Is the welfare state relevant for economic growth? Evidence for Portugal[1]

Pedro Bação

Univ Coimbra, CeBER, Fac Econ, Portugal

Marta Simões

Univ Coimbra, CeBER, Fac Econ, Portugal

Abstract

In this paper we describe the evolution of welfare state spending in Portugal in the period 1980-2018, and consider its implications for economic growth. Overall, welfare spending in Portugal increased over this period as a percentage of GDP, but stagnated or even declined in recent years. Our empirical analysis attempts to quantify the contribution of welfare spending to economic growth in that period. We provide a comprehensive robustness check by means of specification-curve analysis. We conclude that the sign of the effect varies with the specification choices, but neither positive nor negative estimates are robust.

1. Introduction

The welfare state is a rather recent institution in Portugal. The current welfare system is a result of the 1974 political revolution and the first steps taken by the democratic regime, which were marked by the creation of the National Health System (NHS), the expansion of the public education system and the introduction of a public social security system for all citizens (Carolo and Pereirinha 2010). However, the 2007-08 financial crisis and the ensuing sovereign debt crisis reduced the fiscal capacity to provide a modern and effective welfare state (Gonzalez and Figueiredo 2015), raising fears that this retrenchment may hamper growth.

The main goals of the welfare state are to increase income equality and equality of opportunities (Van Lancker and Van den Heede 2019). Income equality and equality of opportunities are sometimes cited as determinants of economic growth. Thus, the welfare state may have the ability to influence economic growth (Atkinson 1996) and, in turn, economic growth may help the welfare state pursue its objectives (Tridico and Paternesi Meloni 2018). The literature on the relationship between the welfare state and growth is

[1] This is a pre-print of an article published in Comparative Economic Studies. The final authenticated version is available online at: https://doi.org/10.1057/s41294-020-00126-z .

however somewhat divided. Theoretical predictions point to two opposite sign effects of the welfare state on economic growth: a positive effect since the welfare state creates the conditions for economic agents to make decisions that promote growth, such as investing in human capital or taking more risks associated with innovation; and a negative effect due to the need to finance the welfare state through taxation, which in turn introduces distortions in economic decisions that are detrimental to growth, such as working less and with less effort and reducing savings and thus investment.

This paper investigates whether and how the welfare state impacts growth using 1980-2018 data for Portugal. The empirical approach makes use of a vector autoregression (VAR) model inspired by a Cobb-Douglas aggregate production function that includes physical and human capital stocks, total factor productivity and a measure of welfare state effort (social spending). Thus, we allow the influence of welfare spending on economic growth to occur through both factor accumulation and productivity. In addition, we allow for the possibility that the effects may differ across different components of social spending. This disaggregated analysis might have important implications for the design of more effective economic and social policies that result in a more inclusive society. Since the researcher has a number of degrees of freedom when setting up the empirical analysis (in this case based on a VAR model) and there seems to be a tendency to report only the desired (statistically significant) outcome, in this study we use specification-curve analysis to assess the robustness of the results. This technique involves running all the reasonable/relevant regressions and evaluating the results against a benchmark obtained by simulating the data under the null hypothesis of no effect (Simonsohn et al. 2015). We apply this analysis twice: one from the point of view of a researcher who finds a positive effect of welfare spending on economic growth, and the other from the point of view of an estimated negative effect.

The paper is structured as follows. Section 2 introduces the Portuguese case. Section 3 briefly reviews the theoretical and empirical literatures that debate whether the welfare state is relevant for economic growth. The data and the empirical methods employed in this study are described in section 4. Section 5 presents and discusses the findings of the empirical analyses. Section 6 offers concluding remarks.

2. The Portuguese context: some facts and figures

The period under analysis in this study, 1980-2018, coincides with the early years and deepening of the Portuguese European integration process. Just before, Portugal had gone through a political revolution (in 1974) and was taking the first steps as a democracy after 48 years of dictatorship. By joining the European Economic Community in 1986 (Portugal applied to become a member in as early as 1977), Portugal became officially committed to the European integration process, important in supporting the transition to a developed democracy and achieving higher standards of living. European integration resulted also in the adoption of policy measures aimed at promoting convergence to the European Social Model (ESM), a specific model of economic progress and social cohesion characterizing most European countries. The ESM aims at combining economic growth with a more equal society. In the ESM state intervention is closely linked to the promotion of welfare for all citizens.

Based on data for social expenditure for the period 1938-2003, Carolo and Pereirinha (2010) identify three phases as the most significant in the process leading up to the current welfare state system in Portugal. The first phase coincides with the years of dictatorship, known as *Estado Novo* (New State), spanning from 1938 until the political revolution of 1974. In this period there was some growth and consolidation of welfare provision regarding the major social risks for those working, reflected in a steady growth of social expenditure. The growth of social expenditure was initially due to higher generosity, but towards the end of the period it was associated with wider coverage, incorporating for instance the rural population. According to their data, social expenditure as a percentage of GDP went from 0.38% in 1938 to 6.5% in 1974. The second phase goes from 1974 until the mid-80s, a period when Portugal introduced a social security system for all citizens, including those that had not paid any contributions. In 1986, when Portugal joined the European Union, social expenditure represented 10.82% of GDP. However, according to the authors, p. 495 "(…) in terms of social expenditure it was a period of relative containment in comparison with the previous phase before the transition to democracy. Thus, this result does not provide evidence for the popularly held splurging myth of the 1974 revolution and its aftermath, despite it being a period of egalitarian and universalistic social policies." It is only in the third phase (1986-2003) that there is a clear expansion in social expenditure, a phase characterized also by convergence to the average EU social spending ratios, due mostly to old age pensions but also to health care. In fact,

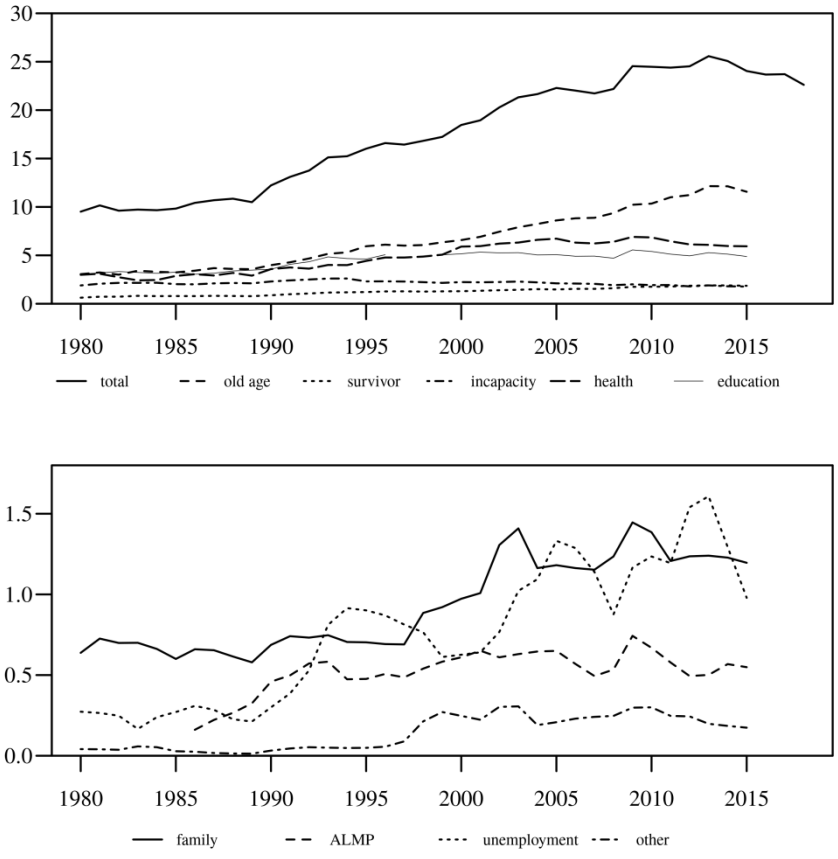
social expenditure represented 20.5% of GDP in 2003, the last year for which the authors provide data. The latest figures from the OECD Social Expenditure database (SOCX) indicate that this ratio stood at 22.6% in 2018. Silva et al. (2014) also pose that it was only with the restoration of democracy and, in particular, with accession to the EU that Portugal built a system of social protection effectively able to protect against the many social risks and coherent in terms of its design. Glatzer (2012) highlights the rapid rise in the number of people receiving social transfers, which expanded by more than 50 times since the 60s until today, while Marinheiro (2014) shows that between 1990 and 2010 social protection, health and education were the public expenditure components that recorded the highest increases.

