managerial inefficiency, given by $\hat{E}[u_{it}|v_{it} + u_{it}]$, enable estimators for statistical noise to be derived residually via

$$\hat{E}[v_{it}|v_{it} + u_{it}] = Eff_{it} - z_{it} \hat{\beta} - \hat{E}[u_{it}|v_{it} + u_{it}] \tag{3}$$

which provide, conditional on $v_{it} + u_{it}$, estimators for v_{it} in (2).

⁴The virtue of using the stochastic frontier approach, as opposed to a limited dependent variable approach, is that the error term is asymmetric, with μ_i following a fixed effects model. However, this model may be restrictive in some respects, if one considers that a specific linear functional form is imposed and that certain assumptions are made regarding the error term. For more on these issues, see Kumbhakar et al. (2007) and Kumbhakar and Tsionas (2005).

Finally, in Stage 3 we repeat the exercise of Stage 1, this time using the adjusted input data x_{it}^A that resulted from the analysis of Stage 2, instead of the original input data. Thus, through this procedure we are able to obtain efficiency scores solely attributable to government managerial practices, which may be directly compared between countries.

A couple of issues probably require some clarification regarding the procedure described above. First, the model will be estimated on ten- and five-year periods, so as to smooth out any country-specific socioeconomic or other shocks that may be correlated with efficiency scores. Second, one may suggest that some endogeneity problems exist in the estimation procedure; for example, it may be the case that less educated and/or more unwell populations cause more education and health spending by the government. Note, however, that this type of endogeneity is between inputs and outputs of the production process, not between any of the inputs or outputs and the error components u_{it} and v_{it} . In addition, if this is an intense short-term phenomenon, it will probably be smoothed out by the use of ten- or five-year periods in the estimation procedure. For more on these issues, see Fried et al. (2002) and Glass et al. (2006).

2.2 Data

The estimation methodology described above is applied on a panel of 19 OECD countries over two decades, the 1980s and the 1990s. These countries are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Luxembourg, Norway, Portugal, Spain, Sweden, Switzerland, The Netherlands, United Kingdom and United States. Using decade averages, instead of higher frequency data, is preferable for three main reasons. First, the effect of various policies is not readily observable, but instead may take time to settle in. Second, most areas of government activity have long-run effects. For example, an increase in educational infrastructure will affect the output of the educational system for more than a single year. Finally and most importantly, we want to exclude potential business cycle effects both from the inputs (public spending) and from the outputs, which consist of economic variables that are highly dependent on the economic cycle (e.g., unemployment, inflation).

Since OECD economies differ substantially in the composition of government spending, developing and examining disaggregated PSE indices appears to be more than important. Otherwise, we may lose valuable information related to public sector efficiency. Therefore, to measure PSE we employ as inputs five alternative accounts of public spending. These accounts correspond to spending on: (i) education (denoted as *EducSp*), (ii) health (*HealthSp*), (iii) economic affairs (*EcaffSp*), (iv) general public services (*GPSSp*), and (v) social security and welfare (*WelfSp*). These are five out of the six largest public spending categories.⁵ The only financially important spending category that we have not included in this study is public spending on defence. This is because it was fairly difficult to substantiate the output of this spending category. Missing data for public spending on health for Belgium in the 1980s and on economic affairs for both Belgium and Portugal in the 1990s prevent us from deriving relevant efficiency estimates for these countries. Data for *EducSp* are obtained from the World Bank's Development Indicators (2004), whereas data for *HealthSp* and *WelfSp* are taken from the OECD's Health Data (2006) and the OECD's Social Expenditure Database (2000), respectively. Finally, data for *GPSSp* and *EcaffSp* are taken from IMF's Government Finance Statistics Database (2002).⁶

⁵For more details on the classification of the functions of government see Classification of Expenditure According to Purpose (United Nations 2000).

⁶All public spending data are consolidated expenditures (over all tiers of government).

Obviously, to measure PSE some “performance measures” are required that could be interpreted as outputs of public spending. Following the reasoning of the relevant literature (see e.g., Afonso et al. 2005; Angelopoulos et al. 2008), we employ the following measures for each category respectively: As an output of *EducSp* we use the gross secondary school enrollment (defined as the ratio of total enrollment, regardless of age, to the population of the age group that officially corresponds to secondary schooling age). In order to capture the potential qualitative differences on education among different countries, we additionally employ as an output of *EducSp* the educational quality indicator developed by Hanushek and Kimko (2000). Also, since the educational systems of the OECD countries examined are far from being homogeneous in the sources of spending (private or public), it is important to account for the different ratios of private to public spending on education.⁷ As such, we multiply the outputs of *EducSp* (i.e. the secondary school enrollment and the quality of education) by the public to total spending ratio. By doing so we isolate the impact of private expenditure on output, and consequently we don’t give countries characterized by heavy private funding on education an undue advantage.⁸

Life expectancy at birth and the inverse of the infant mortality rate (both taken from World Bank’s Development Indicators 2004) are employed as outputs of *HealthSp*. As with the case of education, we account for differences in the ratios of private to public spending on health by multiplying the outputs of *HealthSp* by the ratio of public to total spending on health.⁹ The inverse of income inequality is assumed to be the output of *WelfSp* and is measured by the GINI coefficient taken from the World Institute for Development Economics Research (Luxembourg Income Study) of the United Nations University World Income Inequality Database (2007).

As outputs of *EcaffSp* we use electric power transmission and distribution losses as a share of total electric power output (obtained by World Bank’s Development Indicators 2004), and the standard telephone access lines per 100 inhabitants (taken from the OECD’s Telecommunications Database 2005). These measures reflect the quality of the electricity and telecommunications infrastructures. Our decision to employ these variables as outputs is based on the fact that Economic Affairs spending account includes as major sub-components (i) “spending on construction and operation of electricity supply systems” and (ii) “spending on construction, improvement, operation and maintenance of communication systems” (United Nations 2000).

Finally, we employ as outputs of *GPSSp* (i) the inverse of corruption in government and (ii) bureaucratic quality, both obtained from the International Country Risk Guide (2006) database. The “corruption in government” and “bureaucratic quality” variables are used by Afonso et al. (2005) and Angelopoulos and Philippopoulos (2005) as proxies for public administration performance. Since *GPSSp* is a public expenditure account mainly financing public administration, we employ these measures as outputs of the *GPSSp*.

