BBVA Research

Creating Opportunities

N.º 24/02

Working Paper

Social Welfare and Government Size

J. Andrés, E. Bandrés, R. Doménech and M.D. Gadea

January 2024

Social Welfare and Government Size*

J. Andrés[a], E. Bandrés[b], R. Doménech[a,c] and M.D. Gadea[b]

a University of Valencia
b University of Zaragoza
c BBVA Research

January, 2024

Abstract

In this paper, we analyze the effect of government size on social welfare and GDP per capita growth for the sample of 36 OECD countries in the last six decades. Our results show that the effects are negative but smaller in absolute terms in the case of welfare than for GDP per capita. This result is robust to changes in the estimation method, to the use of smoothed variables, to the inclusion of dummy variables that control for expansions and recessions, and to additional control variables, such as the composition of expenditures and taxes, and public debt. More importantly, we find that the effect of government size follows an inverted U-shape: positive and greater on social welfare than for GDP per capita growth when government size is below 35% to 40%, and negative beyond that threshold. Interestingly, the range of values for which the size of the government positively affects growth and welfare expands significantly with government quality, productive spending, and low levels of debt. An efficient public sector on all these fronts is more important for maximizing social welfare than the size of the government per se.

Keywords: social welfare, GDP per capita, growth, government size.

JEL Classification: E62, H50, O40

1. Introduction

In the last decades, there has been a large literature evaluating the empirical effects of government size on GDP per capita growth (see, for example, the surveys by Johansson, 2016, and Bergh and Henrekson, 2011). The evidence is far from conclusive and varies with model specifications, variable definitions, the sample of selected countries, periods, data quality, or estimation methods. However, for advanced economies, the results con-

* The authors thank the comments of M. Cardoso, A. Gavilán, J. Sicilia, and participants at presentations in the Bank of Spain, BBVA Research, the 2023 EEA-ESEM congress, the 2023 International Association for Applied Econometrics Annual Conference, the 48th Symposium of the Spanish Economic Association, the XXV Applied Economics Meeting, and the 30th Meeting on Public Economics. J. Andrés and R. Doménech thank the financial support of grant PID2020-116242RB-I00 funded by the Spanish Ministry of Science and Innovation (MCIN/AEI/ 10.13039/501100011033), grant GVPROMETEO2020-083 funded by the Generalitat Valenciana, and grant TED2021-132629B-I00 funded by MCIN/AEI/ 10.13039/501100011033 and the "European Union NextGenerationEU/PRTR". María Dolores Gadea acknowledges the financial support from the grants PID2020-114646RB-C44, funded by the Spanish Ministry of Science and Innovation (MCIN/AEI/ 10.13039/501100011033) and grant TED2021-129784B-I00 funded by MCIN/AEI/ 10.13039/501100011033 and the "European Union NextGenerationEU/PRTR".

sistently find a negative effect of government size on growth.

The apparent adverse impact observed may appear counterintuitive from the perspective of the optimality of public sector activity. One plausible explanation for this outcome lies in the scope of the objectives of governments that extend beyond a singular focus on GDP per capita. Rather, their overarching goal may be to optimize a more comprehensive measure of societal well-being, encompassing various dimensions that necessitate the effective provision of substantial government services. Social prosperity depends, among other things, on the quality of public services and regulations, the configuration of the institutional framework, and, most importantly, on the distribution of income among the population. In particular, the literature on the effects of public spending on inequality is also abundant. In general, larger and more effective governments tend to reduce inequality, mainly through spending on education, family and child support, and ex-post distributional policies, such as taxes and transfers.

To test this hypothesis, in this paper, we extend previous findings in the empirical literature by systematically comparing the impact of government size on GDP per capita growth with its effect on the rate of growth of the measure of social welfare proposed by Jones and Klenow (2016) in a sample of 36 OECD countries in the last six decades. These authors show how their welfare index can be rigorously derived from the individual preferences used in the economic analysis of welfare and can be expressed as a function of private and consumption per capita, leisure (which depends on the number of hours worked), the distribution of income and life expectancy.

The preliminary evidence shows a lot of heterogeneity across time and space on the correlation between this welfare index and public spending over GDP, which we consider in our econometric analysis. Panel data estimates using annual data and fixed effects, which control for differences in the steady states, show a negative and statistically significant impact of government size on both GDP per capita and welfare. In general, these effects are smaller in absolute terms in the case of welfare than in the case of GDP per capita. This result is robust to changes in the estimation method, like the Mean Group (Pesaran and Smith, 1995) or the Pooled Mean Group (Pesaran et al., 1999) estimators, to the use of smoothed variables, either the Hodrick-Prescott filter or ten-year averages, and to the inclusion of other variables, such as public debt or the structure of public expenditure and taxes, or dummy variables that control for expansions and recessions over the business cycle.

When we allow the effect of government size to be time-varying, we find that the estimated coefficients are negative, statistically significant, and rather stable until the

Great Recession. Since the beginning of the financial crisis to the sovereign debt crisis, the negative effect of government size on growth has almost doubled. Additionally, we find that the difference between the negative effect of government size and welfare narrows in absolute value in the last six years of the sample.

An alternative explanation of the empirical evidence on the linear negative effects of government size on GDP per capita and welfare growth may be the presence of non-linearities and their interaction with government quality that are omitted in the analysis. This hypothesis suggests that the impact on growth can be negative when the government is either too small, leading to insufficient provision of critical public services, or too large, resulting in high taxes that hinder the efficiency of the private sector, but that it might be positive for intermediate sizes. At the same time, the optimal government size may be different for each country depending on the efficiency level of the public sector. Thus, we expect that countries with more efficient public sectors will have larger governments. Our results show that the effect of government size on welfare and GDP per capita growth is indeed non-linear, following an inverted U-shape. In particular, this effect is positive when the size of the public sector is below 35% to 40%, turning negative beyond that level. We also find that the efficiency with which public resources are used (government quality), the share of public investment in total government spending, and the level of public debt are significant determinants of the range of values for which government size has a positive effect on both indicators of well-being.

The structure of this paper is as follows. In section 2, we present a succinct review of the empirical evidence about the effects of government size on GDP per capita growth and inequality. In Section 3, we survey the literature on social welfare and analyze the properties of the welfare measure used in the rest of the paper. Section 4 presents the main results of our panel data estimates, whereas Section 5 analyses the robustness of these results to two different alternatives to take into account the possible effects of the economic cycle (controlling for expansions and recessions, and averaging the sample every 10 years), the stability over time of the effects of government size, and the sensitivity to the composition of public spending, revenues, and public debt. In Section 6 we discuss the non-linear effects of government size, taking into account government quality. Finally, Section 7 presents the main conclusions.

2. Government size, growth, and inequality: Review of the evidence

Given the novelty of the more recent contributions to the monetary approach to social welfare, there is no empirical evidence on the effects of the size and composition of public expenditure on welfare, in contrast to the large existing literature on their effects on economic growth and income inequality (see, for example, Awaworyi Churchill et al, 2017).

The complexity of the relationship between the size of public spending and growth is reflected in a vast empirical literature that, using different methodologies, time periods, and samples of countries, has tried to assess the extent to which public revenues and expenditures are a driving force for growth. The relationship between public spending and economic growth responds to two opposing forces. On the one hand, the positive effects of public spending on productivity, mainly through the accumulation of physical and human capital, the correction of market failures, and the improvement of social stability by reducing inequality. On the other hand, the distortionary effects of taxes on capital productivity and growth, as well as inefficiencies in public management and the extension of rent-seeking activities. The combined impact of these two types of effects suggests that the relationship between the size of public spending and growth could be nonmonotonic (Tanzi and Zee, 1997).

Early empirical work mostly supported the hypothesis of a statistically significant negative relationship between public spending and economic growth. However, more recent empirical findings have shown that this result is not always robust to the introduction of changes in the control variables and other issues associated with the specification of the models (Slemrod, 1995). Economic prosperity also depends, among other things, on other public activities that are not necessarily reflected in public finances, such as the quality of public services and regulations, or the configuration of the institutional framework.

An examination of the abundant empirical studies on the subject shows that one factor that has a very significant influence on the sign and value of the coefficients that relate the size of public expenditure to growth is the level of development of the countries in the sample. For example, Landau (1983) detected that the negative effect of public consumption on growth was mainly a characteristic of advanced economies, whereas it was not observed for countries whose income was below the median of the more than 100 countries analyzed.

Grier and Tullock (1989) also estimated a negative impact of the increase in the public expenditure/GDP ratio on growth. However, in the sample of less developed countries, the estimated coefficient was much lower and less significant. When the analysis was confined to advanced countries, different econometric methodologies showed very consistent and robust results, pointing to the negative effects of public spending on growth (Saunders, 1985, Hansson and Henrekson, 1994, or Pevcin, 2004). The surveys by Bergh and Henrekson (2011 and 2015) for advanced countries confirm the negative influence of spending on growth, even after correcting for possible reverse causality problems.

As regards the composition of public expenditures, Gemmell, Kneller and Sanz (2016) evaluate the long-run effects on GDP per capita of changes in total government expenditure and in the shares of different spending categories for a sample of OECD countries since the 1970s. The authors take into account the structure of taxes used to finance these expenditures and the potential endogeneity problems of the variables. They find strong evidence showing that reallocating total spending toward productive expenditures, such as infrastructure and education, has positive effects on long-run output levels. On the contrary, the reallocation of public spending towards social expenditures is associated with a modest negative effect on growth.

The results on the effects of redistribution (a reduction in inequality) through taxes and transfers are also ambiguous. Berg et al. (2016) find no significant relationship between redistribution and growth. The explanation would be that the positive and negative effects of redistributive policies on growth would neutralize each other. El-Shagi and Shao (2019) also find no significance in the impact of redistribution on growth, but there are positive effects when the level of education in the country is low and negative when it is high. Finally, Woo (2020) concludes that, taken in isolation, redistribution negatively affects growth in the medium and long term, so there is a trade-off between efficiency and equity, but the final impact depends on the size of the redistribution, the initial level of inequality and how this is reduced precisely because of the redistribution.

The literature on the effects of public spending on inequality is also abundant. Fournier and Johansson (2016) estimate the effect of the size and composition of public spending on inequality in a sample of OECD countries over the 1987-2014 period. Their more relevant findings are the following. First, larger and more effective governments tend to reduce inequality. Second, the size and effectiveness of the government have little effect on the distribution of market income. Thus, ex-post distributional policies are the main instruments used by OECD governments to reduce inequality: about two-thirds of the reduction in inequality between market and disposable income is due to transfers, and one-third is due to taxes. Third, regarding composition, spending on education and family and child support unambiguously reduces inequality. Other subsidies (for example, to firms) and generous pension systems (those cases in which pensioner income is already above the country's average) increase inequality.

Johansson (2016) finds significant distributional effects of tax and transfers across OECD countries, leading to a reduction of up to 40% of the market income Gini coefficient, mostly among European countries. Almost 75% of this reduction is accounted for by transfers, whereas the progressivity of the tax system explains the remaining 25%.

There is also substantial evidence about the positive distributional effect of education, public health, social spending, and labor market policies. Nevertheless, this positive effect does not always respond to the amount of public funds invested but to the quality of the programs involved. In particular, public spending on education and generalized access to schooling is unequivocally equity-enhancing, as both of them facilitate people from poorer backgrounds to move up along the social and economic ladder. However, the impact on the quality of student achievements (PISA scores, for example) seems to be explained by the quality of teaching and the use of other resources at the school rather than by the size of public spending on education.

In this vein, Afonso, Schuknecht, and Tanzi (2010) confirm that social spending and the performance of the education system are critical to redistributing income, but these factors do not work separately. Substantial social spending effectively diminishes inequality solely in nations with an efficient education system. The authors find that the social spending framework displays a great deal of variance in terms of efficiency in OECD countries, in particular comparing Southern Europe with Nordic countries. This variance suggests that better (or similar) results can be achieved with the same (or lower) level of public spending. In addition, the degree of efficiency in social spending is significantly correlated with the level of education achievement and the quality of the institutional framework (civil servants, the judiciary, etc.).