More recently, according to Gonzalez and Figueiredo (2015), there is a new phase in the implementation by Portugal of the ESM. According to the authors, demand for social protection has increased strongly after the 2000s due to structural changes in the Portuguese economy (adoption of the euro, the accession of China to the World Trade Organization, the EU enlargement to Eastern countries), the Great Recession and the austerity measures associated with the financial assistance program following the sovereign debt crisis. However, austerity led to changes in the welfare state structure in order to promote its efficiency and sustainability that in turn resulted in a deterioration of social protection. In these authors' opinion (p. 332), "There is a high risk that blind austerity measures will generate negative dynamic effects in the long term, penalizing growth potential (many people will not return to the labour market) and leading families to inefficient allocation of resources (for example not investing in education)." For instance, in spite of the sharp increase in unemployment following the crises, social expenditure as a percentage of GDP has remained basically unchanged from 2009 until 2016, going from 24.6% to 24.1% of GDP (SOCX), although in 2013–the year after real GDP per capita recorded a drop of 3.6% and unemployment stood at 16.2%–it reached 25.6%. As mentioned before, the 2018 figure is 22.6%, only slightly higher than the 2003 figure, 21.3% in the SOCX and 20.5% in Carolo and Pereirinha (2010).

Figure 1 contains data for public social expenditure as a percentage of GDP, total and by spending categories, from the SOCX database for Portugal over the period 1980-2018, whenever data is available. This data confirms the above description: until 1986 spending ratios are relatively stable but from then onwards, until around the year 2009, they record a clear and steady increase; the period of the Great Recession, sovereign debt crisis and

bailout of the Portuguese economy coincides with stagnation and even decline in total spending as a percentage of GDP, as well as for most spending categories with the exception of old age pensions and unemployment benefits. The more recent years of recovery of the Portuguese economy, from 2014 onwards, have not been accompanied by an inversion of the previous negative trend recorded by social expenditure as a percentage of GDP.

Figure 1: Public Social Expenditure as a percentage of GDP, total and by category, Portugal 1980-2018



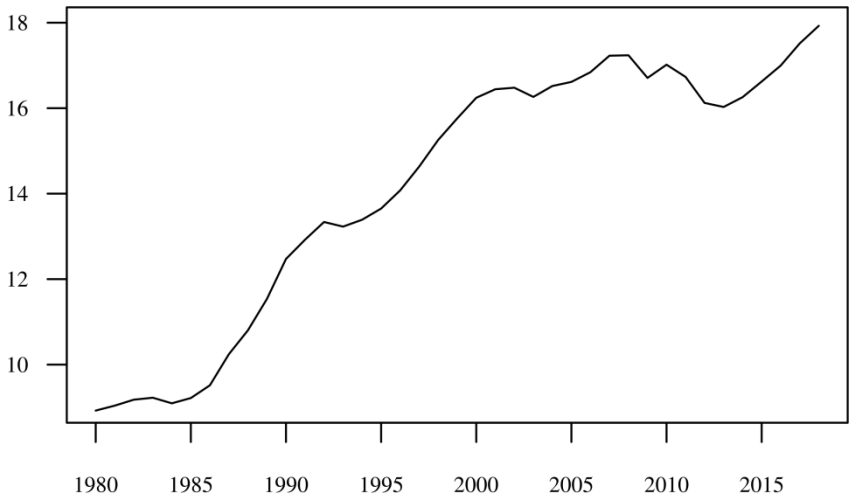
Notes: Total public social expenditure does not include education. Social expenditure on housing is not included due to the low ratios recorded, ranging from a minimum of 0% to a maximum of 0.006%. "ALMP": Active Labour Market Policies.
Source: OECD Social Expenditures database

The 1980s were also a period of great expectations at the economic level. Accession to the EU was accompanied by a growth acceleration of the Portuguese economy relative to the previous decade, 1974-1985, a period during which, following the political turmoil and the concomitant economic hardships, the Portuguese economy became almost

stagnant and balance of payments crises required two IMF interventions, in 1978-79 and again in 1983-85. The first years of European integration were quite favourable for Portugal in terms of output growth and thus created the conditions for political support to joining the Economic and Monetary Union. This became effective as of 1 January 1999, after a decade of preparations. Portugal experienced rapid economic growth in the years that preceded the launch of the euro (between 1995 and 1999-2000). Since then, however, the Portuguese economy experienced a very sluggish rate of economic growth, a scenario aggravated by the 2007-08 financial crisis but foremost by the subsequent sovereign debt crisis that resulted in the third IMF bailout (this time joined by the ECB and the European Commission) that lasted from May 2011 until June 2014, albeit by mid-2013 the Portuguese economy initiated a gradual recovery. This recent evolution of the Portuguese economy has been strongly affected by the austerity policies recommended by the Troika (IMF, ECB and European Commission) due to the need to reduce the deficit and public debt. This in turn has brought the welfare state to the forefront of the debate on Government retrenchment. At a time of unprecedented unemployment rates, access to unemployment insurance was restricted resulting in a strong reduction in the number of the unemployed that could claim unemployment benefits. At that same time there were important restrictions in the access to a modern and effective national health system, increases in the number of students per teacher, changes in the criteria to claim family allowances among others (Gonzalez and Figueiredo 2015).

Figure 2 presents data on real GDP per capita for Portugal over the period 1980-2018 and Table 1 contains the respective average annual growth rates, overall and for different sub-periods. Over the period 1980-2018, real GDP per capita recorded an average annual growth rate of 1.9% that corresponds to quite different performances throughout the period. Immediately before accession to the EU (1980-85), income levels remained basically the same with real GDP per capita growing on average 0.6% a year. In the five years that followed (1985-90) this growth rate increased by a factor of ten, slowed down to 1.8% in the 1990-95 period and again picked up to 3.5% a year over the next 5-years, 1995-2000. Over the course of the new millennium, stagnation was the dominant feature in terms of real GDP per capita, with the growth rate ranging from -0.5% to +0.5% during the three 5-years sub-periods from 2000 until 2015. In the last three years of the analysis, 2015-18, however, real GDP per capita grew at an annual average growth rate of 2.5% and for the first time in 2017 real GDP per capita surpassed the pre-crisis level.