⁷In our sample there are countries characterized by heavy private funding on education (such as Australia or United States) and countries that base the financing of their educational systems on public funds (such as Finland and Denmark).

⁸We note that alternatively the public to total spending ratio could be included as a variable in the second stage correction. However, to our perception is better to derive the first stage PSE scores with the correct measure of output (as far as this is possible). We thank an anonymous referee for bringing this to our attention.

⁹The ratio of public to total spending on health is only 41.6% in United States and 51.9% in Greece, while it reaches 88.42% and 87.32% in Sweden and Norway, correspondingly.

Apart from estimating the five disaggregated PSE indicators described above, we also construct two more general indicators of public sector efficiency that reflect benchmark objectives (or alternatively tasks) of government, namely (i) general economic performance and (ii) economic stability. As Tanzi and Schuknecht (2000: 75) state: “It is difficult or even impossible to consider all the social and economic objectives (and thus all the socioeconomic indicators) that the governments might want to influence with this spending. By necessity, the analysis will include fewer indicators than might have been desirable to include”. Following the reasoning of the relevant literature (see e.g., Afonso et al. 2005) the “performance measures” employed as outputs in the case of general economic performance (denoted as *EcPerf*) are the unemployment rate, the annual GDP growth rate and GDP per capita. Lower scores on the unemployment rate and higher scores in GDP growth and GDP per capita reflect better economic performance. The corresponding outputs in the case of economic stability (denoted as *EcStab*) are the standard deviation of the GDP growth rate, which is interpreted as a measure of economic fluctuations, and the inflation rate. Lower scores for both measures denote improved economic stability. For both measures the relevant spending account that serves as input is total public spending; the data for all the variables related to performance and stability are obtained from World Bank’s Development Indicators (2004).

Except for the variables described so far, in the second stage of the estimation procedure we additionally use a set of control variables (specific to each sub-component of public spending), to carry out the stochastic frontier regressions. This set of control variables captures elements of socioeconomic conditions and economies of scale in the production of public goods and is similar to that employed by the associated literature (see e.g., Hayes et al. 1998; Barankay and Lockwood 2007). More precisely, the efficiency scores for general public services obtained from Stage 1 are regressed on the urbanization rate of each country considered (obtained from World Bank’s Development Indicators 2004); the efficiency scores for education spending on the International Country Risk Guide (2006) socioeconomic condition index and the population density rate (also obtained from World Bank’s Development Indicators 2004); the efficiency score for public health on the socioeconomic conditions indicator, population density and the fraction of the population over 65 years old (obtained from World Bank’s Development Indicators 2004); the efficiency scores for economic affairs on the International Country Risk Guide (2006) index of investment profile, the socioeconomic conditions indicator and on the capital market international integration proxy constructed by Quinn (1997); the efficiency scores for social spending on the socioeconomic conditions and the urbanization rate; and finally in the cases of the efficiency scores for economic performance and stability we employ the socioeconomic conditions and capital market international integration as the socioeconomic control variables.

Table 1 reports the inputs and outputs used, along with some descriptive statistics. Even though the sample of countries considered is rather homogeneous, the size and mix of government spending in these countries are quite different. Thus, there are countries characterized by relatively large public sectors, whose government spending exceeds 40% as a share of GDP (e.g., the Netherlands, Sweden and Denmark), and countries with a public sector size that does not exceed 25% as a share of GDP (e.g., Austria, Switzerland and United States). Moreover, as regards the composition of government expenditures, there are countries that present a clear-cut priority towards social security and welfare spending (e.g., the Nordic countries), and others characterized by heavy spending on public infrastructure (such as Ireland, Norway and Finland, whose public spending on economic affairs exceeds 5% as

Table 1 Summary statistics for outputs and inputs

	Mean	Std. Dev.	Min	Max
Outputs				
Secondary school enrollment (% of gross enrollment)	102.46	15.87	72.9	136.5
Quality of education	55.71	5.65	44.5	64.6
Life expectancy at birth	75.94	1.44	73.31	78.56
Infant mortality rate	8.44	2.12	5.00	12.00
Electric power transmission and distribution losses (% of total electric power output)	7.25	3.14	3.85	22.5
Standard access lines per 100 inhabitants	45.86	10.56	19.58	67.98
Corruption in government	5.42	0.605	4.25	6.00
Bureaucratic quality	3.89	0.246	3.00	4.00
GINI coefficient	29.33	4.74	21.5	38.97
Unemployment rate	7.30	3.67	0.59	15.15
GDP growth rate (annual %)	2.72	1.16	0.93	7.1
GDP per capita (constant 1995 US\$)	25331.79	8036.63	11611	45210
Standard deviation of the annual GDP growth rate	1.65	0.544	0.89	2.98
Inflation rate (consumer prices annual %)	4.45	2.52	1.89	10.25
Inputs				
Public spending on education (% of GDP)	5.60	1.10	3.1	7.8
Public spending on health (% of GDP)	5.83	1.03	4.08	7.94
Public spending on economic affairs (% of GDP)	3.99	1.91	1.45	8.00
Public spending on general public services (% of GDP)	2.07	0.83	0.92	3.9
Public spending on social security and welfare (% of GDP)	16.95	4.78	7.76	26.28
Total public spending (% of GDP)	36.15	9.25	16.43	54.01
'Leveling the playing field' variables				
Quality of socioeconomic environment index (Socioeco)	7.74	1.40	4.5	11
Protection of investment index (Invprot)	7.83	1.58	4	11
Population density-people per sq Km (Density)	133.47	119.06	1.97	460.96
Share of population age 65+ to total population (Old)	13.76	1.97	9.53	17.88
% of total population in urban areas (Urban)	74.69	12.89	32.39	97.11
Openness of capital markets index (capopen)	84.62	15.95	37.5	100

a share of GDP).¹⁰ These differences may have appealing implications for the results of the empirical analysis that follows.