Unemployment benefits and other social protection programs designed to reduce inequality may end up having unintended consequences unless they are accompanied by efficient active labor market policies and other mechanisms to incentivize job search and labor market participation by their recipients. Otherwise, although they may reduce inequality temporarily (for instance in the aftermath of recessions and severe crises), they may end up making poverty and unemployment more persistent.

Agenello and Souisa (2014) look at the effect of fiscal consolidation processes, their structure, and timing on income inequality in a panel of 18 industrialized countries from 1978 to 2009. These events are informative about the connection between fiscal variables and inequality, although the channels through which this connection works may not be the same as those operating between the size and composition of the public sector and the distribution of income. In general, public spending-driven consolidations, even those successful in reducing the debt-to-GDP ratio, are detrimental to the distribution of income, particularly in the aftermath of banking crises. Low-income earners are particularly affected by consolidation in terms of low growth and high inflation. High inflation generally worsens the distribution of income and accentuates the negative effect of consolidations in public spending. The negative effects of tax cuts after a crisis increase with the size of public spending.

Doumbia and Kinda (2019) analyze the effect on income inequality of reallocating different components of total public spending, in a panel of 82 advanced and developing countries. The authors control for an extensive list of variables that have been found to affect income inequality and control for the level of public spending as a share of GDP. The results can be interpreted as the impact on inequality of an increase in a given public spending component (i.e., health or social spending) financed by the reduction in other spending components. As suggested by previous results in the literature, controlling for total public expenditure, an increase in social spending (e.g., family support, poverty programs, or unemployment benefits) and infrastructure reduce inequality, and more so if financed by a reduction in defense or education spending. Interestingly, the reduction in defense spending is only positive in reducing inequality in countries with a high level of internal security and stability. Otherwise, less defense spending is indeed detrimental to income equity, implying that social unrest and political instability undermine mostly the income of poor people. These authors find that increases in spending on education and health produce mixed results. An increase in these components, even financed by (cutting) defense spending, does not help to reduce inequality in the short term. However, when looking at a 10-year horizon, these expenses reduce inequality (although the effect of health spending is barely significant). This implies that the positive impact of these components of public spending accrues first to the middle (mostly urban) class and spreads to the rest of the population as time goes by.

In a meta-regression analysis applied to more than 900 regression results in 84 different studies, including low, middle, and high-income countries, Anderson et al. (2017) also confirm a moderate effect of public spending in reducing inequality. This effect is stronger for social spending, which helps to reduce the gap between high-income and middle-income households, but less so between middle-income and the poorest 20% of the population.

3. Social welfare and GDP per capita

GDP per capita is the most commonly used measure in comparative studies of economic performance, both across economies and time. It summarizes the value of market activities that capital and labor produce in an economy in a specific period. This makes it, at best, a partial measure of economic welfare (Aitken, 2019). For one thing, GDP misses non-market activities, the opportunity cost of leisure foregone in the production process, and the value of some goods and services associated, for example, with the digital economy. From a broader dynamic perspective, a continuous increase in the amount of production also requires a growing environmental cost that will be unequally borne by current and future generations. But most importantly of all, being an average, GDP per capita fails to capture the important effect on personal well-being that unequal distribution of income has on aggregate welfare. Leaving aside preferences for more or less equity in the distribution, risk-averse individuals value more the welfare foregone for being in the lowest deciles of the income distribution than the increase in welfare they may achieve for being in the upper deciles.

Most attempts to provide better indicators of social welfare include some measure of the inequality of income distribution between individuals in the economy. Nevertheless, it is difficult to integrate production or other similar macroeconomic variables along with inequality measures into a simple composite welfare indicator and many such proposals of broader welfare indicators have severe limitations of data gathering, homogeneity, and relative weights given each to their components.

Berik (2020) distinguishes four different recent approaches to deal with the complexity of measuring welfare for comparison purposes: composite indexes (as the UN Human Development Index), subjective evaluations (as the UN World Happiness Report), dashboards (as the OECD Better Life Initiative) and monetary approaches (for example, the Measure of Economic Welfare developed by Nordhaus and Tobin, 1972, or the more recent variant proposed by Jones and Klenow, 2016). The monetary approach is perhaps less comprehensive than other alternatives, but, in contrast to composite indexes or the dashboard approaches, it provides a well-grounded theory-based aggregation procedure of different determinants of welfare. Furthermore, it allows cross-country and intertemporal comparability in a way that subjective evaluation approaches cannot.

In this paper we follow the monetary approach and study the effect of public spending on welfare, the latter being approximated by the Jones and Klenow's (2016) index. This welfare index can be rigorously derived from the individual preferences used in the economic analysis so that welfare can be expressed as a function of private and public consumption per capita (C), leisure (ℓ, which depends on the number of hours worked), the distribution of income and life expectancy (which depends on the probability of survival, S, of living beyond a certain age, a):

$$U = E\sum_{a=1}^{100}\beta^a u(C_a,\ell_a)S(a) \tag{1}$$

The comparison of welfare across countries and time is made in terms of the equivalent annual consumption necessary for a person randomly chosen in any country to be indifferent to living in another (e.g., the United States). Although this form of comparison may seem very different from that carried out using GDP, the underlying principle is similar since with either indicator we are taking into account the capacity to enjoy the set of goods and services produced in each country. Furthermore, this comparison depends on how consumption is distributed among the individuals in an economy. If the average levels of consumption, leisure, and life expectancy are equal, it is preferable to live in a country with less inequality, insofar as people are risk-averse and/or have personal preferences in favor of a more equitable distribution of income. This justifies that the comparison of equivalent annual consumption should take distributional issues into account.

If microeconomic data from representative surveys were available for all countries, a fairly robust approximation of equivalent annual consumption could be made for each country compared to the US. In practice, we do not have this information for all OECD countries, so some additional assumptions have to be made to use information from databases with macroeconomic information for large samples of countries. Jones and Klenow (2016) show that the approximation with macroeconomic data is quite good in the case of the 13 economies (advanced and developing) for which microeconomic data are available: the correlation of welfare levels measured with macroeconomic and microeconomic data is equal to 0.999.

In particular, the relative welfare measure (λ_i) of each OECD country relative to the United States is calculated using the following expression:

$$\begin{aligned}\log\lambda_i = {} & \frac{e_i - e_{us}}{e_{us}}\left(\overline{u} + \log c_i + \nu\left(\ell_i\right) - \frac{1}{2}\sigma_i^2\right)\\ & + \log c_i - \log c_{us}\\ & + \nu\left(\ell_i\right) - \nu\left(\ell_{us}\right)\\ & - \frac{1}{2}\left(\sigma_i^2 - \sigma_{us}^2\right)\end{aligned} \tag{2}$$

where e, c, ν, and σ are, respectively, life expectancy, per capita consumption, a function of leisure, and the variance of consumption among individuals, for country i and the United States (us).

This index is particularly suitable for studying the economic and social impact of public spending. The role of the public sector in an economy is not only to foster productivity growth, through public investment, and the provision of public goods, such as property rights, defense, regulations, or justice. Among its objectives is also to ensure that the benefits of the wealth created are spread among all members of society. Public spending affects market activities directly, whose effect on personal well-being is represented by consumption per capita and leisure (the latter reflecting the efficiency with which a given amount of goods and services are produced). It also influences the distribution of income and hence inequality. Public spending is in particular a key determinant of pre-market inequality (schooling or health) and post-market inequality (taxes and transfers). Finally, health care, and other regulations, affect life expectancy.

When the information needed is available, the welfare measure for OECD countries has been calculated since 1960 or the first available year. Life expectancy at birth (e) is obtained from the Gapminder database (2020). Consumption per capita (c), GDP per capita (gdp), and the number of hours worked by the working-age population have been taken from PWT 10 (Feenstra, Inklaar and Timmer, 2015). For inequality of disposable income after taxes and transfers, we use the Gini coefficient of Eurostat (2020) and the OECD (2020a). Data from SWIID 8.3 (Solt, 2020), Atkinson et al. (2017), and Prados de la Escosura (2008) have been used to extrapolate backward.

Data availability allows us to construct an unbalanced or incomplete panel for 35 OECD countries, with observations since 1960 for Australia, Austria, Belgium, Canada, Switzerland, Germany, Denmark, Spain, Finland, France, United Kingdom, Greece, Ireland, Italy, Japan, Korea, Netherlands, Norway, New Zealand, Portugal, Sweden, and the United States, since 1970 for Colombia, Mexico, and Iceland, since 1980 for Hungary, since 1990 for Estonia, Israel, Lithuania, Slovakia, Slovenia and Turkey, and since 2000 for Luxembourg, Poland and Czech Republic.

The evidence in Figure 1 for the OECD countries using averages from 2010 to 2019 shows that social welfare is closely related to GDP per capita, which can explain 83% of the welfare differences between most countries. However, behind this high correlation, there are interesting differences between these two indicators. For example, in Ireland, GDP per capita is more than 22 percent higher than in the US, as a result of many multinationals having moved their tax bases to Ireland, but the welfare index corresponding

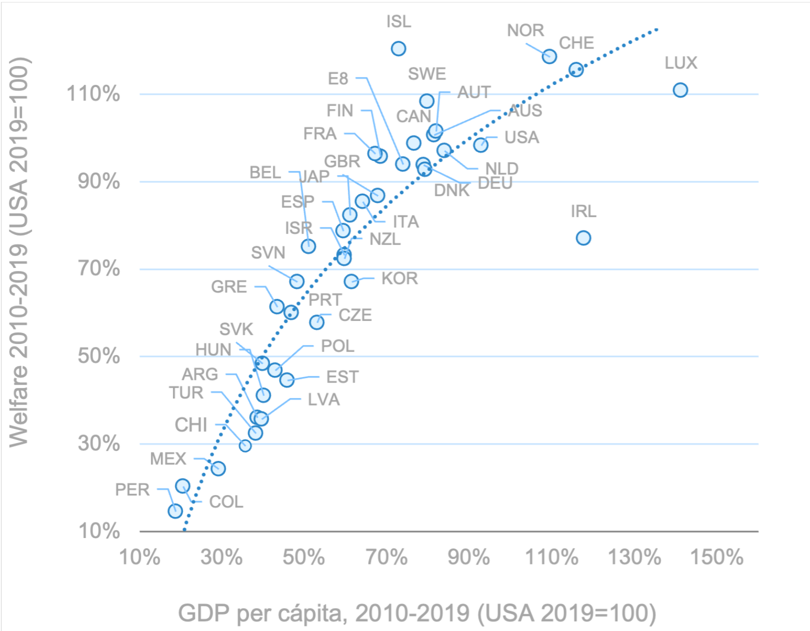


Figure 1: *GDP per capita and welfare in OECD and selected Latam countries, 2010-2019. Source: own elaboration based on Jones and Klenow (2016), PWT10, SWIID, Eurostat, OECD and Gapminder.*

to Ireland is 21 percent lower because GDP per capita does not translate into a similar relative level of consumption per capita (see Byrne, Conefrey and O'Grady, 2021). Something similar happens in the case of Luxembourg: with a GDP per capita 47% higher than in the US, its welfare level is only 13%. In this case, this disparity is mainly explained by the weight of financial services, which represent around a third of Luxembourg's GDP. Except for these particular cases, and to a lesser extent Switzerland, if we compare the most advanced countries with the most backward OECD ones, the differences in welfare are greater than those observed in terms of GDP per capita, since in countries with lower per capita income, inequality and the number of hours worked tend to be greater, and life expectancy lower.

In contrast to less developed countries, in the case of the advanced economies, their differences in well-being with respect to the United States are significantly smaller than the distance measured in terms of GDP per capita. Longer life expectancy, a better distribution of income, and fewer hours worked partially compensate for the advantage of the United States relative to the European economies in GDP and per capita consumption. For example, the average gap between the eight most advanced European economies (E8) with respect to the United States was approximately 23 percentage points in 2018 when measured in income per capita terms, while it was just over 8 points in terms of the welfare index.