Figure 2: Real GDP per capita, 2011 prices (1000 EUR), Portugal 1980-2018



Source: AMECO database, May 2019 release

Table 1: Average annual growth rate of Real GDP per capita, 2011 prices, Portugal 1980-2018

	Overall		**10-years sub-periods**			
Period	**1980-2018**		**1980-90**	**1990-2000**	**2000-10**	**2010-18**
Growth rate	1.9%		3.4%	2.7%	0.5%	0.7%
	5-years sub-periods					
Period	**1980-85**	**1985-90**	**1990-95**	**1995-00**	**2000-05**	**2005-10**
Growth rate	0.6%	6.2%	1.8%	3.5%	0.5%	0.5%
Period	**2010-15**	**2105-18**				
Growth rate	-0.5%	2.5%				

Source: Owns calculation with data from AMECO database, May 2019 release

In late 2015, a new government resulting from an understanding between left-wing parties replaced the government that had been responsible for the implementation of the austerity measures. The economy continued the recovery started in mid-2013, with Portugal recording higher output growth rates, lower unemployment and lower public deficits. The new government introduced some changes in spending and in the structure of the Portuguese welfare system, but a fundamental question arises. Will the rescaling and reorganization of the Portuguese welfare system aggravate further the already dismal long-run growth prospects in a country that ranks as one of the most unequal in Europe and presents still relatively low educational attainment levels? Understanding the different mechanisms that connect the two dimensions and the circumstances under which

they operate is important for sound social policy design and implementation, allowing for a better understanding of how to target social support so that social policies not only improve social cohesion but also sustain long-term growth in Portugal.

From the perspective of social cohesion, Table 2 presents data on selected social indicators for Portugal for some of the years covered in this study. The main message from these data is that progress has been modest when one considers the expected outcomes from the expansion of the welfare state in Portugal in terms of income distribution and the poverty rate: over the period under analysis the Gini index of income distribution remained basically unchanged (1980 – 33.5; 2018 – 33.7), the at-risk-of-poverty rate recorded a slight reduction (2000 – 20%; 2018 – 18.3%) and the top 10% income share rose from 27.5% in 1980 to 32.7% in 2018, although in 2000 it stood at 34.2%. In summary, poverty and inequality remain high. On the other hand, until 2010 improvements in the health status of the population have been quite substantial, due to the national health system created in 1979, which is presently considered at risk on account of under-investment (Gonzalez and Figueiredo 2015). From 1980 until 2010, life expectancy at birth increased from 71.4 years to 80 years and infant mortality decreased from 24.3 deaths per 1000 live births to 2.5, although in 2018 it stood at 3.2. Education is another major area of state intervention in Portugal but one where the accomplishments are less striking. The increase in public education expenditure has contributed to higher educational attainment levels of the Portuguese population, with average years of schooling increasing from 4.65 years in 1980 to 7.52 years in 2010. Nevertheless, in 2018 only around 50% of the population aged 15-64 years had completed upper secondary, post-secondary and tertiary education. Another area of concern is that of long-term unemployment and youth unemployment. Although unemployment for most of the period has been quite low, it started to climb in the mid-2000s and this rise was accompanied by an increase in both the long-term unemployment rate and the youth unemployment rate. In 2018, long-term unemployment represented almost 50% of the unemployed and youth unemployment represented 20.3% of the youth labour force.

Table 2: Social Indicators for Portugal, 1980-2018

	1980	1990	2000	2010	2018*
Income distribution					
Gini index, disposable income	33.5	33.9	33.7	33.7	33.7
Pre-tax national income, Top 10% share (%)	27.5	31.0	34.2	32.5	32.7
Poverty					
At-risk-of-poverty rate after social transfers (%)	---	---	20.0	18.0	18.3
Health status					
Life expectancy and birth (years)	71.4	74.1	76.9	80	81.2
Infant mortality (Deaths per 1 000 live births)	24.3	10.9	5.5	2.5	3.2
Educational attainment					
Population 15-64 years with tertiary education (%)	---	8.5	7.5 [a]	13.9	22.5
Population 15-64 years with upper secondary and post-secondary non tertiary education (%)	---	11.8	13.5 [a]	18.8	27.6
Average years of total schooling for the population aged 15 and above	4.6	6.0	7.4	7.5	---
Unemployment					
Long-term unemployment rate (%)	---	---	1.7	5.7	3.1
Long-term unemployment rate, % of unemployed	---	44.9	42.2	52.2	50.0
Youth unemployment rate, % of youth labour force	---	---	8.9	22.7	20.3

Notes: * data for 2018 or latest year available, usually 2016; [a] break in the time series;
Source: Own calculation with data from the Eurostat, the Barro and Lee education dataset, the OECD Statistics, the Standardized World Income Inequality database and Pordata.

3. Controversies on the relationship between the welfare state and economic growth: theoretical arguments and recent findings

Economic growth, the steady increase of output in the long run, is the immediate result of either higher accumulation of factors of production or improvements in efficiency/productivity of those factors, or, more realistically, both. These direct sources of growth are in turn determined by more fundamental sources, i.e., those features that have an important influence on a country's ability to accumulate inputs and become more productive and efficient, such as the institutional arrangements that frame economic

activity. Unlike the proximate determinants of growth, there is no consensus as to the fundamental sources of growth, an issue also known as open-endedness of growth theory, implying that different growth factors highlighted by different theories are compatible with one another (Brock and Durlauf 2001). The welfare state comes under the classification of fundamental sources, in particular institutions. However, both at the theoretical and at the empirical level, the studies that investigate the impact of the welfare state on economic growth have reached no definite conclusions on the sign, transmission mechanisms and direction of causality of the relationship.

The literature suggests that the welfare state affects aggregate output behaviour but there is no consensus as to the sign of this effect. Critics of welfare intervention by the state pose that the financing side of welfare provision introduces distortions in economic decisions that are detrimental to growth. The traditional view is that the taxes needed to finance social expenditure sap economic efficiency and thus growth. Taxes affect the behaviour of individuals because, for instance, the work-leisure and the consumption-savings decisions become different relative to a situation with no taxes. Higher income taxes encourage people to work less and spend more time in leisure (although this depends on the relative importance of substitution vs. income effects), thus the total amount of output decreases. Through its impact on the fraction of income being saved, taxation influences capital accumulation and in this way growth. Higher income taxes lower the net returns from savings and, if savings rates are higher for richer individuals, a more progressive tax system has even more adverse effects on savings. Investment becomes lower than in the absence of taxes and, since there is less capital accumulation, a negative impact on economic growth emerges. Income taxes also lower the gains from education and result in less human capital accumulation, again hampering growth. Social contributions increase labour costs and in a context of economic globalization reduce competitiveness and growth. But different financing mechanisms for supporting social protection systems might have different outcomes in terms of economic growth since some modes of financing are less distortionary than others. For instance, Arnold et al. (2011) conclude that corporate income taxes have the strongest negative effects on growth, followed by personal income taxes, while consumption taxes have less adverse effects, and finally property taxes appear to produce the least important effect. Additionally, the various ways the financial resources that support the welfare state are spent result in different channels of influence relative to economic growth. A number of

social policies have the potential to produce a positive growth effect. Education and health expenditure promote the accumulation of human capital, overcoming market failures that do not allow talented individuals to have access to education/healthcare, and thus enabling countries to fulfil their human capital potential and grow faster (Benos and Zotou 2014; Bloom et al. 2018); family support schemes, such as childcare or long-term care programs, prevent people from quitting their jobs (or encourage them to work) because they have to raise their children or take care of dependent family members; the funds used to pay for pensions through the public pension system can be used to finance investment projects; unemployment compensation can help unemployed people find a better job by enabling jobseekers to wait for a job that matches their skills, and may even allow them to invest in their human capital; overall, social transfers reduce income inequality, which according to some authors has a positive influence on growth (Galor and Moav 2004; Cingano 2014; Berg et al. 2018). The foregoing arguments support the view that the welfare state leads to improved macroeconomic performance in the long run in the form of faster growth. In summary, the sign of the impact of the welfare state on economic growth depends on how the financial resources that support the welfare state are spent.