3 Efficiency estimates

Space constraints prevent reporting the results of all three stages of analysis, as this would triple the number of tables. Thus, only limited information is provided for the Stages 1 and 2

¹⁰Tanzi and Schuknecht (2000) note that: “[...] the main difference between big-government and small-government OECD countries lies in the spending for social security and welfare. In the 1990s, social security related expenditure in big-government countries was two and one-half times higher than in small-government countries (that is an average of 0.20 as a share of GDP versus 0.08 as a share of GDP)”.

results in the text to allow comparison with the Stage 3 results. The full set of Stage 3 relative PSE scores is reported in Tables 2a and 2b for the 1980s and the 1990s, correspondingly.¹¹ The results are reported for each of the 19 countries of our sample and for each disaggregated public spending account. In the last two columns of Table 2 we present the results for the two general indicators of public sector efficiency that reflect efficiency in achieving competent general economic performance and economic stability. Sensitivity analyzes performed on the Stage 3 results showed that efficient public sectors remained efficient to any simultaneous data changes in the respective inputs (for a detailed discussion of the sensitivity analysis on DEA estimates, see Zhu 2003).

Let us first turn to the discussion of the ranking of countries as derived from the Stage 3 results. Concerning the PSE of education, the frontier is shaped by a number of countries (Denmark, Finland, Norway, Greece, Norway and United States) in the 1980s, and the general ranking remains practically the same in the 1990s. Switzerland, United Kingdom and Spain attain relatively high efficiency scores, while Luxembourg and Germany are the last countries in the ranking. The frontier of the efficiency of public spending on health is shaped by Luxembourg, Portugal and Sweden in the 1980s, and Greece, Ireland, Luxembourg and Sweden in the 1990s. Switzerland, Australia and Spain are ranked among the good performers, whereas France, Germany and Sweden are last in the countries' ranking.

Turning to the spending on economic affairs we note that there exist relatively large differences between the many efficient countries (i.e., countries on the frontier) and countries characterized by considerable inefficiency. As regards the efficiency scores of general public services, the frontier in the 1980s is shaped by Switzerland, with Germany, United States and Spain. In contrast, Luxembourg and Greece appear to be the most inefficient countries. In the 1990s Canada accompanies Switzerland in shaping the frontier, with the general ranking changing only in few cases. Note that since most of the OECD countries examined present satisfactory scores in the outputs of this public sector category, differences in the efficiency scores are driven mainly by differences in spending. Thus, Luxembourg is the last country in the general ranking not because it is characterized by high corruption in government or unsatisfactory bureaucratic quality, but because it spends relatively more in order to achieve these goals.

Social security and welfare is the largest public spending account. According to Tanzi and Schuknecht (1998) this spending account increased the most in the last 35 years, from an average of less than 10% of GDP in 1960 to 23% of GDP in 1995, which corresponds to about three quarters of the total public expenditure increase since 1960. Thus, if one wants to examine the evolution of total spending and to explain its rising trend during the last four decades it is crucial to focus on this account. In this spending category the frontier is shaped by Australia and Finland in the 1980s, and Australia, Belgium, Denmark and United States in the 1990s. Moreover, the countries' ranking does not seem to change substantially through time.¹² Australia, Canada, Finland, Norway and United States present relatively high efficiency scores in both periods, whereas the scores for France, Ireland and The Netherlands reflect poor performance.

The second to last columns of Tables 2a and 2b report the relative efficiency scores for total spending when the government target is economic performance. Under this government

¹¹ The full set of Stages 1 and 2 results is available upon request. Since the outputs used to estimate efficiency do not vary substantially over time we preferred to report the ten-year period results. Yet, the estimations for five-year time periods we use in the following section give us similar results and are also available upon request.

¹² The only exception worth noting is Denmark, which appears to be one of the worst performers in the 1980s and one of the best in the 1990s.

Table 2a Stage 3 efficiency scores by public spending account for 19 OECD countries, 1980–1990

	Education		Health		Economic affairs		General public services		Social security and welfare		Performance		Stability	
	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank
Australia	0.579	12	0.779	7	0.933	2	0.409	7	1.000	1	1.000	1	1.000	1
Austria	0.762	8	0.750	9	0.432	8	0.371	9	0.822	6	0.615	11	0.841	2
Belgium	0.795	5			0.324	10	0.333	12	0.688	10	0.621	10	0.701	9
Canada	0.787	6	0.630	14	0.501	7	0.378	8	0.823	5	0.753	6	0.755	6
Denmark	1.000	1	0.580	16	0.660	4	0.249	15	0.586	14	0.536	14	0.523	13
Finland	1.000	1	0.761	8	0.356	9	0.346	11	1.000	1	0.784	4	0.825	3
France	0.644	9	0.687	13	0.724	3	0.285	14	0.421	17	0.612	12	1.000	1
Germany	0.409	14	0.580	15	1.000	1	0.775	2	0.679	12	0.729	7	1.000	1
Greece	1.000	1	0.834	4	0.282	11	0.211	18	0.639	13	0.762	5	0.756	5
Ireland	0.593	11	0.697	12	0.227	15	0.212	17	0.559	16	0.574	13	0.574	11
Luxembourg	0.348	15	1.000	1	0.227	14	0.196	19	0.708	9	1.000	1	0.546	12
Norway	1.000	1	0.730	11	0.271	12	0.417	6	0.961	2	0.722	8	0.606	10
Portugal	0.523	13	1.000	1	0.269	13	0.311	13	0.885	3	0.822	3	0.822	4
Spain	0.826	4	0.908	2	0.512	6	0.670	4	0.576	15	0.852	2	1.000	1
Sweden	0.618	10	1.000	1	1.000	1	0.359	10	0.724	8	0.524	15	0.732	8
Switzerland	0.947	2	0.804	5	1.000	1	1.000	1	0.681	11	1.000	1	1.000	1
The Netherlands	0.770	7	0.749	10	1.000	1	0.239	16	0.343	18	0.509	16	1.000	1
United Kingdom	0.835	3	0.872	3	0.562	5	0.510	5	0.735	7	0.721	9	0.749	7
United States	1.000	1	0.789	6	1.000	1	0.689	3	0.837	4	1.000	1	1.000	1
Average	0.760		0.786		0.594		0.419		0.719		0.744		0.812	
Spearman correlation with Stage 1 results	0.852		0.925		0.942		0.896		0.866		0.913		0.902	

Table 2b Stage 3 efficiency scores by public spending account for 19 OECD countries, 1990–2000

