Forthcoming in *The European Economic Review*

Growth Effects of Government Expenditure and Taxation in Rich Countries*

by

Stefan Fölster[1] & Magnus Henrekson[2‡]

20 June 2000

Abstract: A number of cross-country comparisons do not find a robust negative relationship between government size and economic growth. In part this may reflect the prediction in economic theory that a negative relationship should exist primarily for rich countries with large public sectors. In this paper an econometric panel study is conducted on a sample of rich countries covering the 1970–95 period. Extended extreme bounds analyses are reported based on a regression model that tackles a number of econometric issues. Our general finding is that the more econometric problems are addressed, the more robust the relationship between government size and economic growth appears. Our most complete specifications are robust even according to the stringent extreme bounds criterion.

JEL Classification: E62, H20, H50, O23, O40.

Keywords: Economic growth, Extreme bounds analysis, Fiscal Policy, Government expenditure, Public sector, Taxation, Cross-country regressions, Panel regressions, Robustness test.

[1]HUI, The Swedish Research Institute of Trade

S-103 29 Stockholm
SWEDEN
Phone: +46-8-762 72 88
Fax: +46-8-679 76 06
e-mail: Stefan.Folster@hui.se

‡Corresponding author
[2]Department of Economics
Stockholm School of Economics
Box 6501
S-113 83 Stockholm
SWEDEN
Phone: +46-8-736 92 02
Fax: +46-8-31 32 07
e-mail: Magnus.Henrekson@hhs.se

---

* We are grateful for excellent research assistance from Per Thulin and for useful comments and suggestions from Ulf Jakobsson, Assar Lindbeck, Erik Mellander, Joakim Persson and two anonymous referees. Needless to say, the usual caveats apply.

1. Introduction

The explosive development in the theory of endogenous growth has stimulated a great deal of empirical work on the determinants of economic growth. In particular, the influential work by Barro (1991), using a data set covering a large cross-section of both rich and poor countries, appeared to present strong empirical evidence favoring the view that a large public sector is growth-impeding. This result has been confirmed in some subsequent studies (e.g., Engen and Skinner, 1992; Grier, 1997; Hansson and Henrekson, 1994; de la Fuente, 1997) but has been challenged in others. For example, Mendoza, Milesi-Ferreti and Asea (1997) and Easterly and Rebelo (1993) find no discernible relation between government spending and growth. An explanation for the diversity of conclusions is illustrated by the extreme bounds analysis (EBA) that Levine and Renelt (1992) report. They analyze a large number of regressions with different combinations of conditioning variables. The negative partial correlation between government size and economic growth does not appear to be robust for some combinations of conditioning variables.[1]

While it may be that theory does not give much guidance as to the ultimate effect of public expenditure on growth, it does give some guidance regarding how empirical studies should be specified. For example, mainstream theory – such as in Barro (1990) and Slemrod (1995) – predicts that we should only expect to find a negative effect in countries where the size of the government sector exceeds a certain threshold.[2] With few exceptions, however, we only observe very large public sectors in rich countries.

A closely related rationale for restricting the empirical analysis to a sample of rich countries is also stressed by Slemrod (1995). It is well known that the scope of government tends to increase with the level of income. This tendency is commonly called Wagner's Law, and is often said to imply that the income elasticity of demand for government is larger than unity. But this relationship is weakened at the highest levels of income.[3] Moreover, one should keep in mind that in the theoretical models, tax rates cause the detrimental growth effects, whereas in the empirical work tax rates are proxied by tax revenues. Since the tax compliance ratio increases with the level of development, tax revenue is a much better proxy for tax rates in rich countries than in poor countries (Easterly, 1995).

The composition of public expenditure also differs between rich and poor countries. The

---

[1] A similar agnostic conclusion is reached in three recent review articles: Slemrod (1995), Atkinson (1995) and Agell, Lindh and Ohlsson (1997).

[2] See also Tanzi and Zee (1997).

[3] Easterly and Rebelo (1993) show that there is a strong positive relationship between government size and per capita income both across a large sample of countries at a point in time (1985) and for a panel of 28 countries from 1870 to 1988. This relationship disappears only at the highest levels of income.

various programs that have been hypothesized in theoretical work to have positive growth effects – e.g., schooling, infrastructure and R&D subsidies – typically amount to less than one-fifth of public expenditure in OECD countries, while they typically amount to more than half of public spending in less developed countries. This means that 80 percent or more of public expenditure in OECD countries consists of expenditure that is not often claimed to have positive growth effects. Moreover, most of the variance in public expenditure between countries is explained by differences in public expenditure that has not been claimed to have positive growth effects. There is an extensive literature indicating that many public programs have negative effects on saving and capital accumulation, and create marginal effects in addition to those that emanate from the tax system.[4]

We conclude from this discussion that analyzing rich countries separately may add to our understanding of whether large public expenditure has negative growth effects. This conclusion is also consistent with the findings by Grier and Tullock (1989) who present evidence showing that data from the OECD and the rest of the world do not share a common set of coefficients and thus should not be pooled. A key question is still how to select rich countries. A common approach has been to use the sample of OECD countries. We also report regressions using the OECD sample. At the same time it is worth noting that this sample is not fully satisfactory. Countries are granted OECD membership partly based on good economic performance and high GDP levels, and partly based on other criteria, such as size, democracy and institutional structure. The effect of this selection could introduce a bias. For this reason we also run regressions using an objective income criterion to select a sample of rich countries. Moreover, we report extensive robustness tests including extreme bounds analyses based on a regression specification that addresses the issue of country selection as well as a number of other econometric issues. The robustness tests are extended by means of the extreme bounds analysis suggested by Sala-i-Martin (1997). The basic idea of this extension is to examine the entire distribution of coefficient estimates rather than using an absolute criterion of robustness.