4. Main results

4.1. Data

In this section, we compare the effects of government size on welfare growth with the effect on GDP per capita growth. Our results are based on data of the welfare index (*WI*), described in section 2, *GDP* per capita and total public spending (*PS*) as a percentage of GDP. The last two variables are taken from Penn World Table 10.0 and the OECD Economic Outlook for a sample of 36 countries. *PS/GDP* has been extrapolated backward, until 1960, whenever possible, using the database of Mauro et al. (2015). The lower case variables $gdp_{i,t}$, $wi_{i,t}$, and $ps_{i,t}$ represent logarithms, $i=1;...,N$, where N is the number of countries and $t=1,...,T$ the time period from 1960 to 2019, although the sample is not complete for all countries. The original data for *WI* and *PS* are shown in Appendix A.

A preliminary analysis of the evidence shows some interesting results. First, there is a great deal of heterogeneity in the sample of OECD countries. Total public spending over *GDP* ranges from 10% in Colombia during the 1960s to close to 70% in Israel around 1980. Second, on average, total public expenditure over *GDP* increased in the 1960s and 1970s but has remained relatively constant since then with countercyclical fluctuations[2]. Nevertheless, again there is heterogeneity across countries with respect to this general trend. For example, in Sweden there was an upward trend until the beginning of the 1990s, approaching 70%, and since then *PS* has gradually converged to 50%. Third, in general, relative welfare to the US in 2019 shows a positive trend. Again, there are some exceptions. In Colombia, relative welfare was almost constant from 1970 to 2005, whereas in the US it increased by 24 percentage points. Something similar happened in Mexico from 1981 to 2010, Israel from 1981 to 2006 or Denmark from 1977 to 2005. Fourth, there are countries and decades in which the correlation between WI and PS is positive, for example, in most European countries in the 1960s and 1970s. However, it is easy to find examples of a negative correlation, as in Canada, Denmark or the Netherlands from 1985 to the present, Poland from 1995 onward, and Sweden since 1993.

[2] It should be noted that, by construction, the variable *PS* is countercyclical as long as public expenditures are at least procyclical but with less volatility than *GDP*. As noted before, given the disconnection between *GDP* and domestic economic activity in Ireland from 2015, the fall of *PS* since then is artificially affected by the large inflows of intellectual property assets and foreign incomes

4.2. Panel data analysis with annual data

We begin our econometric analysis by performing a panel data estimation using annual data and fixed effects of the following equation:

$$\Delta y_{it} = \beta ps_{it} + \rho y_{it-1} + \delta_t + \alpha_i + \epsilon_{it} \tag{3}$$

where $y_{i,t} = ln(Y_{i,t})$, Y_{it} is *GDP* per capita or *WI* (the welfare index), $ps_{i,t}$ is the log of the ratio of public expenditure to GDP of the country i in year t, δ_t is the fixed time effect for year t, which allows correction of common economic cycles between countries, and α_i is the fixed effect for country i that controls differences in GDP per capita and welfare steady states.

Two types of panels have been estimated. The first includes all 36 countries considered in our study, although they have different sample sizes. It is, therefore, an unbalanced panel and we denote it as the *full panel*. The second includes only those countries for which we have data for the full period 1960-2019, in general more advanced economies. In this case, the panel is balanced. We denote it as the *restricted panel* and the number of countries is reduced to 22. The choice of a fixed effects specification instead of random effects is supported by the Hausman test in all estimates. In any case, the results obtained with the random effects model are very similar. Furthermore, the standard errors of the fixed effects model have also been robustly estimated, adjusting the heteroskedasticity for N clusters.

The first column in Table 1 shows the result of the panel regression of equation 3 for GDP per capita, with the unbalanced (top) and balanced (bottom) samples, imposing the same β for all countries and omitting fixed effects for the countries. The effect of government size is negative and statistically significant in the balanced panel, whereas the rate of convergence (the coefficient of $gdp_{i,t-1}$) is very similar. In column (2) we confirm the negative effect of ps, now using the growth of welfare as the dependent variable, although the estimated coefficient is smaller than in column (1). In both cases, we observe that the negative effects of the government size are greater and the convergence rates smaller in the case of the balanced sample. A possible explanation is that, in general, the countries excluded in the balanced sample are those with lower initial levels of GDP per capita and rapid convergence.

Table 1: *Results of panel estimates*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Panel A: full sample															
	Original variables								Smoothed variables							
	Non Fixed-effects		Fixed-effects		Mean group		Pool mean group		Non Fixed-effects		Fixed-effects		Mean group		Pool mean group	
	$\Delta gdp_{i,t}$	$\Delta wi_{i,t}$	$\Delta gdp_{i,t}$	$\Delta wi_{i,t}$	$\Delta gdp_{i,t}$	$\Delta wi_{i,t}$	$\Delta gdp_{i,\bar{t}}$	$\Delta wi_{i,t}$	$\Delta\widetilde{gdp}_{i,\bar{t}}$	$\Delta\widetilde{wi}_{i,t}$	$\Delta\widetilde{gdp}_{i,\bar{t}}$	$\Delta\widetilde{wi}_{i,\bar{t}}$	$\Delta\widetilde{gdp}_{i,t}$	$\Delta\widetilde{wi}_{i,t}$	$\Delta\widetilde{gdp}_{i,\bar{t}}$	$\Delta\widetilde{wi}_{i,t}$
$ps_{i,t}$	−0.011	−0.006	−0.049	−0.033	−0.096	−0.052	−0.028	−0.036	−0.005	−0.003	−0.034	−0.026	−0.087	−0.042	−0.040	−0.043
	(−1.41)	(−0.72)	(3.80)	(−2.91)	(−5.65)	(−4.43)	(−2.50)	(−3.33)	(−1.22)	(−0.37)	(−2.51)	(−1.99)	(−4.13)	(−3.46)	(−2.53)	(−2.16)
$gdp_{i,t-1}$	−0.009		−0.015		−0.026		−0.040									
	(−2.16)		(−1.67)		(−3.23)		(−5.91)									
$wi_{i,t-1}$		−0.010		−0.016		−0.015		−0.033								
		(−3.84)		(−1.76)		(−1.86)		(−4.13)								
$\widetilde{gdp}_{i,t-1}$									−0.010		−0.008		−0.010		−0.028	
									(−3.38)		(−1.67)		(−1.57)		(−3.88)	
$\widetilde{wi}_{i,t-1}$										−0.010		−0.005		0.015		−0.020
										(−4.05)		(−0.72)		(3.08)		(−1.99)
Error Correction term							−0.875	−0.891							−0.007	−0.007
							(−17.39)	(−17.04)							(−10.32)	(−10.32)
R^2	0.33	0.29	0.42	0.32	0.52	0.39	0.59	0.56	0.40	0.37	0.70	0.58	0.82	0.75	0.99	0.99
Countries	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36
Observations	1775	1775	1775	1775	1775	1775	1596	1596	1775	1775	1775	1775	1775	1775	1596	1596
	Panel B: restricted sample															
	Original variables								Smoothed variables							
	Non Fixed-effects		Fixed-effects		Mean group		Pool mean group		Nonfixed effects		Fixed-effects		Mean group		Pool mean group	
	$\Delta gdp_{i,\bar{t}}$	$\Delta wi_{i,t}$	$\Delta gdp_{i,\bar{t}}$	$\Delta wi_{i,t}$	$\Delta gdp_{i,\bar{t}}$	$\Delta wi_{i,t}$	$\Delta gdp_{i,\bar{t}}$	$\Delta wi_{i,t}$	$\Delta\widetilde{gdp}_{i,\bar{t}}$	$\Delta\widetilde{wi}_{i,\bar{t}}$	$\Delta\widetilde{gdp}_{i,\bar{t}}$	$\Delta\widetilde{wi}_{i,\bar{t}}$	$\Delta\widetilde{gdp}_{i,\bar{t}}$	$\Delta\widetilde{wi}_{i,\bar{t}}$	$\Delta\widetilde{gdp}_{i,\bar{t}}$	$\Delta\widetilde{wi}_{i,\bar{t}}$
$ps_{i,t}$	−0.025	−0.019	−0.068	−0.051	−0.078	−0.053	−0.028	−0.040	−0.021	−0.018	−0.057	−0.047	−0.062	−0.046	−0.060	−0.064
	(−2.76)	(−2.20)	(−6.11)	(−4.66)	(−11.32)	(−8.22)	(−3.77)	(−5.67)	(−2.38)	(−2.00)	(−6.01)	(−4.73)	(−13.35)	(−6.44)	(−9.88)	(−5.51)
$gdp_{i,t-1}$	−0.007		−0.011		−0.029		−0.031									
	(−1.05)		(−1.14)		(−2.92)		(−4.26)									
$wi_{i,t-1}$		−0.010		−0.009		−0.013		−0.024								
		(−2.02)		(−1.07)		(−1.50)		(−3.17)								
$\widetilde{gdp}_{i,t-1}$									−0.007		−0.006		−0.011		−0.025	
									(−1.16)		(−1.06)		(−1.60)		(−3.22)	
$\widetilde{wi}_{i,t-1}$										−0.010		−0.001		0.011		−0.012
										(−2.03)		(−0.21)		(2.28)		(−1.33)
Error Correction term							−0.924	−0.950							−0.008	−0.007
							(−15.28)	(−15.64)							(−8.80)	(−7.02)
R^2	0.42	0.37	0.52	0.42	0.55	0.44	0.60	0.55	0.58	0.55	0.79	0.69	0.85	0.80	0.99	0.99
Countries	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22
Observations	1298	1298	1298	1298	1298	1298	1188	1188	1298	1298	1298	1298	1298	1298	1188	1188

Cluster robust t-ratios in brackets; all estimates include time effects; $\Delta gdp_{i,t}$ and $\Delta wi_{i,t}$ denote the annual growth rate of *GDP per capita* and *WI* growth; $ps_{i,t}$ denotes the public spending GDP ratio in logs; $\widetilde{gdp}$, $\widetilde{gdp}$ and $\widetilde{gdp}$ represent the log of GDP, WI and PS/GDP smoothed with the Hodrick-Prescott filter ($\lambda = 100$); The model with fixed effects estimates equation 3 which includes time effects; in columns (1) and (2) cross-country effects are removed; mean-group models are estimated according to equation 4 and pooled mean group according to equation 7.

In columns (3) and (4) we now include fixed effects to control for omitted variables that determine the steady state of the growth equation. As we observe, the effect of the government size is larger in absolute terms and statistically significant in both the balanced and unbalanced panels. As in columns (1) and (2), the effects of PS are smaller in absolute terms in the case of welfare than for GDP per capita.[3]

As noted above, in columns (1) to (4) in Table 1 we have assumed that β is the same between countries. In columns (5) to (8) we relax this restriction using two different approaches. In the first one (columns (5) and (6)), in the same vein as Gemmell, Kneller and Sanz (2011), we estimate a specification similar to the Mean Group (MG) estimator (Pesaran and Smith, 1995).[4] In particular, we allow now in equation (4) for country-specific effects of government size on GDP per capita and welfare (β_i), that is

$$\Delta y_{it} = \beta_i ps_{it} + \rho y_{it-1} + \delta_t + \alpha_i + \epsilon_{it} \tag{4}$$

Then we define the Mean Group estimate of β as:

$$\widehat{\beta}_{MG} = N^{-1} \sum_{i=1}^{N} \widehat{\beta}_i \tag{5}$$

and its standard error as

$$se(\widehat{\beta}_{MG}) = \left\{ [N(N-1)]^{-1} \sum_{i=1}^{N} (\widehat{\beta}_i - \widehat{\beta}_{MG})^2 \right\}^{-1/2} \tag{6}$$

In the second approach (columns (7) and (8)), the Pooled Mean Group (PGM) (Pesaran et al., 1999), we start from an ARDL specification that estimates the long-run relationship, including the short-run dynamics. The PMG can be considered as an intermediate estimator, which involves both pooling and averaging approaches. In our specification we allow the intercepts (α_i) and the long-run coefficient of the government size (β_i) to differ freely across countries, but we constrain the error correction and the short-run

[3] We have also estimated the dynamic impulse responses of GDP per capita and welfare growth to changes in the government size based on the local projections methodology proposed by Jordá (2005), adapted to our panel data, in which we include time and country fixed effects, as in columns (3) and (4) in Table 1. The results of this exercise for the unbalanced and balanced panels show a negative and statistically significant effect of the unexpected component of government size on growth, for both GDP per capita and welfare. The negative effects are of a magnitude similar to those in Table 1.