	Education		Health		Economic affairs		General public services		Social security and welfare		Performance		Stability	
	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank
Australia	0.843	5	0.865	4	0.836	3	0.390	12	1.000	1	0.942	2	1.000	1
Austria	0.780	8	0.758	10	0.470	8	0.287	14	0.502	16	0.636	10	0.702	10
Belgium	1.000	1	0.717	11	1.000	1	1.000	1	1.000	1	0.798	3	1.000	1
Canada	1.000	1	0.675	14	0.858	2	1.000	1	0.843	5	0.786	4	0.985	2
Denmark	0.795	7	0.682	13	0.798	4	0.419	9	1.000	1	0.597	12	0.767	7
Finland	1.000	1	0.771	9	1.000	1	0.683	4	0.867	4	0.515	14	0.656	12
France	0.651	12	0.664	15	0.619	5	0.239	16	0.504	15	0.571	13	1.000	1
Germany	0.564	15	0.582	16	1.000	1	0.771	2	0.970	2	0.696	7	0.775	6
Greece	1.000	1	1.000	1	0.347	9	0.393	10	0.577	11	0.731	6	0.759	8
Ireland	0.661	11	1.000	1	0.233	11	0.156	17	0.686	10	1.000	1	0.805	4
Luxembourg	0.695	9	1.000	1	0.221	12	0.251	15	0.839	6	1.000	1	0.757	9
Norway	1.000	1	0.689	12	0.235	10	0.442	8	0.878	3	0.639	9	0.606	13
Portugal	0.636	13	0.947	2	0.764	3	0.764	3	0.772	7	1.000	1	1.000	1
Spain	0.690	10	0.864	5	0.499	7	0.669	5	0.687	9	0.769	5	0.782	5
Sweden	1.000	1	1.000	1	1.000	1	0.584	6	0.691	8	0.495	15	0.487	14
Switzerland	0.952	3	0.910	3	1.000	1	1.000	1	0.562	12	1.000	1	0.819	3
The Netherlands	0.816	6	0.859	6	1.000	1	0.390	11	0.516	13	0.629	11	1.000	1
United Kingdom	0.941	4	0.853	7	0.511	6	0.492	7	0.514	14	0.656	8	0.664	11
United States	0.970	2	0.819	8	1.000	1	0.336	13	1.000	1	1.000	1	1.000	1
Average	0.842		0.824		0.684		0.540		0.758		0.761		0.819	
Spearman correlation with stage 1 results	0.925		0.896		0.873		0.913		0.798		0.902		0.902	

target the frontier is shaped by Australia, Switzerland, Luxembourg and United States in the 1980s, and Switzerland, Ireland, Portugal and United States in the 1990s. Although the country rankings appear to change through time, we note that Australia, Canada, and Spain represent relatively efficient countries, whereas France, Denmark, Sweden and The Netherlands are characterized by poor performance. Yet, these results should be interpreted with caution since the countries examined present small differences in GDP per capita, as well as in their GDP growth and unemployment rates (assumed to be the outputs).

The last column reports the relative efficiency scores for total spending when the government target is economic stability. In this case Australia, Switzerland, France, The Netherlands, Spain and United States in the 1980s, and Australia, Belgium, France, The Netherlands and United States in the 1990s shape the frontier. Although the country ranking appear to change substantially through time, note that Canada, Switzerland and Germany perform well, whereas Norway and Sweden present low efficiency scores. Again these results should be interpreted with caution since most OECD economies are characterized by low inflation and stable growth rates during the period examined. Thus, countries with relatively large public sectors may present worse efficiency scores, if they achieve similar economic targets by “paying” relatively more for them.¹³

Comparing the Stage 1 with the Stage 3 results we note that leveling the playing field between OECD countries does not affect significantly the country ranking. The Spearman correlation of the rankings formed by the first and the third stage efficiency scores in both the 1980s and the 1990s (see last rows of Tables 2a and 2b) is quite high and always significant at the 5% level.¹⁴ This result has clear implications regarding the forces that shape PSE, since luck and superior socioeconomic environments appear to be less important than sound governance. In turn, this reflects the idea that governments have much to gain by observing and implementing the strategies followed by the efficient governments, in an attempt to improve their own PSE scores.

Given the rankings of the three-stage analysis, in what follows we try to identify the political and institutional sources of the observed differences, thus (i) inquiring into the usefulness of the indices, (ii) relating the findings on the efficiency scores to the relevant public choice literature and (iii) explaining the variation of *PSE* across countries/within time.

4 Political determinants of public sector efficiency

Since we have “leveled the playing field” across counties our PSE index can be a valuable tool in examining various theoretical relationships in the public choice literature. In this section, we use the implications of this literature on the political determinants of public sector efficiency so as to examine the empirical validity of our index. Note that the previous empirical literature has relied on local government data, and therefore the analysis below can

¹³It is worth observing that PSE scores sometimes exhibit notable variation across different spending accounts within the same country. At first sight, this may appear as a puzzling result. However, these findings can be attributed to political economy forces. For example interest groups involved in the production of each different public good (e.g., infrastructure and health) may have different political power. To the extent that groups with greater political influence receive more favorable treatment on the part of the government (see e.g., Grossman and Helpman 2001), there will be different levels of budgetary slack across different spending accounts.

¹⁴A point worth noting is that the Spearman correlation of the ranking of countries produced by our Stage 1 results (when the target is economic performance) with the public sector performance indicator ranking reported in Afonso et al. (2005) is 0.846.

be thought as a test of whether the relationships examined hold when using country-level data.

Our analysis is based on the following Tobit model:¹⁵

$$\begin{aligned} Eff_{i,t}^* &= 0 & Eff_{i,t}^* &< 0 \\ Eff_{i,t}^* &= \lambda_t + \beta X_{i,t} + \varepsilon_{i,t} & \text{if } 0 \leq Eff_{i,t}^* \leq 1 & \\ Eff_{i,t}^* &= 1 & Eff_{i,t}^* &> 1 \end{aligned} \quad (4)$$

where $Eff_{i,t}$ is the efficiency indicator of country i at time t obtained from the three-stage analysis above, $X_{i,t}$ is a vector of explanatory political variables and $\varepsilon_{i,t}$ is the error term. To account for time-specific shocks we add dummy variables (λ_t) for each time period t . Again, our dataset consists of the 19 OECD countries for the period 1980–2000. Yet, in order to increase the size of our sample and thus the robustness of our results obtained from the estimation of (4), we re-estimate the PSE scores by using the methodology described in Sect. 2 and five-year data averages. Therefore, the maximum number of observations in our sample is 76 (note that a few observations are missing for some of the variables used).