In section 2 we present our basic empirical analysis, where we address a number of econometric issues. Section 3 contains an examination of potential business cycle effects and the results from estimations including additional rich countries. Section 4 reports the results of our robustness tests. Section 5 concludes.

2. Our Basic Test

Many scholars (e.g., Plosser, 1993) have found a negative bivariate relationship between the

---

[4] See Fölster and Henrekson (1999) for evidence and references corroborating the assertions in this paragraph.

rate of growth and the measure of government size. It is well known that the inclusion of particular control variables can wipe out this bivariate relationship (e.g., Easterly and Rebelo, 1993). Thus, one needs to carefully consider what variables to include in a growth regression. Sala-i-Martin (1997) collected around 60 variables, which had been found to be significant in at least one growth regression.

Since growth theory suggests that growth is driven by accumulation, growth regressions usually include measures of the growth of the factors of production: physical capital, labor and human capital. In every regression below we therefore include gross investment as a share of GDP (*INV*), the growth rate of the labor force (*DLAB*), and the growth of human capital (*DHUM*), measured as the growth rate of the average years of schooling.[5] Given the overwhelming support for (conditional) convergence in the empirical growth literature (Barro and Sala-i-Martin, 1995), we also include initial income (Y_0) among the regressors that always appear in the regressions.[6] Two measures of government size will be used: total taxes as a share of GDP (*TAX*) and total government expenditure as a share of GDP (*GEXP*).

When this basic set-up is applied in a pure cross-section framework on a sample of OECD countries, we detect no effect of the government size variables (*Table 1*). As in most cross-section studies initial income Y_0 is taken as the beginning-of-period value, while government expenditure and taxes are measured as averages over the observed period.[7]

The argument for cross-section studies over long time spans has been that less interesting short-to-medium term effects, such as business cycle effects, are thereby eliminated. We will return to this issue in more detail. First, however, a number of problems with cross-section studies using long time spans need to be discussed.

The most important may be a potentially severe simultaneity problem. The cross-country regressions are usually based on average values of government spending and growth over long time periods, typically twenty-year periods or longer. Over long time spans the level of government spending is likely to be influenced by demographics, in particular an increasing share of elderly. Just to give an example, for OECD countries Agell *et al*. (1997) report a

---

[5] This choice of explanatory variables can be derived from an aggregate production function. Alternatively, one could use the specification derived from the Solow growth model, which implies that the level rather than the change in human capital enters the regression.

[6] In regressions that include both rich and poor countries it is common to use the log of Y_0 as the measure of initial income. Since we focus on rich countries only in this paper, the differences in Y_0 are comparatively small, and it turns out that it makes little difference whether initial income is logged. Therefore, we report only regressions where Y_0 is not expressed in logarithms.

[7] Since government expenditure and taxes change so much over a time span of several decades it would hardly be meaningful to regress beginning-of-period values of, say, taxes, over income growth. For initial income the argument is different, since subsequent changes in income are captured by the dependent variable, GDP-growth.

correlation of 0.72 between the tax ratio and the percentage share of population aged 65 or above. At the same time the share of elderly is closely correlated with GDP. Higher incomes increase expected life spans. This means that if GDP increases faster over the twenty-year period, growth will be higher, but the share of elderly also increases and government spending rises. Thus, errors in the growth variable will affect GDP, demographics and taxes or government spending. As a result the independent variable, taxes or government spending as a share of GDP, is correlated with the error term in the growth regression. This bias could easily give rise to positive coefficient estimates for the effect of taxes on GDP growth, such as those indicated by Table 1.

A second problem is that cross-section studies using long observation periods give rise to an endogenous selection of tax policy. Countries that do raise taxes and experience lower growth during the observation period are more likely to change policy stance and reduce taxes, such as Ireland did during the 1980's. In contrast, countries that raise taxes without experiencing a negative growth effect (such as Norway, which discovered oil along the way) are more likely to continue having high taxes. This means that cross-section studies over long time spans may fail to capture growth effects of fiscal policy due to such endogenous policy determination.

Table 1 Cross-country OLS Regressions for the Growth Effect of Public Sector Size in 22 or 23 OECD Countries 1970–95.

Explanatory variables	*Expenditure*		*Taxation*	
Constant	0.0086	(0.50)	0.012	(0.76)
GEXP	0.018	(0.90)		
TAX			0.015	(0.64)
Y_0	–0.015**	(–2.90)	–0.014*	(–2.59)
INV	0.082(*)	(1.84)	0.074	(1.63)
DLAB	0.227	(0.84)	0.154	(0.53)
DHUM	–0.065	(–0.27)	–0.0056	(–0.02)
No. of obs.		22		23
Adjusted R^2		0.30		0.25

Note: (*), * and ** denote significance at the 10, 5 and 1% levels, respectively. *t*-statistics in parentheses. All variables are measured as averages over the entire period. Turkey and Mexico are excluded from the sample, since they cannot be considered to be rich countries. A complete list of the included OECD countries is provided in the Appendix.

A third, related problem with cross-section studies over long time spans is that they may be inefficient since they discard all information on within-country variation.[8] Exploiting within-

[8] To our knowledge this point was first made by Grier and Tullock (1989).

country variation is particularly interesting, since the dispersion across OECD countries of total government outlays as a share of GDP has increased substantially since 1960. In some countries, such as Sweden and Portugal, government size has continued to increase up to the present, while in others, such as the U.K. and the Netherlands, there has been little change or even a decline over the last 15 years. As a result, the expenditure ratio as of the mid 1990s varied between roughly 65 percent for Sweden and some 35 percent for the U.S. and Japan. An additional reason for using a panel data approach, as stressed by Islam (1995), lies in its ability to allow for differences in the aggregate production function across countries.