It is thus not surprising that empirical studies on the relationship between welfare state effort and economic growth are somewhat divided. This lack of consensus from the existing empirical analysis is made clear by recent meta-regression analyses by Awaworyi Churchill and co-authors, aimed at synthesizing the evidence and accounting for the sources of heterogeneity among reported findings on the link between main components of social expenditure and economic growth (Awaworyi Churchill et al. 2017; Awaworyi Churchill and Yew 2017; Awaworyi Churchill et al. 2015). The main takeaway from these meta-regression analyses is that a deeper understanding of the welfare state-economic growth nexus demands disaggregating the former according to its different categories. Furthermore, research on the relationship between particular components of the welfare state and growth is more likely to produce policy-relevant findings than studies that focus on overall measures of welfare effort.

Awaworyi Churchill and Yew (2017) focus on 149 estimated coefficients for the relationship between government transfers (GTRAN) and growth, retrieved from 23 different studies. These government transfers refer to different types of social expenditure, in particular social security, pensions and unemployment benefits, not including education nor health. A first inspection of these coefficients leads the authors

(p. 272) to conclude that "(…) many of the empirical studies found a statistically insignificant effect of GTRAN on economic growth, while others found significant, either positive or negative, estimates. For studies that reported significant estimates, there was considerable disagreement concerning the size of the effect of GTRAN on economic growth." They then go on to apply meta-regression analysis to statistically examine these differences in results, concluding that the dominant evidence, indicating the existence of a negative relationship, is due to publication bias. However, for developed countries the negative sign remains beyond publication bias. The authors also found robust evidence of a negative growth impact of unemployment benefits, while for social security transfers the effect is positive. Other characteristics of the empirical approach that influence the results are: the use of more recent data, in particular from the 90s and the 2000s, that leads to less adverse effects of GTRAN on growth; time series data results in more intense negative effects while cross-section data leads to results that are not statistically significant.

In Awaworyi Churchill et al. (2017) the focus is on education expenditure, for which they use 237 estimated coefficients from 29 selected empirical studies on the relationship between education expenditure and economic growth. The main conclusion is that the relationship is positive but statistically significant only for developed countries. Additionally, study characteristics such as data type, period of data averaging and variables included in the econometric specification have an impact on the size of the effect found. Studies that use data from 1990 and beyond tend to report less positive effects of education expenditure on growth. In particular, studies that adopt specifications based on an endogenous growth model report less positive effects of education expenditure on growth; studies that use data averaging for periods equal to or greater than 5 years tend to report more positive effects of education expenditure on growth and the same applies to studies that use cross-section data (as opposed to panel data); studies that control for population growth rate, political instability, tax and government quality (as opposed to those that do not) report less positive effects, while those that control for life expectancy and inflation tend to report more positive effects.

Awaworyi Churchill et al. (2015) synthetize the evidence regarding the impact of health expenditures on economic growth using meta-regression analysis based on the results from 12 studies with a total of 69 estimates, of which 37.58% are statistically insignificant estimates, 56.52% are negative estimates, and only four estimates,

representing 5.80% of the total, are positive. The meta-analysis tests performed indicate that there is a genuine negative effect of health expenditures on growth. This effect is more adverse when: the datasets used are older (end before the 2000s); the samples consist of OECD countries only; specifications are based on predictions from endogenous growth models; data is averaged over periods equal to or longer than five years; and private investment is included as an explanatory variable.

The opposing theoretical arguments on the impact of the welfare state on economic growth and the varied evidence thus claim for more empirical research on the subject, in particular country-specific analysis since it does not seem possible to derive one-size-fits-all policy implications given the variety of welfare state models adopted by different countries. The theory and previous evidence provide a guide to interpret our empirical approach aimed at verifying and quantifying the importance of the various channels through which welfare effort might influence economic growth. As the results from meta-analysis studies show it is important to control for the robustness of the results found. Since one possible source of the heterogeneity of the results found in previous studies concerns the options made by researchers in terms of processing the data, when different approaches to running the analysis can be justified, we apply the specification-curve analysis proposed by Simonsohn et al. (2015) to mitigate this problem.

4. Empirical modelling and estimation strategy

To determine the sign of the relationship between the welfare state and economic growth it is common to use a reduced form equation that relates output growth to initial income per capita, a variable capturing the welfare state, and a vector of control variables. A common problem faced by empirical growth studies is that of model uncertainty due to the large number of features with the potential to influence economic growth; different empirical models lead to different conclusions concerning the same growth determinants. In order to overcome to some extent this problem we investigate the impact of the welfare state on growth through the more consensual proximate determinants, factor accumulation and productivity. Therefore, the sign and magnitude of the welfare state growth impact is estimated through the use of a VAR model defined according to a standard Cobb-Douglas aggregate production function with human capital, as in Hall and Jones (1999).

Hall and Jones (1999) assume that output, *Y*, is produced according to the following production function:

$$Y_t=K_t^{\alpha}(A_tH_t)^{1-\alpha} \qquad (1)$$

where *K* is the stock of physical capital, *H* is the amount of human capital-augmented labour used in production, *A* is total factor productivity and α is the capital share.

Output per worker, *y*, can thus be written as:

$$y_t=k_t^{\alpha}(A_th_t)^{1-\alpha} \qquad (2)$$

where *k* is the stock of physical capital per worker and *h* is the amount of human capital per worker. Thus equation (2) can be rewritten as:

$$y_t=(K_t/Y_t)^{\alpha/(1-\alpha)}A_th_t \qquad (3)$$

Taking logs and first differences of both sides yields:

$$g_t^y= [\alpha/(1-\alpha)]g_t^{K/Y}+g_t^A+g_t^h \qquad (4)$$

where the g's are the log-growth rates of the variables in superscript. From this point of view, what matters for GDP growth is the growth of productivity and of factor inputs (proximate sources of growth). Our approach is based on the analysis of the relation between the welfare state and these variables. In other words, does the welfare state influence any of the determinants of GDP growth? To answer this question, we resort to a VAR model where we include the growth rates of total factor productivity, the capital-output ratio and human capital, alongside a variable related to welfare state effort. The growth rate of total factor productivity must be estimated; we do so by setting α, the capital share, to one third, as is customary.

The general form of the VAR model of order *p* that we use to analyse the relationship between the welfare state and real GDP per worker growth can thus be written as:

$$X_t = \beta_0 + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_p X_{t-p} + \varepsilon_t \qquad (5)$$

where the vector X contains the variables under analysis (the capital-output ratio, human capital, total factor productivity, all in log-growth rates) and a measure of social expenditure. Details on the variables used and respective sources are provided in Table 3. Figures 3 and 4 show the behaviour of the aggregate production function variables in levels and in log growth rates, respectively. Note that the estimated VAR models are

stationary, which suggests that the series used in the estimations may also be stationary. Therefore we did not perform stationarity/unit root tests for the individual series.