[4] We include in Appendix B additional country analysis based on DOLS estimates

dynamics coefficients and the time dummies to be the same between countries:

$$\Delta\Delta y_{i,t} = \theta(\Delta y_{i,t-1} - \alpha_i - \rho y_{i,t-2} - \beta_i ps_{i,t-1} - \delta_t) + \sum_{p=1}^{P} \lambda_p \Delta\Delta y_{i,t-p} + \sum_{q=1}^{Q} \gamma_q \Delta ps_{i,t-q} + \epsilon_{it} \quad (7)$$

Columns (5) to (8) in Table 1 summarize the results obtained with these two models. The main conclusion is that, regardless of the estimation method, we find a negative and significant effect of the size of the public sector on the growth of GDP per capita and of the welfare index, once country-fixed effects are taken into account, although this negative effect is usually smaller in the latter case. Additionally, the negative sign and the strong significance of the error correction term confirm the hypothesis that we are working with stationary variables or that there is a long-run cointegration relationship between them.[5]

The negative sign of the government size coefficient in columns (1) to (8) may be a consequence of the business cycle. For example, as growth is negative in recessions, GDP falls, and government expenditure over GDP increases, even in the case of a neutral fiscal policy. As a result, *PS* is countercyclical. As a first attempt to correct for this potential bias, in columns (9) to (16) we replicate the previous estimations of columns (1) to (8) using the trend component obtained after smoothing the logs of GDP per capita, welfare and government expenditures over GDP with the Hodrick and Prescott filter. The dependent variables correspond to the annual rate of growth of the trend components of GDP per capita and welfare. As expected, the coefficients of *PS* are smaller in absolute value (except for the Pool Mean Group estimator), but the effect is still negative and statistically significant in all specifications in which country-fixed effects are included.

The results in Table 1 can be summarized in two conclusions. First, when we replace *GDP* with the welfare index (*WI*), the negative effect persists but is smaller in absolute value in most specifications.[6] Second, there is strong cross-country heterogeneity in the results, and taking into account unobserved differences across countries is very relevant, but controlling for these differences does not change the negative effects of government size.

[5] Unit root tests to check if ΔGDP and ΔWI they are integrated of order 1 would not be very reliable in this case due to their known problem of lack of power in finite samples and the more than likely presence of structural breaks.

[6] As shown in Appendix C, the effects of government size on welfare growth are mainly driven by the component associated with the growth rate of per capita consumption, although it is also statistically significant for life expectancy, leisure, and inequality components.

5. Robustness analysis

In this section, we analyze the robustness of previous results to two different alternatives to take into account the possible effects of the economic cycle (controlling for expansions and recessions, and averaging the sample every 10 years), the stability over time of the effects of government size, and the sensitivity to the composition of public spending, revenues, and public debt.

5.1. Controlling for expansions and recessions

In the spirit of Burns and Mitchell (1946), we consider phases of expansion and recession. According to this framework, *GDP* generally grows until it reaches its (local) maximum and a contraction phase begins as *GDP* starts to decline. After some time of negative growth, *GDP* reaches its (local) minimum, and an expansion begins as *GDP* growth rates become positive. We estimate the specific business cycle turning point chronologies by applying a yearly adaptation of the non-parametric dating procedure that was introduced by Bry and Boschan (1971) at the monthly frequency.[7] This algorithm consists of a set of filters and rules, based on moving averages of the data with different windows, which isolates the local minima and maxima in the log levels of the national series of seasonally adjusted *GDP*, subject to constraints on both the length and amplitude of expansion and contraction periods. Figure 5 in Appendix D shows the result of the business cycle dating for different countries.

The panel data model is now re-estimated with an additional variable *bc* that takes the value 1 in times of recession and 0 in times of expansion:

$$\Delta y_{i,t} = \beta ps_{i,t} + \rho y_{i,t-1} + \alpha_i + \theta bc_{i,t} + \delta_t + \epsilon_{i,t} \tag{8}$$

In all the specifications in Table 2 the business cycle is, as expected, clearly significant. Nevertheless, the effects of *ps* are statistically significant and only slightly smaller, in absolute value, than in fixed-effect models in Table 1, both in the full and balanced samples. Therefore, after including *bc* in the complete sample, β is now equal to -0.036 for GDP per capita growth (compared to -0.049 in column (3), Table 1), and -0.024 for welfare growth (-0.033 in column (4), Table 1).

[7] See Harding and Pagan (2002) for an extension to a quarterly frequency.

Table 2: *Panel estimation with a business cycle dummy*

	Full sample		Balanced sample	
	$\Delta gdp_{i,t}$	$\Delta wi_{i,t}$		
$ps_{i,t}$	−0.036 (−3.335)	−0.024 (−2.505)	−0.055 (−6.372)	−0.043 (−5.060)
$gdp_{i,0}$	−0.016 (−2.204)		−0.013 (−1.579)	
$wi_{i,0}$		−0.015 (−2.137)		−0.010 (−1.347)
$bc_{i,t}$	−0.042 (−12.641)	−0.028 (−7.770)	−0.037 (−12.954)	−0.023 (−6.137)
$\overline{R}^2$				
Countries	36	36	22	22
Observations	1775	1775	1298	1298

Robust t-ratios in brackets. See also the note in Table 1.

5.2. Averaging the sample every 10 years

An alternative way to eliminate the influence of the economic cycle is to average the sample every 10 years. In Table 3 we present the results of the reestimation of columns (1) to (8) in Table 1, using rolling 10-year averages. In particular, in the first fourth columns, we estimate the following equation

$$\frac{y_{it} - y_{it-10}}{10} = \beta \overline{ps}_{it}^{10} + \rho y_{it-10} + \delta_t + \alpha_i + \epsilon_{it} \tag{9}$$

where the dependent variable is the average rate of growth for a 10-year interval and $\beta \overline{ps}_{it}^{10}$ is the average level of public spending for each 10-year interval.

Columns (1) to (2) do not include fixed effects. The government size coefficients in the equations for GDP per capita and welfare are close to zero or even positive in the unbalanced sample. However, they are not statistically significant. As we control for fixed effects, in columns (3) to (4), the coefficients of *ps* become negative and statistically significant. These results suggest that the omission of country-specific effects may bias the estimated coefficients of government size towards zero, in contrast to the negative coefficients estimated in Table 1.

In columns (5) to (6) in Table 3 we re-estimate the Mean Group (MG) estimator, using rolling 10-year averages, that is,

$$\frac{y_{it} - y_{it-10}}{10} = \beta_i \overline{ps}_{it}^{10} + \rho y_{it-10} + \delta_t + \alpha_i + \epsilon_{it} \tag{10}$$

The estimated values of β are also negative and statistically significant. Again, the coef-

Table 3: *Results of panel estimates with 10-year averages*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Full sample							
	Non-fixed effects		Fixed Effects		Mean Group		Pool Mean Group	
	Δgdp	Δwi	Δgdp	Δwi	Δgdp	Δwi	Δgdp	Δwi
$\overline{ps}_{it}^{10}$	-0.001	0.001	-0.033	-0.024	-0.085	-0.047	-0.010	-0.111
	(-0.09)	(0.16)	(-2.48)	(-1.65)	(-5.43)	(-2.77)	(-0.18)	(-3.16)
$gdp_{i,t-11}$	-0.013		-0.026		-0.04		-0.01	
	(-3.57)		(-3.18)		(-5.57)		(-1.67)	
$wi_{i,t-11}$		-0.013		-0.026		-0.039		0.02
		(-4.69)		(-2.57)		(-3.95)		(1.98)
ECM coefficient							-0.152	-0.145
							(-10.62)	(-10.06)
$\overline{R}^2$	0.43	0.43	0.74	0.68	0.84	0.76	0.52	0.41
Countries	36	36	36	36	36	36	36	36
Observations	1451	1451	1451	1451	1451	1451	1272	1272
	Balanced sample							
	Non-fixed effects		Fixed Effects		Mean Group		Pool Mean Group	
	Δgdp	Δwi	Δgdp	Δwi	Δgdp	Δwi	Δgdp	Δwi
$\overline{ps}_{it}^{10}$	-0.015	-0.014	-0.051	-0.043	-0.054	-0.042	-0.028	-0.064
	(-2.64)	(-1.18)	(-5.10)	(-3.31)	(-7.50)	(-6.24)	(-4.24)	(-5.28)
$gdp_{i,t-11}$	-0.013		-0.018		-0.037		-0.006	
	(-2.34)		(-2.70)		(-4.99)		(-1.07)	
$wi_{i,t-11}$		-0.012		-0.015		-0.025		0.022
		(-2.33)		(-2.05)		(-2.66)		(2.63)
ECM coefficient							-0.155	-0.169
							(-9.15)	(-9.98)
$\overline{R}^2$	0.59	0.58	0.81	0.72	0.87	0.79	0.52	0.46
Countries	22	22	22	22	22	22	22	22
Observations	1100	1100	1100	1100	1100	1100	990	990

Robust t-ratios in brackets. See also the note in Table 1.

ficient of government size is smaller in absolute terms for the welfare equation (column (6)) than for GDP capita (column (5)). The last two columns of Table 3 show the results of the Pool Mean Group and, as in columns (3) to (6), the coefficients of *ps* are negative in both specifications.

5.3. The stability over time of the effects of government size

Following previous contributions to the empirical growth literature, such as Acemoglu and Molina (2022) on the effects of democracy upon economic growth, we have reestimated equation (9) also allowing the effect of government size to vary over time, that

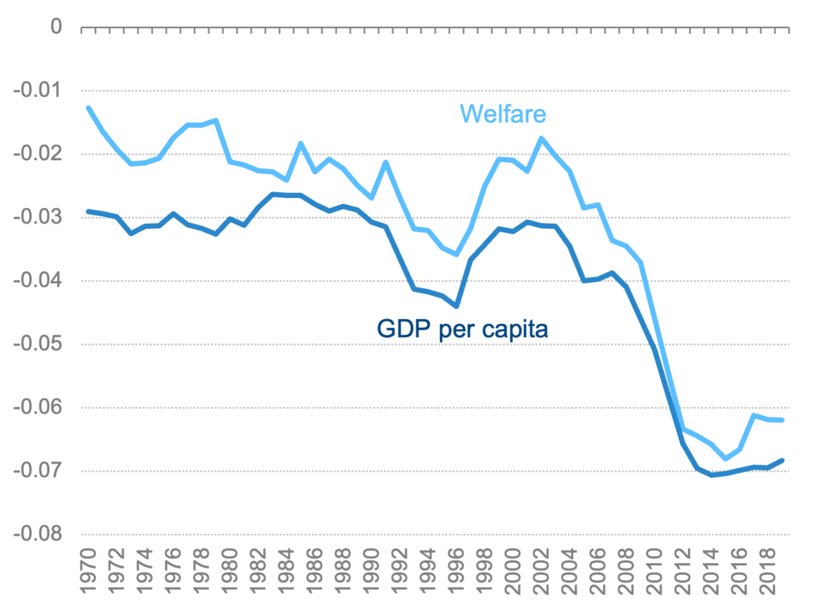


Figure 2: *Coefficients of government size over time in welfare and GDP per capita regressions. Source: own elaboration.*

is:

$$\frac{y_{it} - y_{it-10}}{10} = \beta_t \overline{ps}_{it}^{10} + \rho y_{it-10} + \delta_t + \alpha_i + \epsilon_{it} \tag{11}$$

In Figure 2 we represent the estimated values of β_t for the equations with GDP per capita and welfare. The estimated government size coefficients are negative, statistically significant, and fairly stable until the Great Recession. From the beginning of the financial crisis to the sovereign debt crisis, the negative effect of $\overline{ps}_{it}^{10}$ almost doubled. This might be due to that this was just a period of higher-than-normal government size and lower-than-normal growth rates. The financial crisis reduced the potential growth of GDP and welfare, but countries maintained the rate of growth of public spending during the next decade, then $\overline{ps}_{it}^{10}$ increased, reinforcing the effect estimated from government size to growth and welfare estimated in previous decades. For this argument to be valid, it is necessary to assume that the time dummies included in our specification do not capture this singularity well.