While both the simultaneity effect and the use of within-country variation are arguments in favor of panel regressions with shorter time spans, there are also risks. When the period of observation is short, it is much less likely that the error in the growth regression will affect life expectancy and government spending in the same period. But in a panel regression with annual data it would be very important to estimate the proper lag structure. This would mean that several years of lagged government expenditure would have to be included in the regression, giving rise to multicollinearity. In addition, if the lag length varies over time and between countries, the lagged effects may not be captured properly anyway.

As a compromise we focus mainly on combined cross-section time-series regressions using five-year periods. In keeping with most of the previous literature a control variable such as initial income is defined as GDP per capita at the beginning of the five-year period, while the explanatory variables government expenditure and taxes are averages over the respective five-year period. Over the course of five years a good deal of the lagged effects are captured.

This still leaves another important risk with panel regressions, namely the occurrence of short term covariation such as business cycle correlations that may, for example, give rise to increasing public expenditure for unemployment when the growth rate falls. We will address the issue of business cycle covariation carefully in the next section, where we show instrument variable regressions and regressions with panels using annual observations.

Not all immediate effects need to be related to the business cycle however. Some long-term effects of changed fiscal policy may materialize quite quickly. For example, if a government announces expenditure or tax hikes there can be a rapid reaction in savings and investment and thus GDP growth, which might occur simultaneously or even before the change in fiscal policy is actually implemented. For this reason it would not be satisfactory to use only lagged fiscal variables as proxies, since these would fail to capture such immediate effects.

Before proceeding to the results of the panel regressions, and further considering the issue of business cycle covariation, a number of other econometric issues should be mentioned. Heteroscedasticity most often appears in a form where the error term is correlated with one of the independent variables or with the dependent variable. This does not seem to be a problem in our data, however.[9] We do have a potential problem with heteroscedasticity between countries.[10] A standard solution to this problem, which we apply, is to use a weighted least squares procedure that weights countries inversely proportional to the standard deviation of the error term.[11] An additional problem that we attempt to come to grips with is that panel data estimations may yield biased coefficient estimates when lagged dependent variables are included. In our case, initial income is a regressor which is also present in the dependent variable, the rate of growth per capita. We have therefore reestimated our regressions using the corrected least squares dummy variable estimator suggested by Kiviet (1995).[12] This procedure led, however, to quite similar estimates in all specifications, so they are not reported here.

In *Table 2* the regression results are presented. Fixed country effects and fixed period effects are here taken into account by including dummies. The inclusion of period dummies prevents us from picking up a spurious correlation that could arise because most countries have experienced a reduction in the growth rate in the 1970s and 1980s. Country dummies take account of country-specific effects, such as culture and social norms. As shown by Islam (1995), neglecting unobserved differences in the aggregate production function between countries induces an omitted variables problem.

As Table 2 shows, the panel estimation yields a highly significant negative growth effect for *GEXP*. The tax variable is negative, but not quite significant at the 10% level in a two-tail test (in the weighted regression the significance level is 10.2%). The estimated effects of *GEXP* are also somewhat larger, implying that an increase in the expenditure ratio by 10 percent of

---

[9] Not surprisingly, therefore, common corrections for heteroscedasticity along these dimensions, such as the White (1980) and Newey-West (1987) corrections hardly change the results. A White test for heteroscedasticity yields an *F*-statistic of 1.07 in the tax equation (and 1.04 in the government spending equation), implying a probability of 0.41. Moreover, testing the relation between the error term and the independent and dependent variables one at a time does not yield a significant relationship in any instance.

[10] The likelihood ratio test for groupwise heteroscedasticity suggested by Fomby, Hill and Johnson (1984) yields a χ^2 of 139.7, which is significant at the 1% level.

[11] It is worth noting that the weighted least squares procedure we use is not biased even if the standard deviation of the error is correlated with the explanatory variable. To see this intuitively one need only recall that the central idea in the correction of traditional heteroscedasticity – where the error term is correlated with an explanatory variable – is precisely to weight observations by the values of the independent variable such as the tax rate. In our case, two points may be made. First, the correlation between the standard deviation of the error term and the tax rate is extremely weak and by no measure significant. Second, even if there was such a correlation, the weighted least squares procedure would not be biased.

[12] Kiviet (1995) derives a formula for the small sample bias of the within-group, or least squares dummy variable, estimator for the coefficients of a first-order dynamic panel data model, and shows that correcting the estimator with the calculated bias gives more robust results than various GMM or IV estimators.

GDP is associated with an annual growth rate that is 0.7–0.8 percentage points lower. A straightforward explanation of this difference may be that a budget deficit has growth effects similar to that of taxation, which implies that government expenditure is a better measure of current and future taxation.

Table 2 Panel Regressions for the Growth Effect of Taxation and Public Consumption in 23 OECD Countries and Public Expenditure in 22 OECD Countries 1970–95 (including country and period dummies).