As dependent variable, we employ two alternative overall efficiency indicators: (i) the average of the efficiency indices on Performance, Stability and General Public Services (denoted as Eff) and (ii) the average of all seven PSE indices (denoted as $Eff2$). In our view Eff rests on a stricter definition of governmental tasks whereas $Eff2$ captures PSE on a broader but more heterogeneous range of public sector activities.¹⁶

Our set of explanatory variables consists of political variables that previous research has identified as important determinants of public sector efficiency on the basis of analyzes of local government (see e.g., Ashworth et al. 2006; Geys 2006; Borge et al. 2008).¹⁷ Hence, we employ a measure of democratic participation (denoted as *Turnout*); a cabinet political ideology measure (*Ideology*); the share of excess seats in the parliament held by the government (*powergov*)¹⁸ and a dummy variable taking the value 1 if a coalition government is in power (*coalition*). Finally, following Barankay and Lockwood (2007), we add the variable *decentr* to account for the degree of fiscal decentralization. Table 3 contains an analytical description of the variables employed, as well as the data sources.¹⁹

In Table 4, we report results from the estimation of (4) when the dependent variable is Eff . More precisely, in column I, Eff is regressed on the core set of political variables described above. In the remaining columns of Table 4, we examine the sensitivity of our

¹⁵The most-often encountered approach to modeling the DEA scores against exogenous variables is Tobit regression, which is suitable when the dependent variables are either censored or corner solution outcomes, of which DEA scores falls within the second category (Wooldridge 2002; Hoff 2007). As DEA efficiency scores are continuous on the interval (0; 1], and takes on the value 1 with positive probability, it seems obvious to use a two-limit Tobit technique for modeling the scores as a function of the exogenous variables. Tobit has as such been adopted as the natural ‘choice’ for modeling DEA scores in second stage evaluations in a large number of studies (see e.g. Coelli et al. 2005; Bravo-Ureta et al. 2007). For sensitivity analysis we have also estimated a simple linear regression using OLS estimation and our results remain qualitatively intact.

¹⁶Our results remain qualitatively intact when we use as dependent variable each efficiency indicator separately. All results are available from the authors.

¹⁷Note that as we have accounted for the socio-economic characteristics in the second stage of our computation of the efficiency index, there is no reason to include a set of variables to account for differences in the socio-economic environment as in Hayes et al. (1998).

¹⁸This refers to the seats beyond the number needed for a simple parliamentary majority.

¹⁹In the Appendix we present the correlation matrix of the variables.

Table 3 Variable definitions and statistics

Variable	Definition	Source	Summary	Statistics
<i>Eff</i>	Average of the efficiency indices on general public services, stability and performance	Authors' calculations	Mean Min, Max Obs	0.75 0.32, 1 73
<i>Eff2</i>	Average of all efficiency indices	Authors' calculations	Mean Min, Max Obs	0.75 0.48, 1 84
turnout	The proportion of voters who actually voted in the presidential (United States and Switzerland)/parliamentary (rest) elections. When two elections were held in the same five year period, we used the average ^a	Institute of Democracy and Electoral Assistance (2009)	Mean Min, Max Obs	0.78 0.42, 0.96 84
ideology	Ideological position of the government on a political spectrum from 1 to 10 with higher values denoting more extreme right-wing governments	Mierau et al. (2007)	Mean Min, Max Obs	5.55 0.14, 8.43 84
powergov	Number of surplus seats (i.e., seats beyond the number needed for a simple parliamentary majority) of the government in parliament divided by the number of total seats	Mierau et al. (2007)	Mean Min, Max Obs	0.06 −0.15, 0.34 84

results first by introducing one explanatory variable at a time (columns II to V)²⁰ and then by using alternative measures of government ideology (denoted as *ideo_CM*), fiscal decentralization (denoted as *taxdecentr*) and voter participation (denoted as *vapturnout*) (columns VI to VIII).

As can be seen, *Turnout* bears a positive and highly significant coefficient (at all relevant levels of statistical significance) in all alternative specifications, highlighting the positive effect of increased democratic participation on public sector efficiency. This result is in line with previous findings of empirical studies (see e.g., Borge et al. 2008) and suggests that a higher degree of democratic participation gives the politicians incentives to implement policies that improve efficiency and benefit broader categories of the population, at the expense of policies aiming to satisfy special interest groups (e.g., public sector unions and other lobbies). The coefficient on *powergov* appears to be positive and significant at the 1% level in all five estimations. This implies that strong political leadership appears to have an advantage in imposing hard budget constraints (see e.g., Roubini and Sachs 1989) and in

²⁰We begin this exercise in column II, where we regress *Eff* only on *decentr*. We have verified that the sign of the coefficients and their obtained statistical significance remain the same in all possible regressions between *Eff* and each independent variable separately.

Table 3 (Continued)

Variable	Definition	Source	Summary	Statistics
coalition	Dummy variable taking the value 1 if a coalition cabinet, which includes ministers from two or more parties, is in power	Tavares (2004)	Mean Min, Max Obs	0.56 0, 1 84
decentr	Sub-national expenditures (% of total expenditures)	Government Finance Statistics. IMF (2002)	Mean Min, Max Obs	30.98 4.31, 57.87 73
taxdecentr	Sub-national revenues (% of total revenues)	Government Finance Statistics. IMF (2002)	Mean Min, Max Obs	22.70 14.17, 3.45 73
ideo_CM	Castles and Mair (1984) ideological position of the government Ideo_CM locates the cabinet ideology on a 0 to 10 political spectrum with higher values denoting more right-wing governments	George Tsebelis' veto players dataset	Mean Min, Max Obs	5.10 2.9, 7.8 59
vapturnout	Total votes divided by voting age population in the presidential (United States and Switzerland)/parliamentary (rest) elections. When two elections were held in the same five year period, we used the average	Institute of Democracy and Electoral Assistance (2009)	Mean Min, Max Obs	0.73 0, 0.94 84

^a The classification of countries in parliamentary and presidential democracies follows Persson and Tabellini (2003)

implementing efficiency-enhancing policies even when the latter are not in agreement with public sector unions' interests (e.g., providing incentives for increased performance).