An alternative interpretation suggests that the expansion of government size during the Great Recession, undertaken to forestall a more substantial economic contraction and mitigate the adverse distributional impacts of the downturn, compelled many administrations to augment transfers at the cost of reducing public investment and other productive expenditures. This was particularly pronounced amid a severe credit crunch, where such transfers exhibited stronger multiplier effects. This change in the composition of public spending, coupled with the potentially adverse growth effects associated with historically elevated levels of public debt in numerous countries, weighed down economic growth and consumption in the subsequent years. As we show in the next subsection, there exists supporting evidence for this explanation, pointing to a combination of reduced public investment and increased public debt as contributing factors to the observed pattern in Figure 2.

Figure 2 also shows that the negative effect of government size on welfare is smaller in absolute value than for GDP per capita, during most of the sample period, although the difference almost disappeared in the final part of the sample.

5.4. Controlling for the composition of expenditures and taxes, and public debt

Previous theoretical and empirical contributions have pointed out significant differences between the effects of productive and unproductive government spending on economic growth (see, for example, Barro and Sala-i-Martin, 1992, Romero-Ávila and Strauch, 2008, Gemmell, Kneller and Sanz, 2011, or Fournier and Johansson, 2016). In general, the results of this literature confirm that public investment and other productive government expenditures have a positive effect on economic growth, whereas transfers, subsidies and other social expenditures have negative effects. Other studies have also controlled for the inclusion of public debt and the tax structure.

We focus on the robustness of the coefficient of government size in the welfare growth equation (3), with annual data and fixed effects, when we control for the government expenditure composition, the public debt to GDP ratio, and the tax structure:

$$\Delta\widetilde{wi}_{i,t} = \theta\widetilde{wi}_{i,t-1} + \beta\widetilde{ps}_{i,t} + \phi_x\widetilde{\mathbf{x}}_{i,t} + \delta_t + \alpha_i + \epsilon_{i,t} \qquad (12)$$

where $\widetilde{x}$ stands for the trend components of the three shares of public expenditures (public investment ($\widetilde{gi/ps}$), public consumption ($\widetilde{gc/ps}$) and the rest of expenditures ($\widetilde{gr/ps}$) in total public expenditures), for ratio of public debt to GDP ($\widetilde{d/gdp}$), and for the different tax shares considered. In particular, we include the share over total taxes of social security contributions ($\widetilde{ssc/t}$), indirect taxes ($\widetilde{ti/t}$), and other direct taxes ($\widetilde{tr/t}$), different from corporate taxes ($\widetilde{tc/t}$), which are not included since the sum of the different shares adds up to 1.

In Table 4 we present the panel estimated with these additional controls. The specification in column (1) is the same as in Table 2, column (12), but with a smaller number

Table 4: *Panel estimates with government expenditures, debt, and taxes.*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\Delta\widetilde{wi}_{i,t}$	$\Delta\widetilde{wi}_{i,t}$	$\Delta\widetilde{wi}_{i,t}$	$\Delta\widetilde{wi}_{i,t}$	$\Delta\widetilde{wi}_{i,t}$	$\Delta\widetilde{wi}_{i,t}$	$\Delta\widetilde{wi}_{i,t}$
$\widetilde{wi}_{i,t-1}$	-0.002	-0.009	-0.004	-0.004	-0.004	-0.012	-0.018
	(2.34)	(4.30)	(2.36)	(2.04)	(2.23)	(5.93)	(8.87)
$\widetilde{ps}_{i,t}$	-0.042	-0.033	-0.040	-0.037	-0.030	-0.026	-0.020
	(10.2)	(8.09)	(9.40)	(8.78)	(5.49)	(9.55)	(7.27)
$\widetilde{gi/ps}_{i,t}$		0.010					
		(8.42)					
$\widetilde{gc/ps}_{i,t}$			0.008				
			(2.05)				
$\widetilde{gr/ps}_{i,t}$				-0.012			
				(4.22)			
$\widetilde{d/gdp}_{i,t}$					-0.004		
					(3.80)		
$\widetilde{ssc/t}_{i,t}$							0.005
							(5.55)
$\widetilde{ti/t}_{i,t}$							0.008
							(3.34)
$\widetilde{tr/t}_{i,t}$							-0.001
							(16.1)
R2 adj	0.717	0.730	0.717	0.720	0.721	0.695	0.715
Countries	36	36	36	36	36	36	36
Observations	1279	1279	1279	1279	1279	1483	1483

Robust t-ratios in brackets. See also the note in Table 1.

of observations given the availability of government expenditure and public debt data. Given government size, a larger share of public investment exerts a positive effect on welfare (column (2)).

The effects of the share of public consumption (column (3)) are similar to those of public investment, although not as statistically significant. Column (4) confirms that the effects of the share of the rest of public expenditures (mainly transfers and cash benefits) have a negative and statistically significant effect on welfare, as the public debt to GDP ratio in column (5). The same happens when we control for the composition of taxes in column (7), although in this case, we have more available observations than in columns (1) to (5). For this reason, in column (6) we estimate the base specification with the same number of observations as in column (7).

In all cases, the coefficient of government size continues to be statistically significant and is not much affected by the composition of taxes and expenditures, of the level of public debt, confirming that the main result of Table 2 is robust to these additional controls.[8]

Finally, we have tested to what extent the rolling estimates in Figure 2 are affected by these additional control variables. In particular, we have found that the fall of public investment and the increase of public debt after the Great Recession could explain 1/3 of the increase in the negative effects of government size.

6. Non-linear effects and government quality

The robustness of the negative sign of the estimated average effect of government size on growth and welfare, along with the heterogeneity observed across countries on this effect, leads us to explore the second hypothesis presented in the introduction, namely the presence of strong non-linearities in this effect as well as the relevance of its interaction with government quality in terms of the institutional framework and the efficiency of the public sector in the provision of goods and services.

There are good reasons to think that the effects of government size on welfare and growth may be nonlinear. As pointed out by theoretical contributions such as Barro (1990), there could be an optimal level of government size that maximizes GDP per capita or welfare so that too small or too large government sizes would be less than optimal, or outright detrimental, in promoting GDP and welfare growth. In the former case, due to

[8] According to column (7) in Table 4, for a given level of public spending, financing it through direct taxes would be more harmful to well-being than through indirect taxes.

an insufficient provision of critical public services to the economy, and in the latter case, due to the waste of resources that could be more efficiently used by the private sector. However, this optimal level may be different for each country depending on the level of efficiency of the public sector.

To test to what extent the effects of government size are affected by the quality of the public sector, we have used the Government Quality Indicator from The Quality of Government Institute of the University of Goteborg. This quality indicator is the result of adding three components: the corruption of the political system, the impartiality of the judicial system and compliance with the law, and the quality of public administrations. Scandinavian countries are in the top, followed by some countries in central Europe, New Zealand, Canada, and Australia. We have data for all countries in our sample with the exception of Colombia. For most of them, the initial year is 1984, although for some countries the available data begins in 1993 or even later.

As expected, the correlation between the trend component of the government size log (*ps*), obtained with the Hodrick and Prescott filter, and the log of the Government Quality Indicator (*gq*) is positive and statistically significant. The coefficient of *gq* in a regression equation for *ps* is equal to 0.508 with a $t-$ratio of 16.7. This potential endogeneity problem of government size to government quality causes multicollinearity. Furthermore, *gq* is rather stable in some countries, making it difficult to differentiate its effects from the country dummies included in our specifications. Both problems could make it difficult to estimate significant coefficients for *gq* and its interaction with *ps*. To circumvent these problems, we follow an alternative two-stage approach. First, we regress *ps* on *gq*. Second, we include the square of the residuals of this regression $(ps_{i,t} - \widehat{ps}_{i,t})^2$ in our panel equation, as well as the level of government quality $(gq_{i,t})$ and its interaction with government size $(ps_{i,t} * gq_{i,t})$, that is,

$$\Delta y_{it} = \beta_1 ps_{it} + \beta_2 (ps_{i,t} - \widehat{ps}_{i,t})^2 + \beta_3 gq_{i,t} + \beta_4 ps_{i,t} * gq_{i,t} + \rho y_{it-1} + \delta_t + \alpha_i + \epsilon_{it} \qquad (13)$$

Table 5 shows the estimated parameters of equation (13) for the full sample and the trend components of GDP per capita, welfare and government size. As government quality evolves smoothly and does not depend on the business cycle, we do not filter this variable. For comparability, in columns (1) and (2) we estimate the same equations as in columns (11) and (12) of Table 1, but for the sample with data for government quality. Again, the coefficients of *ps* are statistically significant. When we add the squared residuals $(ps_{i,t} - \widehat{ps}_{i,t})^2$, the statistical significance of the size of the government increases. The

Table 5: *Panel results with the government efficiency variable*

	(1) Δgdp	(2) Δwi	(3) Δgdp	(4) Δwi	(5) Δgdp	(6) Δwi
$ps_{i,t}$	-0.022	-0.032	-0.030	-0.039	-0.039	-0.046
	(-4.26)	(-7.31)	(-7.46)	(-9.64)	(-8.07)	(-10.2)
$(ps_{i,t} - \widehat{ps}_{i,t})^2$			-0.030	-0.024	-0.033	-0.026
			(-4.04)	(-2.93)	(-4.64)	(-3.38)
$gq_{i,t}$					0.072	0.061
					(1.92)	(1.94)
$ps_{i,t} * gq_{i,t}$					-0.020	-0.017
					(-1.95)	(-2.02)
$gdp_{i,t-1}$	-0.019		-0.020		-0.022	
	(-6.66)		(-6.91)		(-6.81)	
$wi_{i,t-1}$		-0.003		-0.003		-0.004
		(-1.16)		(-0.96)		(-1.32)
R^2	0.742	0.719	0.751	0.725	0.752	0.726
Countries	36	36	36	36	36	36
Observationss	1157	1157	1157	1157	1157	1157

Robust t-ratios in brackets. See also the note in Table 1.

squared residuals enter with the expected negative sign. The interpretation of $\beta_2 < 0$ is that the deviation of government size with respect to the level predicted by the quality of government reduces the growth of GDP per capita and welfare.

Finally, in columns (5) and (6) we add the quality of government and its interaction with government size. As expected, we estimate a positive coefficient for government quality. To further analyze the effects of government size, we have simulated its effects on the growth rates of GDP per capita and welfare, holding constant government quality at two different values: its average and the maximum level of quality.[9]

Figure 3 represents the results of this simulation. As we observe, when government quality is at its average value, the effects of government size are positive on growth rates for values of *ps* below 41.1% of GDP. According to our estimates, the maximum positive effects occur when the government size is below 30% of GDP. When government quality is at its maximum level, the curves shift upward and to the right. Positive effects on growth are greater since β_3 is positive. Furthermore, as the quality of government

[9] In this exercise the government size has been expressed in deviation with respect to its sample mean.