Explanatory variables	*OLS regression*	*Weighted regression*	*OLS regression*	*Weighted regression*	*OLS regression*	*Weighted regression*
TAX	–0.046	–0.055				
	(–0.96)	(–1.65)				
GEXP			–0.074*	–0.088**		
			(–2.49)	(–4.32)		
GCONS					–0.25**	–0.28**
					(–3.36)	(–5.23)
Y_0	–0.084**	–0.085**	–0.099**	–0.087**	–0.081**	–0.085**
	(–4.46)	(–5.79)	(–5.35)	(–6.28)	(–4.55)	(–6.46)
INV	–0.0088	0.016	–0.034	–0.020	0.005	–0.037
	(–0.18)	(0.37)	(–0.77)	(–0.45)	(0.10)	(–0.88)
DLAB	0.266(*)	0.190(*)	0.089	0.062	0.20	0.20*
	(1.96)	(1.96)	(0.68)	(0.75)	(1.56)	(2.25)
DHUM	–0.0097	0.025	0.014	–0.0022	–0.011	0.027
	(–0.13)	(0.63)	(0.20)	(–0.06)	(–0.15)	(0.69)
No. of obs.	115	115	109	109	115	115
Adjusted R^2	0.46	0.82	0.56	0.88	0.52	0.83

Note: (*), * and ** denote significance at the 10, 5 and 1% levels, respectively. *t*-statistics in parentheses. The regressors are measured as averages for the respective subperiods, except for Y_0 which measures the income level in the initial year of each subperiod.

Table 2 also shows the regular OLS regressions, leaving heteroscedasticity uncorrected. As expected, the standard errors of the estimates in these regressions are larger. Nevertheless, *GEXP* has a significant negative coefficient, while the coefficient for taxes is negative, but insignificant. In the Appendix we show similar regressions including all the standard regressors that are meaningful and available in a panel of rich countries. Both the tax and government expenditure coefficients are significant in the weighted regression specification. Other conditioning variables such as investment and human capital are not significant in Table 2. This is mainly a consequence of introducing country dummies into the regression equation, while the coefficients on the country dummies are quite significant.[13] As noted above, neglecting differences across countries in the aggregate production function may induce an omitted variables bias. But these unobserved country differences arguably imply cross-country differences in investment opportunities and optimal capital-labor ratios. It is therefore not all that surprising to find that investment loses significance once country dummies are included. It is also the case that the statistical significance of the estimated effect of *TAX* and *GEXP* increases when *INV* is not included among the regressors. This is to be expected, because in this case the indirect effect of taxes on growth via investment is also captured by the government size variable.

The regression specifications in the first four columns of Table 2 are subjected to the robustness tests we report in section 4. Many studies use government consumption rather than government expenditure as the explanatory variable of interest. For example, Grier (1997) finds strong negative effects of government consumption on growth. As shown in Table 2, government consumption (*GCONS*) is significantly negatively related to growth at the 1% level. Yet, in theory, the tax used to finance non-actuarial transfers should have the same growth effect as a tax that finances government consumption. Moreover, public financing of education, health care and other social policies are classified as government consumption in some countries and as transfers in others. For these reasons, we focus on total government expenditure rather than consumption in what follows.

3. Checking for Business Cycle Effects and an Extension of the Sample

As noted above, the use of panel data itself mitigates long-run simultaneity problems that arise because, among other things, the demographic structure and political preferences change with rising income. But, shortening the period of observation may increase the risk of picking up a correlation driven by business cycle effects. However, this cyclical covariation should be at least partially removed by using five-year periods and by controlling for period effects using period dummies.

A typical business cycle correlation might imply that government expenditure increases (e.g. in line with unemployment costs) when growth falls, while tax revenue would typically decrease. Further, an expansive fiscal policy can stimulate demand and thus growth. To check the importance of these correlations, we also entered control variables that vary with the business cycle such as unemployment. In *Table 3* we present results when *UNEMPL* is added among the regressors. This hardly changes the results for *GEXP*, although the effect of the tax

[13] All country dummies are significantly different from zero, and eight country dummies are significantly different from the average dummy coefficient (using the specification of the first column of Table 2).

variable is somewhat weakened. Similar results are obtained using the change rather than the level of unemployment as an explanatory variable (not shown). The same is also true for regressions using beginning-of-period values for the fiscal variables (not shown).

In order to further examine the possibility of business-cycle induced simultaneity, we estimate various specifications using first differences, instruments and potential output. Reporting all these would lead us too far off track. A typical result is shown in columns five and six of Table 3 for a specification with a first differences, two-stage weighted least squares regression using instruments for the tax and government variables.[14] The coefficients on the tax and expenditure variables are still significant.[15]

Table 3 Panel Regressions for the Growth Effect of Public Sector Size in 23/22 OECD Countries 1970–95 Accounting for Business Cycle Effects (including country and period dummies).

Explanatory variables	*OLS regression*	*Weighted regression*	*OLS regression*	*Weighted regression*	*First differences 2SLS*	*First differences 2SLS*
TAX	–0.021	–0.045			–0.12**	
	(–0.39)	(–1.17)			(–3.40)	
GEXP			–0.073*	–0.093**		–0.11**
			(–2.10)	(–4.15)		(–5.59)
Y_0	–0.088**	–0.083**	–0.099**	–0.087**	–0.027*	–0.037**
	(–4.62)	(–5.46)	(–5.30)	(–6.17)	(–2.00)	(–3.12)
INV	–0.025	0.015	–0.035	–0.013	0.071*	0.058*
	(–0.49)	(0.32)	(–0.77)	(–0.29)	(2.83)	(2.88)
DLAB	0.287*	0.194$^{(*)}$	0.092	0.051	0.079	0.053
	(2.10)	(1.91)	(0.68)	(0.61)	(1.47)	(1.17)
DHUM	–0.013	0.027	0.014	–0.0067	0.040	0.005
	(–0.17)	(0.63)	(0.20)	(–0.19)	(1.05)	(0.17)
UNEMPL	–0.071	–0.032	–0.0047	0.025		
	(–1.21)	(–0.73)	(–0.08)	(0.60)		
No. of obs.	115	115	109	109	65	64
Adjusted R^2	0.46	0.81	0.55	0.88	0.61	0.86

Note: $^{(*)}$, * and ** denote significance at the 10, 5 and 1% levels, respectively. *t*-statistics in parentheses. *UNEMPL* is defined as the average for each subperiod.