The coefficient on *Ideology* is positive and significant at the 1% level in all of the estimations, indicating that right-wing governments appear to achieve increased public sector efficiency. Our findings are in accordance with the empirical studies of Borge et al. (2008) and Strøm (1995) who suggest that socialist governments are associated with larger public sectors and increased costs through higher public officials' wages and, thus, are expected to be less efficient. The coefficient on *coalition* is always negative and significant (again at the 1% level). A potential explanation for this result is that the likelihood of a policy change decreases as the number of "veto-players" (i.e., individuals or institutions whose agreement is necessary to change the status quo) increases (Tsebelis 1995). Therefore, since coalition cabinets are characterized by larger numbers of veto players relative to one-party majority governments, they present a stronger "status quo bias" to established policies that holds even when these policies are proven to be inefficient (see e.g., Alesina and Drazen 1991; Spolaore 1993).

Finally, *decentr* bears a positive and significant coefficient in all alternative specifications. This finding is in line with Barankay and Lockwood (2007) and Adam et al. (2008),

Table 4 Tobit regression analysis (dependent variable: *Eff*)

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)
decentr	0.007*** (4.268)	0.018*** (11.902)	0.008*** (4.386)	0.007*** (3.949)	0.007*** (4.486)	0.007*** (4.802)		0.005*** (2.705)	0.005*** (3.150)
taxdecentr							0.006*** (3.961)		
turnout	0.454*** (4.349)		0.584*** (7.065)	0.358*** (3.032)	0.356*** (3.325)		0.485*** (4.640)	0.511*** (4.968)	0.171 (0.579)
vapturnout						0.282** (2.110)			
ideology	0.036*** (2.900)			0.040*** (2.838)	0.037*** (2.605)	0.054*** (3.420)	0.043*** (3.495)		0.017 (1.204)
ideo_CM								0.037*** (2.858)	
powergov	0.767*** (3.897)				0.474** (2.387)	0.710*** (3.176)	0.701*** (3.429)	0.451** (2.443)	0.583*** (3.299)
coalition	-0.144*** (-3.349)					-0.120** (-2.423)	-0.129*** (-2.962)	-0.143*** (-3.127)	-0.144*** (-2.832)
socioeco									0.000 (0.002)
invprot									0.023 (1.017)
density									-0.000 (-0.324)
old									0.008 (1.217)
urban									-0.000 (-0.118)
capopen									0.002 (1.143)
Log-Likelihood	24.261	6.828	13.989	16.757	19.897	19.733	23.072	27.335	29.732
observations	66	66	66	66	66	66	66	53	66
R-sq	0.25	0.19	0.10	0.11	0.15	0.21	0.23	0.33	0.31

Notes: ***, **, * denote statistical significance at the 1, 5 and 10% level, respectively. Robust t-statistics are reported in parentheses. R-sq is computed as the square of the correlation coefficient between actual and predicted dependent variable. All models estimated with no constant but with time dummies

and highlights the beneficial effect of greater fiscal decentralization on public sector efficiency. This result can be attributed either to (i) increased electoral control that comes as a result of fiscal decentralization (see e.g., Hindriks and Lockwood 2005; Seabright 1996) or to (ii) the yardstick competition argument (see e.g., Besley and Case 1995).²¹

²¹ According to this theory citizens have an advantage in evaluating the performance of their policy makers when they are able to compare the policy choices of their own political representatives with the corresponding choices of neighbor regions' policy makers. Therefore, fiscal decentralization may raise PSE, since it provides citizens the chance to compare public services and taxes across jurisdictions and helps them to judge whether their government wastes resources through low human capital capacity or rent-seeking.

Table 5 Tobit regression analysis (dependent variable: *Eff2*)

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)
decentr	0.006*** (6.341)	0.015*** (12.016)	0.007*** (6.383)	0.006*** (6.024)	0.006*** (6.425)	0.007*** (7.437)		0.004*** (3.949)	0.005*** (5.638)
taxdecentr							0.005*** (6.095)		
turnout	0.519*** (7.868)		0.623*** (13.585)	0.480*** (7.544)	0.478*** (7.724)		0.541*** (8.366)	0.583*** (8.160)	0.246** (2.161)
vapturnout						0.371*** (3.646)			
ideology	0.023*** (3.129)			0.026*** (3.588)	0.025*** (3.242)	0.037*** (3.010)	0.029*** (3.905)		0.010 (1.071)
ideo_CM								0.021*** (3.055)	
powergov	0.244* (1.705)				0.144 (1.044)	0.175 (0.999)	0.172 (1.139)	0.102 (0.674)	0.048 (0.375)
coalition	-0.049* (-1.885)					-0.019 (-0.589)	-0.033 (-1.227)	-0.036 (-1.337)	-0.043 (-1.524)
socioeco									0.014 (0.995)
invprot									0.015 (1.090)
urban									0.001 (1.021)
Log-Likelihood	64.762	9.450	59.713	62.743	63.499	55.560	62.435	56.519	71.665
observations	73	73	73	73	73	73	73	57	73
R-sq	0.22	0.19	0.05	0.06	0.08	0.09	0.10	0.16	0.29

Notes: ***, **, * denote statistical significance at the 1, 5 and 10% level, respectively. Robust t-statistics are reported in parentheses. R-sq is computed as the square of the correlation coefficient between actual and predicted dependent variable. All models estimated with no constant but with time dummies

In the equation presented in the last column of Table 4 we include the variables used above to “level the playing field”. Even though this regression is miss-specified by construction, it represents a rough test of whether we have indeed “leveled the playing field”. Evidently, the variables used to correct the first-stage efficiency scores are all insignificant. On the other hand, the signs of the coefficients of the political variables do not change, although their statistical significance drops.²²

In Table 5 we report the results from the estimation of equation (4) when we employ *Eff2* instead of *Eff* as dependent variable. As before, in column I *Eff2* is regressed on the core set of political variables, whereas in columns II to VIII we examine the sensitivity of our results first by introducing one explanatory variable at a time (columns II to V) and

²²A similar exercise can be carried out by including country fixed effects in the model. However, in the present setting this is not possible as a fixed effects (conditional) Tobit model cannot be estimated parametrically. However, if we estimate the model with random effects the results remain similar to those reported in column I. These results are available on request.