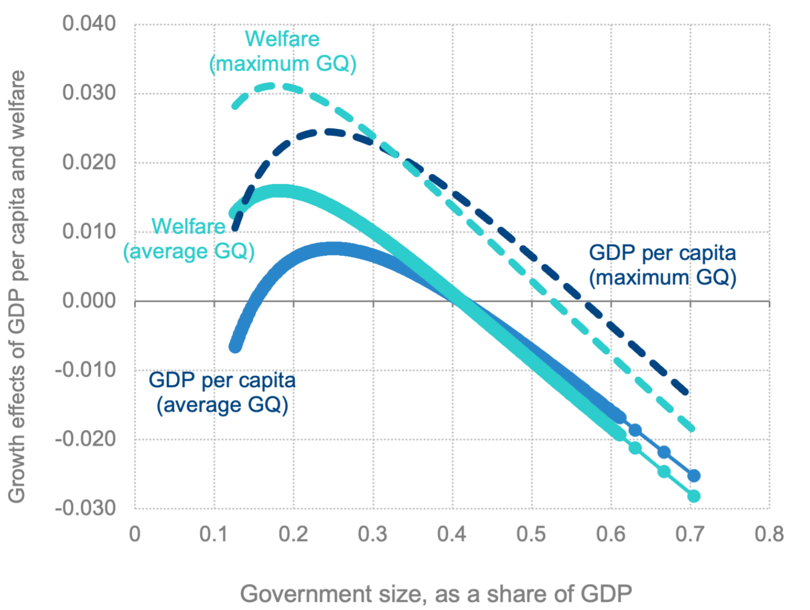


Figure 3: *Effects of government size on GDP per capita and welfare growth rates for two different values of government quality.*

improves, the range of government sizes that have positive effects on growth and well-being expands.

We have also checked the robustness of this result to additional controls. In particular, as in the previous section, when we include the composition of expenditures and taxes, and public debt, the inverted U-shape represented in Figure 3 holds the region in which government size has a positive effect on welfare expands. This result indicates that the effects of the public sector on welfare depend on the quality with which the public sector provides goods and services, the composition of spending, how it is financed, and whether it does so without incurring excessive levels of debt. An appropriate government behavior on all these fronts is crucial when it comes to maximizing social well-being.

7. Conclusions

In this paper, we have systematically compared the effects of government size on social welfare growth with that on GDP per capita growth, using different econometric techniques and robustness analysis for the sample of 36 OECD countries in the last six decades. For this purpose, we rely on the welfare measure proposed by Jones and Klenow

(2016).

Panel data estimates using annual data and fixed effects show negative and statistically significant effects of government size on both GDP per capita and welfare. These effects are smaller in absolute terms in the case of welfare than in the case of GDP per capita. This result is robust to changes in the estimation method, to the use of smoothed variables, to the inclusion of dummy variables that control for expansions and recessions throughout the business cycle, and to additional control variables, such as the composition of public expenditures and taxes, and public debt.

When we allow the effect of government size to be time-varying, we find that the estimated coefficients are negative, statistically significant, and rather stable until the Great Recession. Since the beginning of the financial crisis to the sovereign debt crisis, the negative effect of government size on growth has almost doubled. Additionally, we find that the effect on welfare is smaller in absolute value than for GDP per capita, during most of the sample period.

Finally, we have tested the presence of non-linear effects of government size on growth, controlling for government quality. We have found that the effects of government size follow an inverted U-shape. The effects on social welfare are positive and greater than for GDP per capita growth when the government size is below 40%, and improvements in government quality, as well as higher public investment and lower public debt, can significantly increase the range of government size values for which its effect on growth and welfare is positive. An appropriate government behavior on all these fronts is key to maximizing social well-being.

References

Afonso, A. y Furceri, D. (2010): "Government size, composition, volatility and economic growth." *European Journal of Political Economy*, 26, 517-532. http://t.ly/87iua

Afonso, A., L. Schuknecht, and V. Tanzi (2010): "Income distribution determinants and public spending efficiency", *Journal of Economic Inequality*, 8, 367–389. https://rb.gy/fyznb6

Agnello, L. and R. Sousa (2014): "How does Fiscal Consolidation Impact on Income Inequality", *Review of Income and Wealth*. Series 60, Number 4, December. https://rb.gy/mnmevg

Aitken, A. (2019): "Measuring Welfare Beyond GDP", National Institute Economic Review No. 249 August. https://rb.gy/dpnvjd

Anderson, E., M. Jalles D'Orey, M. Duvendack, and L. Esposito (2017): "Does Government Spending affect Income Inequality? A Meta-Regression Analysis", *Journal of Economic Surveys*, 31(4), 961–987. http://t.ly/bGQHG

Awaworyi Churchill, S., Ugur, M., and Yew, S. L. (2017): "Does Government Size Affect Per-capita Income Growth? A Hierarchical Meta-regression Analysis." *Economic Record*, 93(300), 142-171. http://t.ly/tVJTJ

Bandrés, E. and Gadea, M. D. (2019): "Investigating casual relations between public spending and economic growth in Europe." *Revista de Economía Mundial*, 51, 51-78. http://t.ly/tWuNk

Barro, R. J. (1990): "Government spending in a simple model of endogenous growth." *Journal of Political Economy*, 98, S103-S125. http://t.ly/SvgYo

Barro, R. J. and Sala-i-Martin, X. (1992): "Public finance in models of economic growth." *Review of Economic Studies*, 59, 645-661. http://t.ly/uADdZ

Berg, A., Ostry, J. D., Tsangarides, C. G., and Yakhshilikov, Y. (2018): "Redistribution, inequality, and growth: new evidence." *Journal of Economic Growth*, 23, 259-305. http://t.ly/ecJ3n

Bergh, A. and Henrekson, M. (2011): "Government size and growth: a survey and interpretation of the evidence." *Journal of Economic Surveys*, 25 (5), 872-897. http://t.ly/l66IZ

Bergh, A., and Henrekson, M. (2016): "Government size and growth: A rejoinder." *Journal of Economic Surveys*, 30(2), 393-396. http://bit.ly/3s1433x

Berik, G. (2020): "Measuring what matters and guiding policy: An evaluation of the Genuine Progress Indicator." *International Labour Review*, Vol. 159, No. 1. http://t.ly/SGbxH

Bhardwaj, R. (2018): "Explaining Income Inequalities: A Cross Sectional-Study of OECD Countries". Mimeo. http://bit.ly/3IhGzgh

Byrne, S.; T. Conefrey and M. O'Grady (2021): "The Disconnection of GDP from Economic Activity Carried out in Ireland," Quarterly Bulletin – Q4 2021, Central Bank of Ireland. http://bit.ly/3LE3vXA

de la Fuente, A. (2003): "Convergence equations and income dynamics: the sources of OECD convergence, 1970-95." *Economica*, 70(280), 655-71. doi.org/10.1046/j.0013-0427.2003.00001.x

Doumbia, D. and T. Kinda (2019): "Reallocating Public Spending to Reduce Income Inequality: Can It Work?", IMF Working Paper, WP/19/88. https://rb.gy/ei67eo

El-Shagi, M. and Shao, L. (2019): "The impact of inequality and redistribution on growth", *Review of Income and Wealth*, 65 (2), 239-263. http://t.ly/x55s8

Fournier, J.M. y Johansson, A (2016). The effect of the size and the mix of public spending on growth and inequality. OECD Economics Department Working Papers, 1344. https://rb.gy/izj2ls

Gemmell, N., Kneller, R., and Sanz, I. (2011): "The timing and persistence of fiscal policy impacts on growth: Evidence from OECD countries." *The Economic Journal*, 121(550), F33-F58. http://t.ly/7vaYm

Gemmell, N., R. Kneller and I. Sanz (2016): "Does the composition of government expenditure matter for long-run GDP levels?" *Oxford Bulletin of Economics and Statistics*, 78(4), 522-547. http://bit.ly/3LFOSEy

Grier, K.B. and Tullock, G. (1989): "An empirical analysis of cross-national economic growth, 1951-80." *Jour-*

nal of Monetary Economics, 24, 259-276. http://t.ly/aLRV4

Hansson, P. and Henrekson, M. (1994): "A new framework for testing the effect of government spending on growth and productivity." *Public Choice*, 81, 381-404. http://t.ly/nECJR

Johansson, A. (2016): "Public Finance, Economic Growth and Inequality: A Survey of the Evidence", OECD Economics Department Working Papers, 1346. https://rb.gy/mehawm

Jones, C.I. and P. Klenow (2016): "Beyond GDP? Welfare across Countries and Time." *American Economic Review*, vol. 106(9), 2426-2457. http://bit.ly/3H2O09k

Jordà, Ò. (2005): "Estimation and inference of impulse responses by local projections." *American Economic Review*, 95(1), 161-182. http://t.ly/fQF4p

Landau, D. L. (1983): "Government expenditure and economic growth: a cross-country study." *Southern Economic Journal*, 49, 783-792. https://doi.org/10.2307/1058716

Pevcin, P. (2004): "Economic output and the optimal size of government." *Economic and Business Review*, 6(3), 213-227. http://t.ly/iBQBC

Saunders, P. (1985): "Public expenditure and economic performance in OECD countries." *Journal of Public Policy*, 5, 1-21. http://t.ly/kPmw6

Slemrod, J. (1995): "What do cross-country studies teach about government involvement, prosperity, and economic growth." *Brookings Papers on Economic Activity*, 2, 373-431. http://t.ly/UEGCh

Tanzi, V. and Zee, H. H. (1997). "Fiscal policy and long-run growth." *IMF Staff Papers*, 44 (2), 179-2019. http://t.ly/xxkw6

Woo, J. (2020): "Inequality, redistribution, and growth: new evidence on the trade-off between equality and efficiency", *Empirical Economics*, 58, 2667-2707. http://t.ly/meKj6

Appendix A Welfare and government size in OECD countries

In Figure 4 we show for each country in the sample the welfare index relative to the US in 2019 and public spending over GDP from 1960 to 2019.

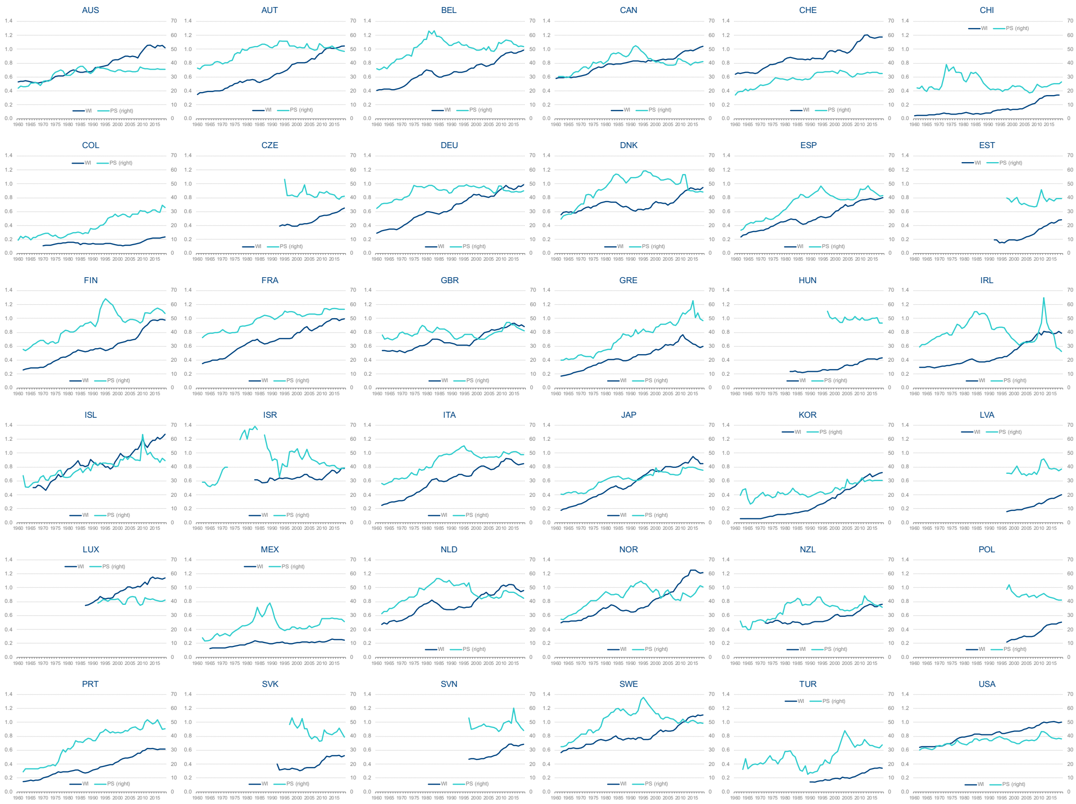


Figure 4: *relative welfare to the* $US_{2019} = 1.0$ *and public spending over GDP in percentage (right scale), OECD, 1960-2019. Source: see main text.*