[14] The use of first differences is often considered to be a more effective way of correcting for fixed country effects. The first difference of the tax and public expenditure variables are instrumented by the lagged levels of taxes and public expenditure, respectively, fixed country effects, and levels and first differences of the population and initial GDP variables. As the dependent variable we use growth of potential GDP per capita (from OECD, *Economic Outlook*).

[15] They become insignificant when the second lags of taxes and government expenditure are used. This is hardly surprising given that the number of degrees of freedom becomes very small.

At this point one might reconsider the issue of using five period averages as observations. Presumably, if business cycle covariation were an important explanation for the link between government expenditure and growth, one would expect regressions using one-year periods to yield even more significant and larger coefficient estimates than our five-year periods.[16] But, using one-year periods does not yield a stronger or more significant correlation than the five-year estimates reported in Table 2. For example, the estimated coefficient for *TAX* in an OLS-regression corresponding to the first column in Table 2 is –0.021 (t = –0.45). And the point estimate for *GEXP* corresponding to column 3 in Table 2 is actually positive, 0.025 (t = 0.78).

In sum, there are serious issues of endogeneity both in cross-section studies using long periods of observation and in panel studies using short periods of observation. Our compromise of using five-year periods hardly settles this issue for good, but we hope to have shown that the results are not based on very short-term covariation over the business cycle.

A further methodological issue of great potential relevance is the selection of a sample of rich countries. OECD countries are themselves selected among high-income countries, in part for their good growth performance, and in part according to other criteria, such as the existence of democracy. It would therefore be natural to analyze our question using a sample of rich countries which is not restricted by OECD membership. To examine this issue we extend our sample to all non-OPEC/non-tax haven countries that have a PPP-adjusted GNP per capita in 1995, the final year of our inquiry, comparable to the OECD countries. These countries have been identified from World Bank (1997, table 1.1) and the most recent version of *Penn World Tables*. The countries thus included are Hong Kong, Singapore, Israel, Mauritius, Korea and Taiwan. The poorest of these countries is Korea with a PPP-adjusted GDP per capita of USD 260 below the level in Greece. Next in line after Korea in terms of income per capita in 1995 is Chile, USD 1,930 below the Korean level. This large gap between these two countries and a formal cluster analysis confirm that the 23 richest OECD countries plus the additional six

[16] A number of studies, e.g., Blanchard & Perotti (1999), find a positive impact of fiscal policy on output using quarterly data.

countries listed above constitute a reasonably well defined group of rich countries.[17] Some interesting features of the additional countries plus Chile are displayed in *Table 4*.

Table 4 Rich non-OECD Countries Excluding OPEC Countries and Tax Havens. Countries with a PPP-adjusted GNP per Capita above the OECD Average Minus Two Standard Deviations in 1995, Average Growth Rate 1980–95 (%) and Government Expenditure Share in 1995 (%).

Country	GNP per capita 1995 PPP dollars	Growth rate of GDP per capita 1980–95	Government expenditure as a share of GDP in 1995
Chile	9,520	3.19	19.2
Hong Kong	22,950	4.75	14.5
Israel	16,490	2.16	44.7
Korea	11,450	7.32	17.7
Mauritius	13,210	4.65	23.3
Singapore	22,770	5.64	14.4
Taiwan	13,490	6.21	30.0

Note: Tax havens excluded are Bahrain, Barbados, the Bahamas and St. Kitts and Nevis. GNP per capita in Taiwan in 1995 has been estimated – on the basis of *Penn World Tables* version 5.6 and World Bank (1997) – to be 50% of the U.S. level in 1995.
Source: See Appendix.

Table 5 Panel Regressions for the Growth Effect of Public Sector Size in 29/28 Rich Countries 1970–95 (including country and period dummies).

Explanatory variables	*OLS regression*	*Weighted regression*	*OLS regression*	*Weighted regression*
TAX	–0.112**	–0.107**		
	(–2.64)	(–3.76)		
GEXP			–0.099**	–0.106**
			(–4.27)	(–5.66)
Y_0	–0.043**	–0.040**	–0.050**	–0.042**
	(–3.00)	(–3.51)	(–3.55)	(–4.00)
INV	0.060(*)	0.076*	0.032	0.019
	(1.68)	(2.30)	(0.96)	(0.58)
DLAB	0.172	0.172(*)	–0.013	0.023
	(1.22)	(1.70)	(–0.10)	(0.26)
DHUM	0.080	0.071(*)	0.101	0.043
	(1.14)	(1.86)	(1.52)	(1.21)
No. of obs.	145	145	139	139
Adjusted R^2	0.70	0.79	0.74	0.82

Note: (*), * and ** denote significance at the 10, 5 and 1% levels, respectively. *t*-statistics in parentheses.

[17] The formal cluster analysis also shows that it is reasonable to exclude the two poorest OECD countries, Turkey and Mexico, from the OECD regressions.

Regressions for this extended sample of rich countries presented in *Table 5* show an overwhelmingly strong relation between *TAX* or *GEXP* on the one hand, and growth on the other hand. OLS or weighted regressions notwithstanding, the estimated effect is highly significant. Quantitatively, the effect is estimated to be somewhat larger than before; a 10 percentage points increase in public sector size is associated with a reduction of the growth rate of roughly one percentage point.