Table 3. Variables and sources

Variable	Definition	Units	Source
Welfare effort	Public social expenditure as a percentage of GDP (total and by spending category: old age, survivors, incapacity-related benefits, health, family, active labour market policies, unemployment, housing, and other social policy areas)	Percentage	OECD Social expenditures database accessed on 13-05-2019
Output	Gross domestic product at 2010 reference levels (AMECO notation OVGD)	Mrd national currency (EUR)	AMECO May 2019 release
Capital stock	Net capital stock at 2010 prices: total economy (AMECO notation OKND)	Mrd national currency (EUR)	AMECO May 2019 release
Hours worked	Total annual hours worked (AMECO notation NLHT)	Millions	AMECO May 2019 release
Human capital	Average years of schooling of the population aged 25 and above.	Years	Teixeira & Loureiro (2019)
TFP	Computed as the residual of the aggregate Cobb-Douglas production function setting $\alpha=1/3$.	Index (2010=100)	Own calculations with data from AMECO.

Figure 3. Aggregate production function variables in levels

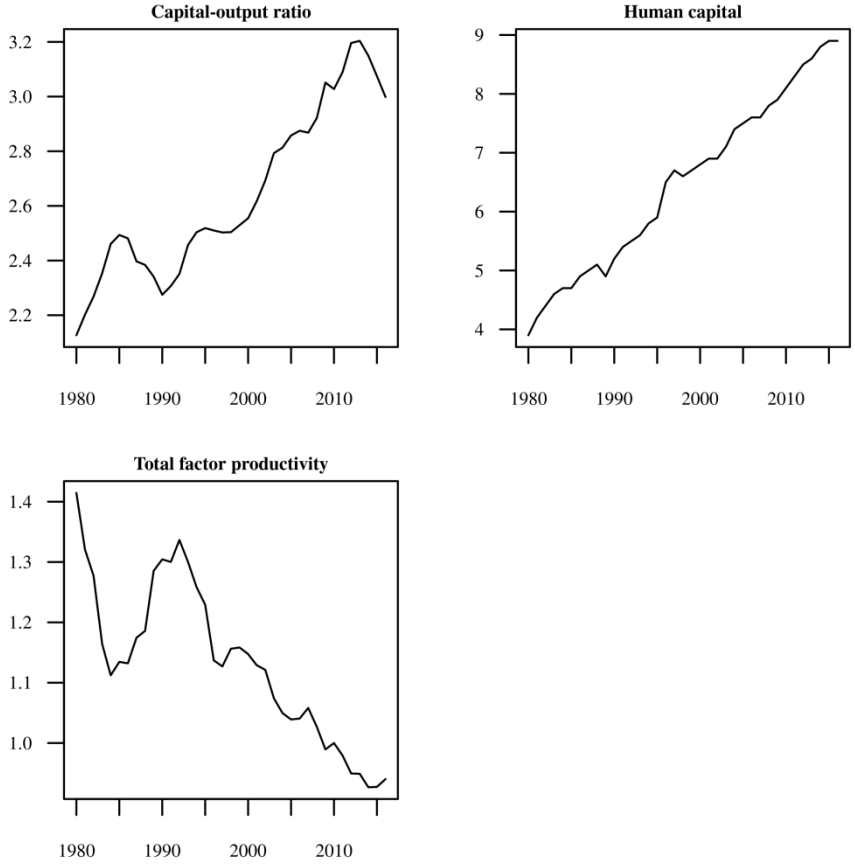
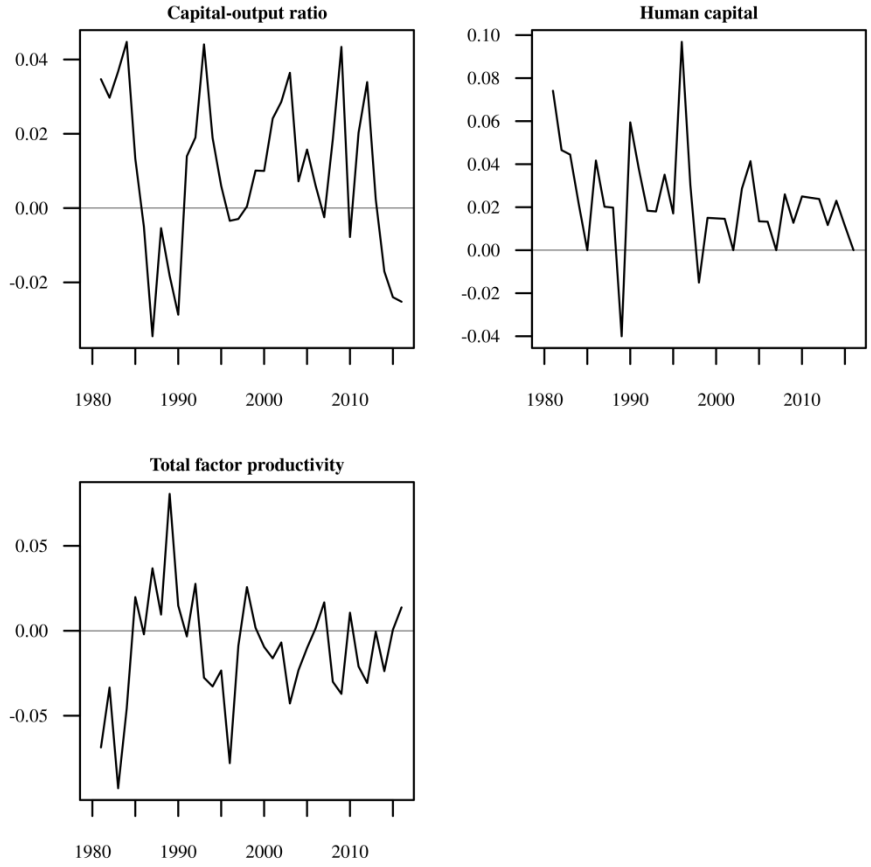


Figure 4. Aggregate production function variables in annual log growth rates



As discussed in section 2, public expenditure associated with the notion of "welfare state" comprises many different elements. Different researchers (or the same research at different times) may want to analyse the relation between different components of welfare-state expenditure and economic growth. In this context, one may say that researchers have a degree of freedom in the choice of the measure of the welfare state. However, there are many other choices to be made before finally getting and analysing the estimates of an econometric model such as that in equation (5). More immediate choices concern whether to take the logarithm of the measure of public expenditure and whether to use the level or the first difference of the series. Equation (5) hints at the existence of another degree of freedom: the number of lags in the VAR model (parameter p). Other choices may emerge as a result of the definition of the parameter of interest. In this paper we define the parameter of interest (the quantity to be estimated in order to gauge the impact of welfare-state expenditure on economic growth) to be the long-term impact of a permanent shock to the level of welfare-state expenditure (as a percentage of

GDP) on the level of output (in logarithms). We approximate the long-term impact by examining the impact after 100 periods. The impact will be measured by the impulse-response function, aggregated according to equation (4) to obtain the impact on output. We use the standard Cholesky decomposition to estimate the impulse-response functions. Since the results obtained using the Cholesky decomposition depend on the ordering of the variables in the VAR model (see Christiano et al. 1999) and there is no obvious ordering, there is an additional degree of freedom available to the researcher. Nevertheless, we will restrict the orderings so that $g_t^{K/Y}$ always comes last, that is to say, it responds faster to shocks than the other variables in the model. The reason for this assumption is that this variable depends on output and output depends on the other variables in the model (except, possibly, welfare-state spending, which is the hypothesis we wish to test).

Given the number of choices that the researcher must make, it is possible that two researchers working on the same dataset but using two different sets of modelling choices will arrive at different conclusions. Often, each researcher will present the main econometric results alongside robustness checks. However, the robustness checks will typically correspond to only a small fraction of all the choices available to the researcher. Inspired by the "reproducibility crisis" in psychology, Steegen et al. (2016) suggested instead the use of a "multiverse analysis". This consists in identifying all the choices made in processing the data and redoing the statistical analysis under all the possible alternatives (which produces a "data multiverse"). The result will be a set of results obtained under different data processing choices. One can then check whether the result reported by the researcher is specific to the choices made by that researcher or whether it is indeed robust.