Appendix B Additional country analysis: DOLS estimates

As an additional analysis of country heterogeneity, we have performed an individualized analysis country by country, using the Dynamic Ordinary Least Squares (DOLS) estimator, which allows for up to a maximum of four lags and four leading values of the independent variable, the public spending to GDP ratio. The number of lags and leads -between one and four- is selected using the Schwartz information criterion, so that a different number is allowed for each country. The estimated equation for $i = 1...N$ is as follows:

$$\Delta y_{i,t} = \alpha_i + \beta_i ps_{i,t} + \rho_i y_{i,t-1} + \sum_{j=-q}^{r} \delta_{ij} \Delta ps_{i,t+j} + \epsilon_{i,t} \tag{14}$$

where $y_{i,t} = ln(Y_{i,t})$, Y_{it} is *GDP* per capita or *WI* (the welfare index), $PS_{i,t}$ is the public spending to GDP ratio of country i in year t, q is the number of delays and r is the number of leads. Compared with the Mean Group estimation in Table 1, DOLS estimation allows for country heterogeneity also in the convergence coefficient (ρ) and for additional dynamics of *PS*, but at the cost of omitting the common time effects (δ_t) estimated in equations (4) and (5).

The DOLS technique has the advantage of correcting for serial autocorrelation, considering the possible endogeneity of the relationship and thus the causality of Δy_{it} to ps_{it}, and it is an alternative to other methods of estimating the cointegrating relationship, which prevents the non-stationarity of the series (Phillips and Hansen, 1990 and Stock and Watson, 1993). In this way, we take into account the dynamic effect of government spending on growth and welfare, which may be extended over time.

What we are looking for is not the Keynesian-type impact of spending on the rate of change of *GDP* per capita or the Welfare index over a short period of time, but a more long-lasting type of effect, in line with models that try to explain the determinants of long-run growth. It is quite possible that the effects of an increase in spending, even if larger in the early periods of its implementation, will end up maintaining a dynamic of persistence in the medium and long term. The research by Gemmell, Kneller and Sanz (2011) and Bandrés and Gadea (2019) using impulse-response functions provides evidence in this sense.

The results obtained with the DOLS estimation (see Table 6 in the Appendix) show much more heterogeneity in the estimates of the coefficients (β_i) of the government size than the panel specifications in Table 1. In the case of the filtered data in the equation for GDP per capita , the correlation of β_i between the DOLS and the Mean Group estimates of column (13) in Table 1 is negative (-0.35), with some outliers as Poland and Czech Republic. The average value of the DOLS estimates of β_i is -0.048, below the Mean Group estimate (-0.087). In the case of welfare, this correlation is positive but low (0.21), with the outliers of Latvia and Hungary. However, the average value of the DOLS estimates of β_i is positive (0.002), whereas in column (14) of Table 1 the Mean Group estimate is negative and equal to -0.042.

Table 6: *Estimation of DOLS regressions with business cycle*

countries/variables	constant	β	θ	constant	β	θ
		GDP growth			Welfare index	
AUS	10.0212	-0.1902	-3.6688	-1.6748	0.0856	1.0852
	(6.0720)	(-3.7487)	(-9.1016)	(-0.5923)	(0.9877)	(0.7004)
AUT	7.1528	-0.0920	-3.4493	5.6527	-0.0742	-0.2510
	(0.9538)	(-0.6398)	(-4.6650)	(1.2768)	(-0.8166)	(-0.2074)
BEL	11.0786	-0.1631	-3.1478	-1.0807	0.0478	-0.4417
	(6.1128)	(-4.7327)	(-3.8703)	(-0.1817)	(0.4293)	(-0.5886)
CAN	9.2511	-0.1437	-5.2815	1.8736	-0.0239	0.9522
	(4.0641)	(-2.6210)	(-7.0844)	(1.0038)	(-0.5536)	(1.0078)
CHL	34.5260	-1.3145	-5.0171	21.2820	-0.6934	-6.3656
	(5.9282)	(-5.2400)	(-5.3949)	(1.4242)	(-1.0497)	(-1.3566)
COL	3.1852	0.0221	-6.4721	-2.7878	0.1824	-3.1375
	(3.9091)	(0.5430)	(-6.5776)	(-0.7846)	(1.1434)	(-1.1658)
CZE	16.7007	-0.3087	-5.0789	-14.3839	0.3856	0.1677
	(1.9149)	(-1.5494)	(-4.5540)	(-1.5635)	(1.8363)	(0.1130)
DNK	2.1039	0.0047	-3.5551	5.1031	-0.0847	0.5121
	(1.0591)	(0.1129)	(-7.0659)	(1.7182)	(-1.3040)	(0.4897)
EST	46.4924	-1.1157	-4.8021	25.6591	-0.5681	0.7911
	(4.7719)	(-4.3302)	(-1.8594)	(2.7713)	(-2.1874)	(0.3884)
FIN	6.7332	-0.0713	-4.2917	5.6958	-0.0695	-0.3965
	(3.8024)	(-1.8892)	(-2.6715)	(2.1941)	(-1.2221)	(-0.3877)
FRA	15.9865	-0.2597	-3.8882	10.3942	-0.1711	0.9415
	(24.8009)	(-19.7724)	(-5.9693)	(3.3435)	(-2.8154)	(1.0369)
DEU	13.4032	-0.2389	-3.8749	0.2081	0.0329	-0.1373
	(4.5043)	(-3.7181)	(-7.0108)	(0.0327)	(0.2379)	(-0.1533)
GRC	9.3633	-0.1441	-5.6832	7.5256	-0.1384	-1.1005
	(6.2148)	(-3.8067)	(-5.1336)	(5.8698)	(-4.1763)	(-0.7920)
HUN	9.6169	-0.1309	-7.2043	24.5038	-0.4477	-0.1980
	(1.0833)	(-0.7062)	(-4.1803)	(1.2703)	(-1.1182)	(-0.1592)
ISL	8.6843	-0.1029	-6.3276	3.9744	-0.0311	-5.0380
	(4.1940)	(-2.0934)	(-7.7145)	(0.8736)	(-0.2885)	(-2.9454)
IRL	13.7161	-0.2033	-3.0387	4.8049	-0.0597	-2.2798
	(3.5697)	(-2.2159)	(-1.9667)	(1.4903)	(-0.7658)	(-1.5532)
ISR	3.9574	-0.0052	-1.2468	6.4603	-0.1198	1.6166
	(0.6455)	(-0.0370)	(-0.7412)	(0.7741)	(-0.6259)	(0.9713)
ITA	9.8494	-0.1644	-3.9035	9.5759	-0.1714	-0.6536
	(5.6508)	(-4.2633)	(-5.8024)	(3.4120)	(-2.8200)	(-0.5522)
JPN	19.9528	-0.4987	-3.2458	11.5903	-0.2881	-0.7337
	(10.2807)	(-8.3152)	(-3.7666)	(7.1461)	(-5.3172)	(-0.6763)
KOR	23.7337	-0.6817	-11.0410	11.9799	-0.2884	-6.8266
	(20.4721)	(-16.5904)	(-8.8177)	(3.1077)	(-1.7522)	(-3.3652)
LVA	50.0588	-1.1908	-5.9987	-7.1443	0.2207	-1.1520
	(3.5185)	(-3.2563)	(-3.7925)	(-0.5852)	(0.6593)	(-0.3621)
LUX	63.1779	-1.4453	-4.8473	17.5732	-0.3295	0.4088
	(2.8135)	(-2.6418)	(-4.5209)	(0.4054)	(-0.3125)	(0.1474)
MEX	6.8393	-0.1082	-6.7827	4.0139	-0.0910	-4.5505
	(3.7308)	(-1.4145)	(-4.9132)	(1.2209)	(-0.6775)	(-3.1633)
NLD	8.6302	-0.1173	-3.8379	8.0061	-0.1444	-0.5317
	(3.7002)	(-2.5515)	(-6.6733)	(1.9247)	(-1.6479)	(-0.4666)
NZL	11.8076	-0.2339	-5.0438	-3.9343	0.1345	-1.5759
	(3.9821)	(-3.0236)	(-7.2716)	(-0.9210)	(1.1725)	(-1.7721)
NOR	7.5479	-0.0954	-3.8895	2.4154	-0.0153	-0.6181
	(6.0379)	(-3.2369)	(-4.3022)	(1.0769)	(-0.3157)	(-0.2185)
POL	-	-	-	-	-	-
	-	-	-	-	-	-
PRT	8.3671	-0.1295	-3.7842	4.2334	-0.0240	-2.7185
	(6.1506)	(-3.6700)	(-4.8359)	(1.4409)	(-0.3714)	(-1.8338)
SVK	16.5756	-0.2854	-6.1224	23.8583	-0.5054	-0.6409
	(1.7022)	(-1.2678)	(-2.4545)	(1.9711)	(-1.7766)	(-0.1770)
SVN	11.9174	-0.1763	-6.3431	-6.6298	0.1920	-1.8381
	(1.4666)	(-1.0362)	(-2.1950)	(-0.8347)	(1.1017)	(-1.5898)
ESP	10.1103	-0.1741	-2.7264	4.6808	-0.0736	-0.1184
	(8.6014)	(-5.3634)	(-3.8622)	(2.4564)	(-1.2630)	(-0.0707)
SWE	6.3576	-0.0667	-4.0367	4.9371	-0.0729	0.6773
	(4.1025)	(-2.2083)	(-3.8658)	(2.7286)	(-2.0067)	(0.8821)
CHE	1.1874	0.0304	-3.4228	2.0758	-0.0330	0.0009
	(0.4690)	(0.3802)	(-3.8692)	(0.5960)	(-0.2869)	(0.0010)
TUR	4.2921	0.0559	-9.3043	4.0032	0.0146	-7.5096
	(3.6232)	(1.2975)	(-7.3561)	(1.2865)	(0.1390)	(-4.4183)
GBR	10.7368	-0.1990	-2.9961	6.3882	-0.1443	2.0754
	(4.8692)	(-3.5781)	(-5.4975)	(0.8079)	(-0.7351)	(1.4974)
USA	15.1317	-0.3178	-3.6478	1.8882	-0.0365	1.3994
	(7.6101)	(-5.8418)	(-7.8910)	(0.8160)	(-0.5864)	(1.9753)

Appendix C Effects on welfare growth components

From equation (2) it is clear that welfare growth (in relation to the United States) can be decomposed as the sum of four components. The first reflects the effects of variations in life expectancy (Δe). The second, per capita consumption growth ($\Delta \log c$). The third, changes in welfare due to leisure ($\Delta \nu(\ell)$). And finally, (minus) changes in income inequality ($-\Delta\frac{1}{2}\sigma^2$). Using trend components of these variables, we have:

$$\Delta \widetilde{wi}_{i,t} = \Delta \tilde{e}_{i,t} + \Delta \widetilde{\log c}_{i,t} + \Delta \widetilde{\nu(\ell)}_{i,t} - \Delta \frac{1}{2} \tilde{\sigma}^2_{i,t} \tag{15}$$

Therefore, to assess the effects of government size on welfare growth components, we can regress each component in equation (15) on the initial level of welfare ($\widetilde{wi}_{i,t-1}$) and on government size ($\widetilde{ps}_{i,t}$), controlling for country fixed effects and time dummies, as in Table 2, column (12). As shown by de la Fuente (2003) the sum of the estimated coefficients for the government size in these four regressions must be equal to the estimated coefficient of $\Delta \widetilde{wi}_{i,t}$.