4. Robustness Tests

The purpose of this section is to investigate the robustness of the regression results presented above. The point of departure for our robustness tests is Leamer's (1983) extreme bounds analysis (EBA), and Levine and Renelt's (1992) empirical application of this test. Adopted to our context, this implies estimation of regressions of the form:

$$\lambda = a_j + b_{yj} y + b_{zj} z + b_{xj} x_j + \varepsilon \tag{1}$$

where y is a vector of fixed variables that always appear in the regressions (Y_0, *INV*, *DLAB* and *DHUM*), z denotes the variable of interest (*TAX* or *GEXP*) and x_j is a vector of three variables taken from the pool X of additional plausible control variables. The regression model has to be estimated for the M possible combinations of $x_j \in X$. For each model j one estimates b_{zj} and the corresponding standard deviation σ_{zj}. The lower extreme bound is defined as the lowest value of $b_{zj} - 2\sigma_{zj}$ and the upper extreme bound is defined to be the largest value of $b_{zj} + 2\sigma_{zj}$. If the lower extreme bound is negative and the upper extreme bound is positive, the variable is considered not to be robust.

Sala-i-Martin (1994) has two important objections against the Levine–Renelt methodology. First, he notes that there is a "reverse data-mining" problem. The control variables are samples drawn with some error from the true population. Therefore, if one keeps trying different combinations of control variables one is almost guaranteed to find one or a combination of several control variables for which the error is such that it renders the coefficient of interest insignificant or even causes it to change sign. "The implication is that the extreme-bounds test may be too strong" (p. 743). Second, Sala-i-Martin points out that Levine and Renelt in fact always find some group of policy variables that matter. The policy variables are so highly correlated that one often cannot distinguish between them, and the proxies used are always imperfect measures. Depending on the sample and the specific choice of explanatory variables, the data are likely to pick one variable or another because they are all close and imperfect indicators of the same phenomenon.

Sala-i-Martin (1997) has moved away from the EBA by looking at the whole distribution of the estimates of b_z. We adhere to that approach in this paper. More specifically, we will reestimate the regressions above with all possible triplets of conditioning variables. From this exercise we can (i) conduct the EBA and (ii) compute the share of all regressions that result in a statistically negative effect of the government size variable.[18]

Sala-i-Martin (1997) applied 59 control variables used in the literature. On the other hand he limits himself to robustness tests of a cross-section of countries including both rich and poor countries. Many of his variables are irrelevant for a sample of rich countries, e.g., the black market premium, the number of revolutions and coups and the degree of civil liberties, or they are not available for several time periods. Furthermore, many variables are constant over the sample period. In our case, these variables are implicitly captured by the country dummies. These considerations have limited the number of control variables we can use considerably. We have collected the following eleven control variables: *DEPPOP, EXP, FERT, IMP, INFL, OPEN, POP, SAV, TYR, URBAN, UNEMPL.* The mnemonic names are largely self-explanatory, but the interested reader is referred to the Appendix for full definitions of the variables. Eleven conditioning variables imply $\binom{11}{3} = 165$ possible combinations of $x_j \in X$.[19]

The results from the robustness test for weighted regressions on the OECD sample using 11 conditioning variables is presented in *Table 6*. We may first note that the *GEXP* coefficient is generally more robustly negative than the *TAX* coefficient; 73.8 percent of the *GEXP* estimates are negative and significant compared to 43.9 percent for the *TAX* estimates. The estimated effects are not robust with respect to the stringent EBA criterion.

Upon closer inspection one can detect a strong negative correlation between savings and government expenditure. This is not all that surprising since national saving actually includes government saving directly via an accounting identity, giving rise to multicollinearity between a conditioning variable and the variable of interest.

To check how important this multicollinearity problem is, we conduct two further sensitivity analyses. First, we exclude *SAV* from the set of conditioning variables. As reported in *Table 6a* the *GEXP* coefficient now becomes robust according to the Levine-Renelt EBA criterion

---

[18] Sala-i-Martin (1997) also suggests a method for computing the fraction of the cumulative distribution of b_z lying to the left of zero. This computation requires assumptions regarding the distribution of the estimator b_z and a choice of an appropriate weighting scheme. The statistical foundation of this procedure is yet unclear so we refrain from reporting these computations here, but they are available upon request from the authors.

[19] Since *OPEN* is a linear combination of *EXP* and *IMP* the regression containing these three variables together cannot be estimated. Thus, the total number of equations is reduced to 164.

as well. The robustness of the *TAX* coefficient is increased considerably. Now 59.7 percent of all estimates are negative and statistically significant at the 5% level. Second, in *Table 6b* we report the results from a full robustness test with eleven conditioning where *SAV* is replaced by private saving as a share of GDP (*PSAV*). The results in this case are stronger than in the case when *SAV* is just excluded: *GEXP* is still robust according to the stringent EBA criterion and more than 70 percent of the estimates of the *TAX* coefficients are negative and statistically significant.

Table 6b Robustness Tests for the OECD Sample with Eleven and Ten Conditioning Variables.

	Eleven Conditioning Variables		*Ten Conditioning Variables*	
	TAX	*GEXP*	*TAX*	*GEXP*
EBA lower bound	–0.202	–0.164	–0.202	–0.164
EBA upper bound	0.103	0.043	0.044	–0.028
% significant†	43.9	73.8	59.7	100.0

†The share of all regressions resulting in an estimate of b_{zj} that is negative and significant at the 5% level.

Table 6b Robustness Test for the OECD Sample when Private Saving Substitutes for Total Saving (as a share of GDP).