Simonsohn et al. (2015) go even further and propose a "specification-curve analysis". The first step in this analysis is similar to the multiverse analysis: estimate the parameter of interest under all reasonable alternative specifications. By doing this one obtains a curve with the estimates derived from the alternative specifications. In the second step, Simonsohn et al. (2015) suggest that one builds simulated datasets in which the null hypothesis concerning the parameter of interest is true. The third step is to estimate the parameter of interest using these simulated datasets under all reasonable alternative specifications. At the end of this step, one has many curves of estimates of the parameter of interest, one curve for each simulation of the dataset under the null hypothesis. The

issue then becomes whether, under the null hypothesis, observing a curve such as the one estimated in the first step is likely or not. This likelihood is assessed via the computation of certain statistics for the curve estimated using the actual dataset, and calculation of the share of curves estimated using the simulated datasets that report statistics of similar magnitude. If the values of the statistics corresponding to the observed curve are also frequent in the curves derived from the simulated datasets, then probably the null hypothesis is true.

In this text we apply the specification-curve analysis to assess the robustness of results concerning the impact of welfare-state spending on economic growth in the context of the VAR model described above. Our procedure for constructing the simulated datasets under the null hypothesis of no effect of welfare-spending on economic growth is the following. The VAR model is estimated on the actual dataset (which covers the period 1980-2018 for Portugal) using the preferred (or reference) specification. From this estimation we obtain an estimate of the variance-covariance matrix of the residuals, say *V*. We apply the Cholesky decomposition to this matrix to obtain a matrix *S* such that $SS'=V$. With *S* and the residuals of the VAR model we can compute an estimate of the underlying structural shocks. We create a modified *S* matrix such that the structural shock associated with welfare spending does not affect the other variables contemporaneously. We use the modified *S* matrix to compute modified residuals which conform to that restriction. We also modify the estimated matrix of coefficients of the VAR model – the betas in equation (5) – so that lagged welfare spending does not affect the other variables. Note that we are in effect imposing no short-run, as well as no long-run, impact of welfare spending on output. Allowing for short-run effects while ruling out long-run effects would require complex procedures that would make harder to apply the large-scale approach involved in the specification-curve analysis.

Given the modified residuals and the modified matrix of coefficients, we generate alternative datasets (in which the null of no impact of welfare spending on growth is imposed) by bootstrapping the modified residuals and using them to feed the VAR model. As in Simonsohn et al. (2015), we construct 500 alternative datasets. The modelling choices are listed in Table 4. They give rise to a total of 4064 alternative specifications, which are applied to each of the 500 simulated datasets, besides being applied to the original dataset.

Table 4: Modelling choices

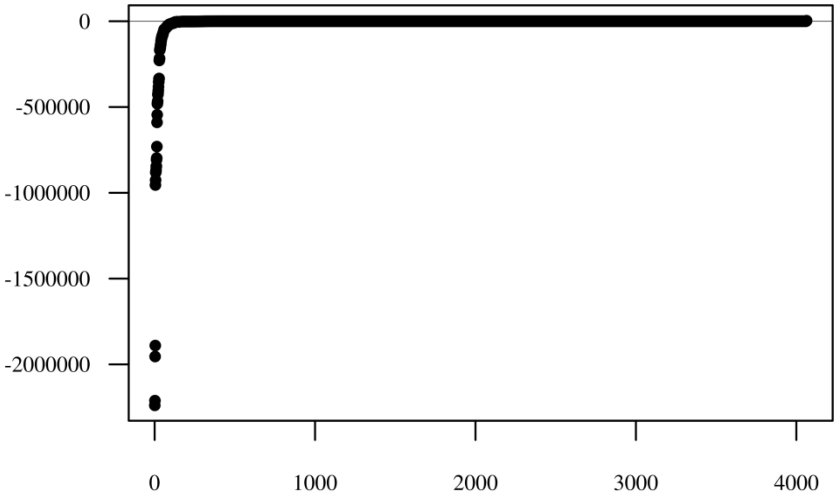
Decision	Options
Which measure of welfare-state spending?	Sum of any subset of the following: old age pensions, survivors' pensions, incapacity related, health, family allowances, unemployment benefits and other social policy areas.*
Take the logarithm of the measure of welfare-state spending?	Yes or No
Take the first difference of the measure of welfare-state spending?	Yes or No
How many lags should the VAR model include?	1 or 2 (given the number of observations, inclusion of longer lags is not desirable)
What should be the ordering of the variables in the VAR model?	Four combinations: welfare state spending in position 2 with human capital and TFP alternating in positions 1 and 3; welfare state spending in position 1; welfare state spending in position 3. In the last two cases the ordering of human capital and TFP is indifferent. The capital-output ratio always comes last.

Notes: * education expenditure is not considered due to lack of data for the years 1997 and 1998 and because it is not included in total public social expenditure in the OECD SOCX database. Housing and active labour market policies expenditures take on very small values and therefore were included in the category of other social policy areas.

5. Results

As per the roadmap set out in the previous section, we begin by estimating the impact of welfare spending on output using the original dataset for all possible combinations of the modelling choices listed in Table 4. The estimates are reported in Figure 5, ordered by ascending value. A minority of extremely large negative estimates dominate the plot, obscuring the behaviour of the rest of the estimates.

Figure 5: Estimates of the impact of welfare state spending on output



Notes: The estimates are ordered by ascending value. The estimates concern the long-run impact on output of a unit increase in the measure of social spending.
Source: authors' computations.

Given the modelling choices, what estimate would a researcher probably arrive at? If the researcher focuses on finding estimates derived from a VAR model in which lags of the welfare spending variable are statistically significant in the equations of the other variables – i.e., welfare spending Granger-causes at least one of the other variables – then the researcher would end up with one of 152 possible specifications that reject the null hypothesis of no Granger causality for at least one of the other variables, at a significance level no larger than 10 percent. In most of those 152 specifications either the largest component of social expenditure (old age pensions) or the smallest components (grouped in "other social policy areas" – recall Table 4) appear. If the researcher decides to focus on old age pensions, then there are 64 specifications in which it appears (sometimes on its own, other times summed with other elements) and in which welfare spending Granger-causes at least one of the other variables. Narrowing further to those specifications in which welfare spending is measured by old age pensions alone – a natural focus point, given the concerns that problems caused by ageing populations have been raising recently – there are only four specifications at which the researcher might arrive. Those specifications differ only in the ordering of the variables in the VAR model. In all of them, the researcher would take the logarithm of old age pensions, first-difference

it, and estimate a VAR model with one lag. The corresponding estimates are reported in Table 5.