Table 6 Jackknife outlier analysis

	Minimum		Maximum	
	Excl. country		Excl. country	
<i>Dep. variable: Eff</i>				
decentr	Switzerland	0.005*** (3.08)	Canada	0.008*** (5.61)
turnout	Spain	0.382*** (3.74)	Switzerland	0.567*** (5.10)
ideology	USA	0.028* (1.88)	Luxemburg	0.043*** (2.68)
powergov	Switzerland	0.476** (2.21)	Canada	0.906*** (4.59)
coalition	Canada	-0.192*** (-3.93)	Spain	-0.094** (-2.06)
<i>Dep. variable: Eff2</i>				
decentr	Switzerland	0.004*** (4.87)	Canada	0.007*** (7.19)
turnout	Spain	0.487*** (6.52)	Switzerland	0.598*** (8.01)
ideology	USA	0.019* (1.91)	Belgium	0.027*** (2.81)
powergov	Switzerland	0.035 (0.24)	Austria	0.386*** (3.08)
coalition	Canada	-0.076** (-2.31)	Denmark	-0.025 (-0.74)

Notes: ***, ** and * denote statistical significance at the 1, 5 and 10% level, respectively. t-statistics are reported in parentheses

then by using alternative control variables (columns VI to VIII). Finally, as in Table 4, column IX includes the variables used to “level the playing field”. As can be easily verified our main results remain qualitatively similar, with the exception of the variables *powergov* and *coalition*, which become marginally statistically insignificant although they retain their predicted signs.

A final sensitivity analysis is provided in Table 6, where in order to determine whether our results are sensitive to the exclusion of particular countries we performed a jackknife analysis (Efron and Tibshirani 1993). This method involves estimating the initial equation by excluding in each replication one cross sectional unit (country). In Table 6 we report the maximum and the minimum estimated coefficients, as well as the excluded countries that exert this extreme identified impact. Comparing these coefficients with the ones presented in column I in Tables 4 and 5 we can conclude that our results are robust to the exclusion of particular countries.

All in all, our analysis in this section demonstrates two interrelated merits of the new PSE index. In particular, for the first time the index allows testing well-established relationships at a national government level, while it proves successful in providing findings that are in line with comprehensive theoretical and empirical evidence of the literature.

5 Conclusions

Recent developments in data envelopment analysis allow the disentangling of efficiency scores that are due to managerial practices from those that are due to the macroeconomic environment or luck. In such a fashion, we obtained relative PSE scores across 19 developed OECD countries. The evidence we presented adds credence to the proposition that individual countries have much to gain by looking at the relative efficiency in the provision of public services of other countries. Yet, the exploration of the forces that shape PSE showed that individual country characteristics and the effect of mere luck are less important than governance in determining PSE. By additionally looking at the political determinants of PSE we identify the political and institutional sources of the observed differences, thus (i) inquiring into the usefulness of the indices constructed and (ii) relating the findings on the efficiency scores to the relevant public choice literature, which has been using local government data to examine the relationships in hand.

We feel that these results point to a new agenda for research on how government efficiency affects economic performance. This agenda may simply examine the interrelationship between the efficiency of public spending, growth and economic volatility, but may well have other repercussions as well. The implications of the estimated scores for bureaucratic efficiency and capacity, the heavily debated link between democratic institutions and the public sector and the potential benefits of political reform, are only a few of the possible ways in which the present analysis can be extended. The further progress of such theoretical and empirical studies may significantly improve our understanding of the challenges faced in governing the public sectors as well as the monitoring of economic development.

Acknowledgements We are indebted to Harold Fried, Knox Lovell, William Shughart II, Efthymios Tsionas and three anonymous referees for valuable suggestions and discussions. We also thank Konstantinos Angelopoulos, Sophocles Brissimis, Vangelis Dioikitopoulos, George Economides, Thomas Moutos, Nikos Papanikolaou, Apostolis Philippopoulos and Nikos Tsakiris. Any remaining errors are ours.

Appendix: Correlation Matrix

	decentr	taxdecentr	turnout	vapturnout	ideology	ideo_CM	powergov	coalition
decentr	1							
taxdecentr	0.930	1						
turnout	-0.001	0.016	1					
vapturnout	-0.094	-0.092	0.558	1				
ideology	0.025	-0.032	-0.101	0.051	1			
ideo_CM	0.019	-0.050	-0.154	0.122	0.872	1		
powergov	-0.207	-0.101	-0.071	0.116	0.073	0.057	1	
coalition	-0.159	-0.166	0.213	0.183	0.073	-0.011	0.312	1