	TAX	*GEXP*
EBA lower bound	–0.243	–0.172
EBA upper bound	0.044	–0.023
% significant	70.7	100.0

Finally, *Table 7* reports analogous robustness tests for the sample of all rich countries excluding Taiwan (due to the extremely limited data availability). Since data for all countries were not available for *SAV* and *UNEMPL*, we only have nine conditioning variables. Both the *TAX* and *GEXP* coefficients are robustly negative according to the EBA criterion, which is the same as to say that all possible regressions yield negative and statistically significant estimates for the government size variables.

Table 7 Robustness Test for all Rich Countries with Nine Conditioning Variables.

Weights	*TAX*	*GEXP*
EBA lower bound	–0.199	–0.179
EBA upper bound	–0.010	–0.052
% significant†	100.0	100.0

†The share of all regressions resulting in an estimate of b_{zj} that is negative and significant at the 5% level.

In sum, the robustness tests seem to imply that there is a robust relation between high public expenditure and lower growth. Even the stringent EBA criterion is met for the OECD sample with ten conditioning variables. The robustness results are less clear-cut for the *TAX* variable. The EBA criterion is not satisfied for the OECD sample, although it is noteworthy that in the case when *SAV* is excluded from the set of conditioning variables 60 percent of the *TAX* coefficients are significant. For the extended sample of rich countries, the tax variable also satisfies the stringent EBA criterion.

5. Conclusion

Empirical studies of the relation between government size and economic growth have come to widely different conclusions. In this paper an econometric panel study on a sample of rich countries covering the 1970–95 period is conducted. A main motivation for our analysis is that tests of robustness such as extreme bounds analyses now are used routinely to examine various relationships. Often the basic regression used in these tests contains numerous, and unnecessary, econometric problems. Our contention is that extreme bounds analyses based on such regression specifications are highly doubtful. In general, it is hardly possible to solve all econometric problems. But it is informative to examine what happens to robustness tests such as the EBA when at least some of the econometric issues are addressed.

In the case of the relationship between public expenditure and economic growth it appears that exploiting within-country variation by means of panel regressions, correcting for heteroscedasticity between countries, and addressing the issue of country selection, in fact permits a more robust conclusion. The results point to a robust negative relationship between government expenditure and growth in rich countries. The size of the estimated coefficients imply that an increase of the expenditure ratio by 10 percentage points is associated with a decrease in the growth rate on the order of 0.7–0.8 percentage points. When the rich country sample is extended to non-OECD countries both government expenditure and taxation are found to be negatively associated with economic growth. These findings are robust even according to the stringent extreme bounds criterion.

Appendix: Data description and supplementary regressions

Average annual growth rates for the relevant variable *X* was computed as $\left(\frac{X_E}{X_B}\right)^{\frac{1}{E-B}} - 1$; *B* = beginning of period, *E* = end of period.

Data for government expenditure were missing for New Zealand, and for Luxembourg they were missing after 1986. As a result the *GEXP*-regressions contains one country less throughout, and likewise there are only four observations for Luxembourg.

There are no data available for *DHUM* and *TYR* for Luxembourg. Instead we have used the average for Belgium and the Netherlands. Several measures were taken to check whether our results could be misleading as a result of the use of this proxy. The exclusion of Luxembourg altogether either strengthened or did not affect the results. We also applied the suggested method of Pindyck and Rubinfeld (1991, pp. 222–3) where *DHUM* for Luxembourg was estimated from a regression of *DHUM* against all other explanatory variables. In no case did the point estimate change at all, and the *t*-value was either unchanged or changed by no more than 0.01 in either direction. All these additional results are available upon request.

TYR is only available every five years and the latest observation is for 1990. Thus, *TYR* takes the value of the first year in the respective periods, and *DHUM* is lagged one period.

No observations for *SAV* were available for 1971. The average for the 1971–75 period is therefore calculated as the average for the 1972–75 period.

The included OECD countries are: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Switzerland, Sweden, the UK and the US.

Variable	*Definition*	*Source*
Dependent variable		
DGDP	Average annual growth rate of GDP per head, 1990 prices and exchange rates	OECD (2), WDI, IMF Fin., Taipei, Penn World Tables ver. 5.6
Variables of interest		
TAX	Total taxes as a fraction of GDP, current prices	OECD (7), WDI, IMF Gov., Hong Kong Trade Development Council, Taipei
GEXP	Government expenditure as a fraction of GDP, current prices	OECD (4), IMF Gov., Hong Kong Trade Development Council, Taipei
Variables always included		
Y_0	Initial GDP per head, current prices and current PPPs, OECD = 1, initial year for each subperiod	OECD (1), Penn World Tables ver. 5.6
INV	Investment as a fraction of GDP, current prices	OECD (6), IMF Fin., Hong Kong Trade Development Council, Taipei
DHUM	Annual growth rate of the average years of schooling in the total population,	Barro and Lee (1996), data downloaded from the NBER home page
DLAB	Average annual growth rate of the labor force	OECD (4), WDI, Taipei
EBA variables		
DEPPOP	Population aged 0–15 and 65– as a fraction of total population	WDI
EXP	Export of goods and services as a fraction of GDP, current prices	WDI, IMF Fin.
FERT	Fertility rate, births per woman	WDI
IMP	Import of goods and services as a fraction of GDP, current prices	WDI, IMF Fin.
INFL	Percentage change p.a. in the consumer price index	WDI
OPEN	Export plus import of goods and services as a fraction of GDP, current prices	WDI, IMF Fin.
POP	Total population, in thousands	WDI

PSAV	Gross private saving as a fraction of GDP, current prices	OECD (3), OECD (5)
SAV	Gross national saving as a fraction of GDP, current prices	OECD (3), OECD (5)
TYR	Average years of schooling in the total population	Barro and Lee (1996), data downloaded from the NBER home page
UNEMPL	Unemployment as a share of the labor force	OECD (4)
URBAN	Urban population as a fraction of total population	WDI