Table 5: Estimates of the impact of welfare spending (old age pensions) on output

Ordering				Estimate
Human capital	TFP	Welfare	Capital-output	0.041
Human capital	Welfare	TFP	Capital-output	-0.285
TFP	Welfare	Human capital	Capital-output	0.016
Welfare	Human capital	TFP	Capital-output	-0.347

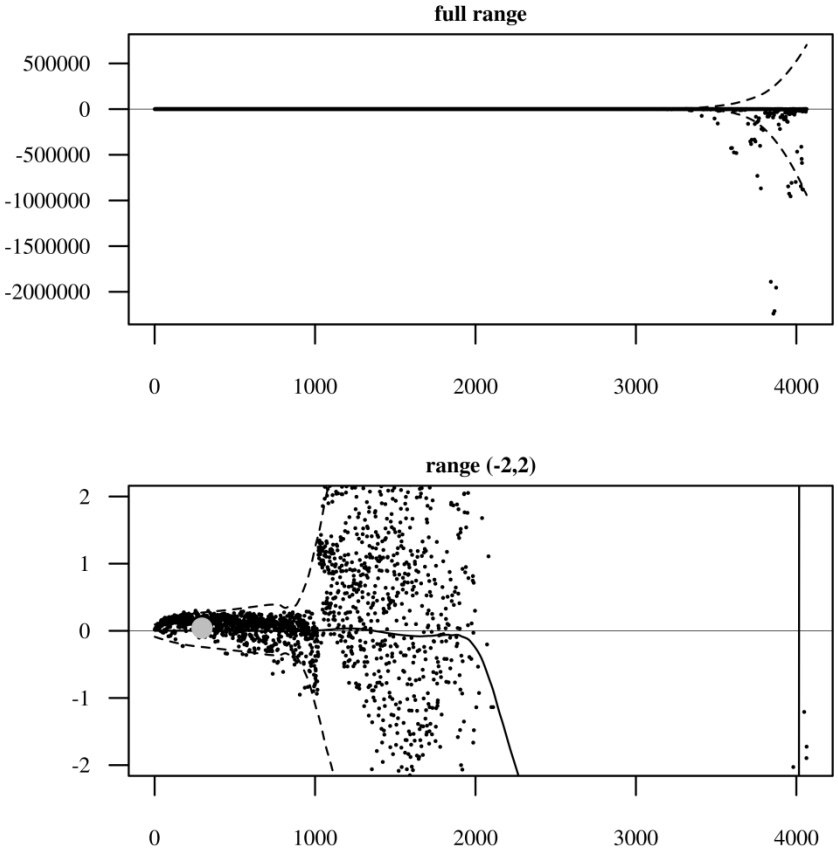
Notes: estimates obtained when the welfare variable in the VAR model is the first difference of the logarithm of old age pensions, and the VAR model includes only one lag.

Interestingly, the sign of the estimate varies with the ordering of the variables in the VAR model. This is the point at which the researcher's preconceptions (ideology) might come into play and tip the choice of the model to report in the research paper into one direction or the other. Let us suppose that the researcher decides to employ the first ordering, either because of preconceptions or because the researcher believes this model provides a better fit to the data. The researcher then reports that, if welfare spending goes up by one percentage point, output will increase by 0.04 percentage points in the long run. The researcher also reports that this result comes from a VAR model in which welfare spending Granger-causes at least one of the other variables, lending credence to the conclusion that welfare spending influences output (in this case, positively).

What does specification-curve analysis tell us about the robustness of this conclusion? Assuming that the VAR model chosen is correct, we need to construct the alternative datasets imposing the null hypothesis of no impact of welfare spending on the other variables, as detailed in the previous section. We then need to estimate the 4064 specifications on those 500 datasets. Figure 6 shows the median of the 500 estimates for each specification, the 2.5 and the 97.5 percentiles for each specification, as well as the estimates obtained under each specification using the original data (black dots). Note that the median and the percentile curves were obtained by smoothing the actual medians/percentiles. Also note that now the specifications are ordered from the specification with the lowest amplitude (difference between the 97.5 and 2.5 percentiles) to the specification with the highest amplitude (instead of from the specification that

yields the lowest estimate to the specification that yields the highest estimate, as was the case in Figure 5).

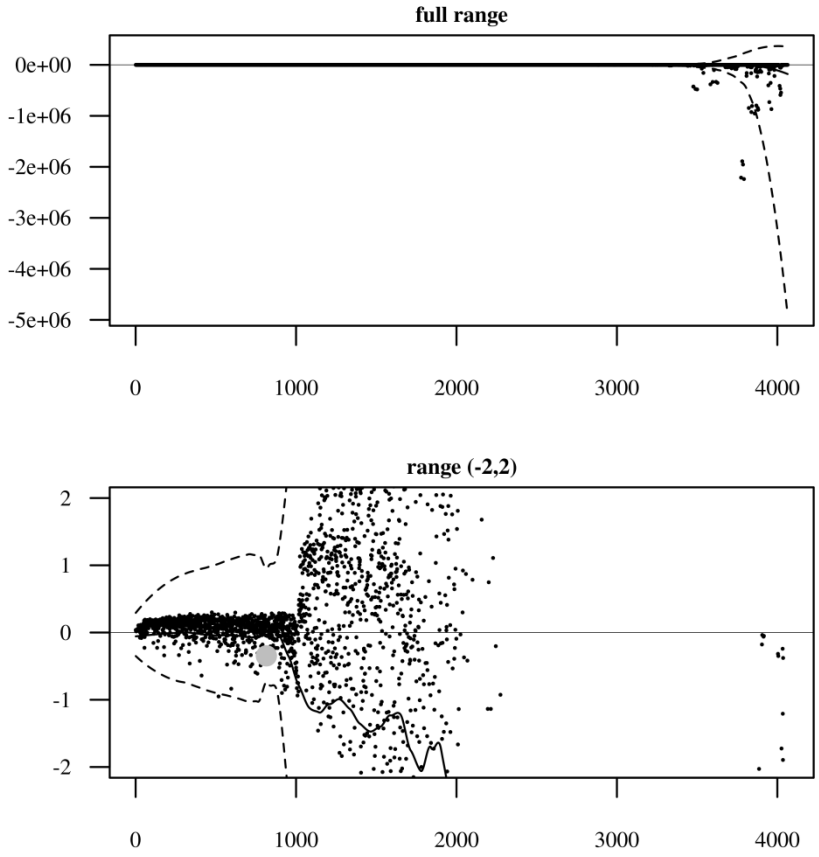
Figure 6: Specification-curve analysis under assumed positive impact



Notes: The specifications are ordered by ascending amplitude of the estimates in the simulations. The grey dot corresponds to the matching estimate in Table 5 (first ordering of the variables).
Source: authors' computations.

Figure 6 suggests that the estimates obtained with the original dataset do not differ much from the estimates obtained assuming that in the original dataset welfare spending has a positive impact on output and estimating the impact with simulated datasets in which that impact is eliminated. Figure 7 provides the same kind of results for the case where the researcher chooses the fourth ordering in Table 5 and reports a negative impact of welfare spending on output.

Figure 7: Specification-curve analysis under assumed negative impact



Notes: The specifications are ordered by ascending amplitude of the estimates in the simulations. The grey dot corresponds to the matching estimate in Table 5 (last ordering of the variables).
Source: authors' computations.

As we mentioned in the previous section, the figures provide an impression of the similarity (or dissimilarity) between the estimates from the actual dataset and the estimates from the simulated datasets. However, one should quantify the degree of similarity by computing statistics on the two sets of results. Table 6 provides this quantification along four dimensions. First, we compute the percentage of simulations in which the median estimate across the specifications is at least as large (when analysing the robustness of the positive estimate; "as small" in the case of the negative estimate) as in the estimates on the actual dataset (where the median is -3.85). Then we compute the percentage of simulations in which the share of positive (negative) estimates across the simulations is at least as large as in the estimates on the actual dataset (positive: 1428/4064; negative 2636/4064). Thirdly, we do the same but restricting to positive (negative) estimates that are statistically significant (positive: 105/4064; negative:

47/4064). Finally, we compute the percentage of simulations in which the share of estimates across the simulations that is above (below) the 97.5 (2.5) percentile is at least as large as in the estimates on the actual dataset (positive: 26/4064; negative: 21/4064). These numbers suggest that the observed positive estimate of the impact of welfare spending on output, although based on a VAR model in which the tests indicate Granger causality from welfare spending to output, is compatible with the null hypothesis of no impact being true: the percentage of cases in the simulations in which the same magnitudes of the statistics are attained is very high. The same can be said about the observed negative estimate. Therefore, the conclusion seems to be that, with our dataset and the specifications employed, the impact of welfare spending on growth is indistinguishable from zero.