References

- Adam, A., Delis, M., & Kammas, P. (2008). *Fiscal decentralization and public sector efficiency: evidence from OECD countries*. CESifo, Working Paper no. 2364, August.
- Afonso, A., Schuknecht, L., & Tanzi, V. (2005). Public sector efficiency: an international comparison. *Public Choice*, 123, 321–347.
- Alesina, A., & Drazen, A. (1991). Why are stabilizations delayed? *American Economic Review*, 81, 1170–1188.
- Angelopoulos, K., Philippopoulos, A., & Tsionas, E. (2008). Does public sector efficiency matter? Revisiting the relation between fiscal size and economic growth in a world sample. *Public Choice*, 137, 245–278.
- Angelopoulos, K., & Philippopoulos, A. (2005). *The role of government in anti-social redistributive activities*. CESifo, Working Paper no. 1427, March.
- Ashworth, J., Geys, B., Heyndels, B., & Wille, F. (2006). *Political competition and local government performance: evidence from Flemish municipalities*. Paper presented at the Annual Meeting of the European Public Choice Society, Turku.
- Balaguer-Coll, M. T., Prior, D., & Tortosa-Ausina, E. (2007). On the determinants of local government performance: a two-stage nonparametric approach. *European Economic Review*, 51, 425–451.
- Barankay, I., & Lockwood, B. (2007). Decentralization and the productive efficiency of government: evidence from Swiss cantons. *Journal of Public Economics*, 91, 1197–1218.
- Besley, T., & Case, A. (1995). Incumbent behavior: vote-seeking, tax-setting, and yardsick competition. *American Economic Review*, 85, 25–45.
- Borge, L. E., Falch, T., & Tovmo, P. (2008). Public sector efficiency: the roles of political and budgetary institutions, fiscal capacity, and democratic participation. *Public Choice*, 136, 475–495.
- Bradford, D., Malt, R., & Oates, W. (1969). The rising cost of local public services, some evidence and reflections. *National Tax Journal*, 22, 185–202.
- Bravo-Ureta, B. E., Solfs, D., Lopez, V. H. M., Maripani, J. F., Thiam, A., & Rivas, T. (2007). Technical efficiency in farming—a meta-regression analysis. *Journal of Productivity Analysis*, 27, 37–72.
- Breton, A., & Wintrobe, A. (1975). The equilibrium size of a budget maximizing bureau. *Journal of Political Economy*, 83, 195–207.
- Castles, F., & Mair, P. (1984). Left–right political scales: some expert judgements. *European Journal of Political Research*, 12, 73–88.
- Coelli, T., Rao, D. S. P., O'Donnell, C. C., & Battese, G. E. (2005). *An introduction to efficiency and productivity analysis*. New York: Springer.
- Duncombe, W., Miner, J., & Ruggiero, J. (1997). Empirical evaluation of bureaucratic models of inefficiency. *Public Choice*, 93, 1–18.
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.
- Falch, T. (2001). Collective bargaining in the public sector and the role of budget determination. *European Journal of Political Economy*, 17, 75–99.
- Fried, H. O., Lovell, C. A. K., Schmidt, S. S., & Yaisawarng, S. (2002). Accounting for environmental effects and statistical noise in data envelopment analysis. *Journal of Productivity Analysis*, 17, 157–174.
- Geys, B. (2006). Looking across borders: a test of spatial interdependence using local government efficiency ratings. *Journal of Urban Economics*, 60, 443–462.
- Glass, J. C., McCallion, G., McKillop, D. G., & Stringer, K. (2006). A 'technically level playing-field' profit efficiency analysis of enforced competition between publicly funded institutions. *European Economic Review*, 50, 1601–1626.
- Grossman, G., & Helpman, E. (2001). *Special Interest politics*. Cambridge: MIT Press.
- Gupta, S., & Verhoeven, M. (2001). The efficiency of government expenditure, experiences from Africa. *Journal of Policy Modelling*, 23, 433–467.
- Hanushek, E. A., & Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American Economic Review*, 90, 1184–1208.
- Hayes, K. J., Razzolini, L., & Ross, L. (1998). Bureaucratic choice and nonoptimal provision of public goods: theory and evidence. *Public Choice*, 94, 1–20.
- Hindriks, J., & Lockwood, B. (2005). *Decentralization and electoral accountability: incentives, separation, and voter welfare*. Warwick, Working Papers in Economics, no. 717.
- Hoff, A. (2007). Second stage DEA: comparison of approaches for modeling the DEA score. *European Journal of Operational Research*, 181, 425–435.
- IMF (2002). *Government Finance Statistics, CD-ROM*. Washington DC, USA.
- Institute of Democracy and Electoral Assistance (2009). IDEA Voter Turnout dataset.
- International Country Risk Guide (2006). IRIS Center, University of Maryland.
- Kumbhakar, S., & Tsionas, E. (2005). The joint measurement of technical and allocative inefficiencies: an application of Bayesian inference in nonlinear random-effects models. *Journal of the American Statistical Association*, 100, 736–747.

- Kumbhakar, S., Park, B., Simar, L., & Tsionas, E. (2007). Nonparametric stochastic frontiers: a local maximum likelihood approach. *Journal of Econometrics*, *137*, 1–27.
- Mierau, J., Jong-A-Pin, R., & de Haan, J. (2007). Do political variables affect fiscal policy adjustment decisions? New empirical evidence. *Public Choice*, *133*, 297–319.
- Migué, J., & Bélanger, G. (1974). Toward a general theory of managerial discretion. *Public Choice*, *17*, 27–43.
- Moene, K. O. (1986). Types of bureaucratic interaction. *Journal of Public Economics*, *29*, 333–345.
- Niskanen, W. A. (1975). Bureaucrats and politicians. *The Journal of Law and Economics*, *18*, 617–643.
- OECD (2000). *Social expenditure database*. Paris.
- OECD (2005). *Telecommunications database*. Paris.
- OECD (2006). *Health data*. Paris.
- Peltzman, S. (1973). The Effect of government subsidies-in-kind on private expenditures: the case of higher education. *Journal of Political Economy*, *81*, 1–27.
- Persson, T., & Tabellini, G. (2003). *The economic effects of constitutions*. Cambridge: MIT Press.
- Quinn, D. (1997). The correlates of change in international financial regulation. *American Political Science Review*, *91*, 531–551.
- Roubini, N., & Sachs, J. (1989). Political and economic determinants of budget deficits in the industrial democracies. *European Economic Review*, *35*, 903–933.
- Seabright, P. (1996). Accountability and decentralization in government: an incomplete contracts model. *European Economic Review*, *40*, 61–91.
- Sijpe, N., & Rayp, G. (2007). Measuring and explaining government efficiency in developing countries. *Journal of Development Studies*, *43*, 360–381.
- Simar, L., & Wilson, P. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, *136*, 31–64.
- Spolaore, E. (1993). Policy making systems and economic efficiency, coalition governments vs majority government. Harvard University. Unpublished paper.
- Strøm, B. (1995). Envy, fairness and political influence in local government wage determination: evidence from Norway. *Economica*, *62*, 382–409.
- Tanzi, V., & Schuknecht, L. (1998). Can small governments secure economic and social well-being. In Grubel, H. (Ed.), *How to spend the fiscal dividend: what is the optimal size of government?* Vancouver: Fraser Institute.
- Tanzi, V., & Schuknecht, L. (2000). *Public spending in the 20th century: global perspective*. Cambridge: Cambridge University Press.
- Tavares, J. (2004). Does right or left matter? Cabinets, credibility and fiscal adjustments. *Journal of Public Economics*, *88*, 2447–2468.
- Tsebelis, G. (1995). Decision making in political systems: veto players in presidentialism, parliamentarism, multicameralism and multipartyism. *British Journal of Political Science*, *25*, 289–325.
- United Nations (2000). Department of Economic and Social Affairs, Statistics Division, Statistical Papers, Series M, No. 84, New York.
- Williamson, O. (1964). *The economics of discretionary behavior*. Englewood Cliffs: Prentice-Hall.
- Wintrobe, R. (1997). Modern bureaucratic theory. In Mueller, D. (Ed.), *Perspectives in public choice: a handbook*. New York: Cambridge University Press.
- Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge: MIT Press.
- World Bank (2004). *World Bank development indicators, CD-ROM*. Washington DC, USA.
- World Institute for Development Economics Research of the United Nations University (UNUWIDER) (2007). *World income inequality database*.
- Wyckoff, G. (1990). Bureaucracy, inefficiency, and time. *Public Choice*, *67*, 169–179.
- Zhu, J. (2003). *Quantitative models for performance evaluation and benchmarking*. New York: Springer.