Publications:
IMF Fin. = IMF, *International Financial Statistics*, various volumes.
IMF Gov. = IMF, *Government Finance Statistics*, various volumes.
OECD (1) = OECD, *National Accounts Main Aggregates, Vol. 1, 1960–1994*, 1996.
OECD (2) = OECD, *National Accounts Main Aggregates, Vol. 1, 1960–1996*, 1998.
OECD (3) = OECD, *Economic Outlook*, Vol. 48, December 1990.
OECD (4) = OECD, *Economic Outlook*, Vol. 60, December 1996.
OECD (5) = OECD, *Economic Outlook*, Vol. 62, December 1997.
OECD (6) = OECD, *Historical Statistics*, various issues.
OECD (7) = OECD, *Revenue Statistics 1965–1996*, 1997.
WDI = World Bank (1997), *World Development Indicators*. Book and CD-ROM. Washington D.C.
Taipei = Taipei Mission in Sweden (all data for Taiwan from this source unless indicated).

Table A1 Panel Regressions for the Growth Effect of Public Sector Size in 23/22 OECD Countries 1970–95 including all control variables (including country and period dummies).

Explanatory variables	*OLS regression*	*Weighted regression*	*OLS regression*	*Weighted regression*
TAX	–0.063	–0.091*		
	(–1.09)	(–2.16)		
GEXP			–0.058	–0.86**
			(–1.35)	(–3.10)
No. of obs.	115	115	109	109
Adjusted R^2	0.56	0.79	0.60	0.85

Note: * and ** denote significance at the 5 and 1% levels, respectively. *t*-statistics in parentheses. The regressors are measured as averages for the respective subperiods, except for Y_0 which measures the income level in the initial year of each subperiod. The control variables included are listed in the data description above. For reasons spelled out in the text *SAV* is not included among the regressors.

References

Agell, J., T. Lindh and H. Ohlsson, 1997, Growth and the public sector: A critical review essay, European Journal of Political Economy 13, 33–52.

Atkinson, A.B., 1995, The welfare state and economic performance, National Tax Journal 47, 171–198.

Barro, R.J., 1990, Government Spending in a Simple Model of Endogenous Growth, Journal of Political Economy 98, S103–S125.

Barro, R.J., 1991, Economic growth in a cross section of countries, Quarterly Journal of Economics 106, 407–443.

Barro, R.J. and J.W Lee, 1996, International measures of schooling years and schooling quality, American Economic Review 86, 218–223.

Easterly, W., 1995, Comment on Slemrod, Brookings Papers on Economic Activity 2, 419–424.

Easterly, W. and S. Rebelo, 1993, Fiscal policy and economic growth: An empirical investigation, Journal of Monetary Economics 32, 417–458

Engen, E.M. and J. Skinner, 1992, Fiscal policy and economic growth, NBER Working Paper No. 4223.

Fomby, T.B., C.R Hill and S.R. Johnson, 1984, *Advanced econometric methods* (Springer Verlag, New York).

Fuente, A. de la, 1997, Fiscal policy and growth in the OECD, CEPR Discussion Paper No. 1755.

Fölster, S. and M. Henrekson 1999, Growth and the public sector: A critique of the critics, European Journal of Political Economy 15, 337–358.

Grier, K.B., 1997, Governments, unions and economic growth, in: V. Bergström, ed., Government and Growth (Clarendon Press, Oxford).

Grier, K.B. and G. Tullock, 1989, An empirical analysis of cross-national economic growth 1951–80, Journal of Monetary Economics 24, 259–276.

Hansson, P. and M. Henrekson, 1994, A new framework for testing the effect of government spending on growth and productivity, Public Choice 81, 381–401.

Islam, N., 1995, Growth empirics: A panel data approach, Quarterly Journal of Economics 110, 1127–1170.

Kiviet, J.F., 1995, “On bias, inconsistency, and efficiency of various estimators in dynamic panel data models, Journal of Econometrics 68, 53–78.

Leamer, E.E., 1983, Let’s take the con out of econometrics, American Economic Review 73, 31–43.

Levine, R. and D. Renelt, 1992, A sensitivity analysis of cross-country growth regressions, American Economic Review 82, 942–963.

Mendoza, E.G., G.M. Milesi-Ferretti and P. Asea, 1997, On the ineffectiveness of tax policy in altering long-run growth: Harberger’s superneutrality conjecture, Journal of Public Economics 66(1), 99–126.

Newey, W. and K. West, 1987, A simple positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix, Econometrica 55, 703–708.

Plosser, C., 1993, The search for growth, in: Policies for Long-Run Growth (The Federal Reserve Bank of Kansas City symposium series, Kansas City, MO).

Pindyck, R.S. and D.L. Rubinfeld, 1991, Econometric models and econometric forecasts (McGraw-Hill, New York).

Sala-i-Martin, X., 1994, Cross-sectional regressions and the empirics of economic growth, European Economic Review 38, 739–747.

Sala-i-Martin, X., 1997, I just ran four million regressions, NBER Working Paper No 6252.

Slemrod, J., 1995, What do cross-country studies teach about government involvement, prosperity, and economic growth?, Brookings Papers on Economic Activity 2, 373–431.

Tanzi, V. and H.H. Zee, 1997, Fiscal policy and long-run growth, IMF Staff Papers 44, 179–209.

White, H., 1980, A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity, Econometrica 48, 817–838.

World Bank, 1997, World development indicators, in book form and on CD-ROM (Washington D.C.).