In Table 7 we present estimates for the four components of welfare growth. To facilitate comparisons, in column (1) we reproduce the results of Table 2, column (12), for $\Delta \widetilde{wi}_{i,t}$. The results show clearly that the negative effects of government size on welfare growth are mainly driven by the growth of consumption per capita (column (3)), despite the fact that includes also public consumption. Consistently with the theory, as consumption growth falls the same happens with leisure growth (column (4)). Finally, the effect of government size on the inequality growth component is positive, although small.

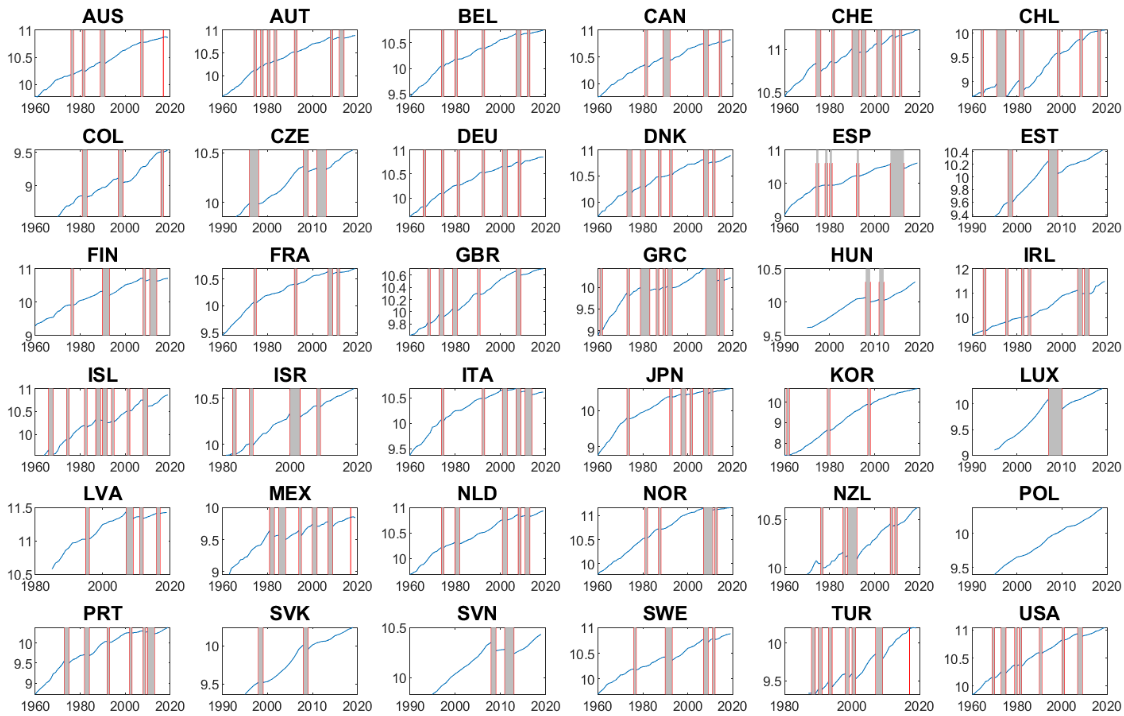
Table 7: *Panel estimates for welfare components*

	(1)	(2)	(3)	(4)	(5)
	$\Delta \widetilde{wi}$	$\Delta \tilde{e}$	$\Delta \widetilde{\log c}$	$\Delta \widetilde{\nu(\ell)}$	$\Delta \frac{1}{2} \tilde{\sigma}^2$
$\widetilde{wi}_{i,t-1}$	-0.005	0.002	-0.010	0.003	-0.001
	(2.42)	(2.39)	(5.21)	(8.00)	(2.96)
$\widetilde{ps}_{i,t}$	-0.026	0.001	-0.026	0.002	0.002
	(8.34)	(0.84)	(8.96)	(6.95)	(3.15)
$R^2 adj$	0.607	0.462	0.575	0.484	0.327
Countries	36	36	36	36	36
Observ.	1775	1775	1775	1775	1775

Appendix D Business cycle dating for different countries

In Figure 5 we have represented for our sample of OECD countries the log of GDP and the specific business cycle turning points estimated by applying the nonparametric dating procedure introduced by Bry and Boschan (1971).

Figure 5: *GDP (in logs) and business cycle turning points, OECD, 1960-2019*



Appendix E Cross-country estimates

The traditional approach of growth equations has used averages for the whole sample period and has estimated cross-country specification. In particular, in columns (1) and (2) of Table 8 we present the estimation results of the following equation:

$$\Delta y_{i,\bar{t}} = \theta y_{i,0} + \beta ps_{i,\bar{t}} + \alpha + \epsilon_i \tag{16}$$

where $y = gdp, wi$, therefore, $\Delta ln(Y_{i,t}) = \Delta y_{i,t}$ represents the average growth rate of *GDP* per capita or *WI* over the whole sample period; $y_{i,0}$ and $wi_{i,0}$ the initial values of the GDP per capita and welfare log for the first year available for each country, and $ps_{i,\bar{t}}$ the average of the government size log, again for the entire sample period.

Table 8: *Results of the cross-section estimation*

	(1)	(2)	(3)	(4)
	$\Delta gdp_{i,\bar{t}}$	$\Delta wi_{i,\bar{t}}$	$\Delta gdp_{i,\bar{t}}$	$\Delta wi_{i,\bar{t}}$
Constant	0.094	0.004	0.258	0.099
	(3.34)	(0.20)	(11.7)	(6.69)
$ps_{i,\bar{t}}$	0.012	0.019	-0.043	-0.011
	(2.20)	(3.10)	(5.99)	(2.30)
$gdp_{i,0}$	-0.012		-0.008	
	(4.67)		(5.89)	
$wi_{i,0}$		-0.015		-0.012
		(6.31)		(9.21)
Fixed effects			0.800	0.509
			(8.07)	(6.39)
Countries	36	36	36	36
R^2 *adj.*	0.350	0.660	0.816	0.790

Contrary to the robust negative effects of *ps* estimated in all previous sections, the results in columns (1) and (2) of Table 8 show now a surprising positive and significant effect for both *gdp* and welfare. However, in these specifications, we are controlling only for the initial levels of GDP per capita and welfare. We have seen that in the panel estimates the inclusion of country fixed effects contributed to uncover the negative effects of government size and their significance in all cases. Therefore, it could be the case that the omission of these country-time invariant effects could bias the coefficient of *ps* toward positive values.

To test this hypothesis, we have recovered the estimated country fixed effects (α_i) in columns (3) and (4) of Table 3, and we have included them as an additional control in the following cross-section equation:

$$\Delta y_{i,\bar{t}} = \theta y_{i,0} + \beta ps_{i,\bar{t}} + \gamma \alpha_i + \alpha + \epsilon_i \tag{17}$$

The results in columns (3) and (4) in Table 8 show that the inclusion of α_i does not affect to much the $\beta-$convergence coefficient of the initial GDP per capita and welfare, but makes again the coefficient of government size negative and statistically significant. As in the panel estimates, the negative effect of ps on welfare was smaller (-0.011) in absolute values than that estimated for GDP per capita (-0.043).

Appendix F Panel estimates with general government debt

Table 9: *Panel estimates with general government debt*

	(1)	(2)	(3)
	$\Delta\widetilde{wi}_{i,t}$	$\Delta\widetilde{wi}_{i,t}$	$\Delta\widetilde{wi}_{i,t}$
$\widetilde{wi}_{i,t-1}$	-0.005	-0.004	-0.008
	(2.49)	(2.35)	(4.61)
$\widetilde{ps}_{i,t}$	-0.042	-0.031	-0.028
	(10.3)	(5.56)	(5.16)
$\widetilde{ggd/gdp}_{i,t}$		-0.004	0.036
		(3.67)	(11.9)
$\widetilde{ggd/gdp}^{2}_{i,t}$			-0.006
			(14.9)
R2 adj	0.714	0.718	0.751
Countries	36	36	36
Observations	1289	1289	1289

Working Papers

2024

24/02 **J. Andrés, E. Bandrés, R. Doménech and M.D. Gadea:** Social Welfare and Government Size.

24/01 **J. Andrés, J.E. Boscá, R. Doménech and J. Ferri:** Transitioning to net-zero: macroeconomic implications and welfare assessment.

2023

23/08 **Ángel de la Fuente y Rafael Doménech:** Renta per cápita y productividad en la OCDE de 1960 a 2022.

23/07 **Ángel de la Fuente:** La evolución de la financiación de las comunidades autónomas de régimen común, 2002-2021.

23/06 **Ángel de la Fuente:** La liquidación de 2021 del sistema de financiación de las comunidades autónomas de régimen común.

23/05 **Ángel de la Fuente:** Las finanzas autonómicas en 2022 y entre 2003 y 2022.

23/04 **J. Andrés, J.M. Barrutiabengoa, J. Cubero and R. Doménech:** Social Welfare and the Social Cost of Carbon.

23/03 **Ángel de la Fuente y Pep Ruiz Aguirre:** Series largas de VAB y empleo regional por sectores, 1955-2021.Actualización de RegData-Sect hasta 2021.

23/02 **Gergely Buda, Vasco M. Carvalho, Giancarlo Corsetti, João B. Duarte, Stephen Hansen, Afonso S. Moura, Álvaro Ortiz, Tomasa Rodrigo, José V. Rodríguez Mora, Guilherme Alves da Silva:** Short and Variable Lags.

23/01 **Ángel de la Fuente:** Series largas de algunos agregados económicos y demográficos regionales: actualización de RegData hasta 2021.

CLICK HERE TO ACCESS THE WORKING DOCUMENTS PUBLISHED IN
Spanish and English

DISCLAIMER

The present document does not constitute an "Investment Recommendation", as defined in Regulation (EU) No 596/2014 of the European Parliament and of the Council of 16 April 2014 on market abuse ("MAR"). In particular, this document does not constitute "Investment Research" nor "Marketing Material", for the purposes of article 36 of the Regulation (EU) 2017/565 of 25 April 2016 supplementing Directive 2014/65/EU of the European Parliament and of the Council as regards organisational requirements and operating conditions for investment firms and defined terms for the purposes of that Directive (MIFID II).

Readers should be aware that under no circumstances should they base their investment decisions on the information contained in this document. Those persons or entities offering investment products to these potential investors are legally required to provide the information needed for them to take an appropriate investment decision.

This document has been prepared by BBVA Research Department. It is provided for information purposes only and expresses data or opinions regarding the date of issue of the report, prepared by BBVA or obtained from or based on sources we consider to be reliable, and have not been independently verified by BBVA. Therefore, BBVA offers no warranty, either express or implicit, regarding its accuracy, integrity or correctness.

This document and its contents are subject to changes without prior notice depending on variables such as the economic context or market fluctuations. BBVA is not responsible for updating these contents or for giving notice of such changes.

BBVA accepts no liability for any loss, direct or indirect, that may result from the use of this document or its contents.

This document and its contents do not constitute an offer, invitation or solicitation to purchase, divest or enter into any interest in financial assets or instruments. Neither shall this document nor its contents form the basis of any contract, commitment or decision of any kind.

ENQUIRIES TO:
BBVA Research: Azul Street. 4. La Vela Building – 4th and 5th floor. 28050 Madrid (Spain).
Tel.:+34 91 374 60 00 y +34 91 537 70 00 / Fax:+34 91 374 30 25
www.bbvaresearch.com