Table 6: Indicators for specification-curve analysis

Indicators	**Positive**	**Negative**
Median at least as large (small)	0.918	0.502
Share of positive (negative) estimates at least as large	0.678	0.570
Share of statistically significant positive (negative) estimates at least as large	0.654	0.920
Share of estimates above (below) the 97.5 (2.5) percentile at least as large	0.400	0.462

6. Conclusion

We looked at the relationship between the welfare state and economic growth in Portugal over the period 1980-2018. This period coincides with the early years and deepening of the European integration process by Portugal, involving also transformations in the welfare state system in this country. Those transformations are the basis for the current social rights and organizational structure that provides welfare in Portugal and made it converge to the ESM. The possibility that welfare state retrenchment resulting from the recent sovereign debt crises has a negative impact on economic growth is an important question worthy of rigorous empirical testing. Our focus has been on the influence of public social expenditure on output growth through factor accumulation and total factor productivity, considering also the impact of different social spending categories. For this

purpose, we estimated a VAR model based on a standard aggregate Cobb-Douglas production function.

A problem that characterizes empirical analyses is that there are usually different ways of testing a relation and researchers tend to report only a subset of results, driven by particular choices and very often a desire for statistically significant results, what Rohrer (2018) calls the "researcher degrees of freedom trap". To overcome to some extent this problem and contribute to the transparency and robustness of the results on the link between the welfare state and economic growth, we applied specification-curve analysis (SCA), proposed by Simonsohn et al. (2015), to our VAR model. SCA broadly consists of specifying and running all reasonable models with the actual data and then comparing the results obtained with a simulated dataset where the null hypothesis of no effect is the true one. We defined the parameter of interest to be the long-term impact of a permanent shock to the level of welfare-state expenditure on the level of output. The modelling choices give rise to a total of 4064 alternative specifications, which were applied to 500 simulated datasets as well as to the original dataset.

Given the potential for bias in the reported results, associated for instance with researchers' preconceptions (ideology), we focused on the estimates derived from a VAR model in which lags of the welfare-state spending variable (in the form of old age pensions alone) are statistically significant in at least one of the equations of the other variables. This led to 4 specifications that differ only in the ordering of the variables in the VAR model, but produce different signs and magnitudes for the impact of social spending on growth. We next investigated what the specification-curve analysis tells us about the robustness of the former estimates. The results suggest that the estimates obtained with the original dataset do not differ much from the estimates obtained on simulated datasets in which the true impact is null. In fact, we computed statistics on the two sets of results to quantify the degree of similarity and the numbers suggest that the observed positive (negative) estimate of the impact of welfare-state spending on output, is compatible with the null hypothesis of no impact being true. Therefore, the overall conclusion seems to be that, with this dataset and the specifications employed, the impact of welfare-state spending on growth is indistinguishable from zero.

From a macroeconomic performance perspective our findings thus do not endorse increasing welfare spending as a means to increase output in Portugal over the long run. Our analysis, however, does not provide a definite answer to whether social spending

impacts economic growth in Portugal. We used a VAR model defined according to a Cobb-Douglas production function to identify the long run impact of social spending on output through factor accumulation and productivity. Alternative approaches include: considering different types of production functions; considering alternative modelling approaches such as an ARDL model with output growth as the dependent variable and additional explanatory variables; or accommodating the potential for a non-linear relationship. Due to the large-scale approach involved in the specification-curve analysis we imposed both no short-run as well as no long-run impact of welfare-state spending on output to obtain the simulated datasets. Allowing for short-run effects while ruling out long-run effects would require complex procedures that would make harder to apply specification-curve analysis. Our aim was to implement the most robust analysis of the research question posed in this study, "Is the welfare state relevant for economic growth?" but there are also issues intrinsic to the data that may constitute important limitations. First, the use of social spending alone as a measure of the welfare state is not problem free. The impact on economic growth might depend also on the financing schemes and organizational structure of the welfare system. Second, carrying out a time series econometric analysis with a short data coverage might hamper the robustness of the results. Finally, the behaviour of some of the series used in the analysis, such as the physical capital series, may have a detrimental impact on the performance of our VAR model.

Acknowledgements

The authors are grateful for the comments received from two anonymous referees and to the Editor of the Special Issue, Andreea Stoian.

Funding

Fnding was provided by COMPETE 2020, Portugal 2020 and the European Union; Fundação para a Ciência e a Tecnologia I.P./MCTES (Grant No. POCI-01-0145-FEDER-029365; PTDC/EGE-ECO/29365/2017).

References

Arnold, J., B. Brys, C. Heady, A. Johansson, C. Schwellnus, and L. Vartia. 2011. Tax Policy for Economic Recovery and Growth. *Economic Journal* 121 (550): F59-F80.

Atkinson, A. 1996. The Economics of the Welfare State. *The American Economist* 40 (2): 5-15.

Awaworyi Churchill, S., M. Ugur, and S.L. Yew. 2017. Government education expenditures and economic growth: a meta-analysis. *The B.E. Journal of Macroeconomics* 17 (2).

Awaworyi Churchill, S., and S.L. Yew. 2017. Are government transfers harmful to economic growth? A meta-analysis. *Economic Modelling* 64 (C): 270-287.

Awaworyi Churchill, S., S.L. Yew, and M. Ugur. 2015. Effects of Government Education and Health Expenditures on Economic Growth: A Meta-analysis. EconStor Preprints, ZBW - Leibniz Information Centre for Economics.

Benos, N., and S. Zotou. 2014. Education and Economic Growth: A Meta-Regression Analysis. *World Development* 64 (C): 669-689.

Berg, A., J. Ostry, C. Tsangarides, and Y. Yakhshilikov. 2018. Redistribution, inequality, and growth: new evidence. *Journal of Economic Growth* 23 (3): 259-305.

Bloom, D.E., M. Kuhn, and K. Prettner. 2018. Health and Economic Growth. IZA Discussion Papers No 11939, Institute for the Study of Labor (IZA).

Brock, W., and S. Durlauf. 2001. Growth Empirics and Reality. *World Bank Economic Review* 15: 229-272.

Carolo, D.F., and J. Pereirinha. 2010. The development of the Welfare State in Portugal: trends in social expenditure between 1938 and 2003. *Revista de História Económica* 28 (03): 469-501.

Christiano, L.J., M. Eichenbaum, and C.L. Evans. 1999. Monetary Policy Shocks: What Have We Learned and to What End? In *Handbook of Macroeconomics*, vol. 1A, ed. Taylor, J.B., and M. Woodford. Amsterdam: Elsevier.

Cingano, F. 2014. Trends in Income Inequality and its Impact on Economic Growth. OECD Social, Employment and Migration Working Papers No 163, OECD Publishing.

Hall, R.E., and C. Jones. 1999. Why do Some Countries Produce So Much More Output Per Worker than Others? *The Quarterly Journal of Economics* 114 (1): 83-116.

Glatzer, M. 2012. Welfare State Growth and the Current Crisis in Portugal: Social Spending and its Challenges. UC Berkeley: Institute of European Studies. Retrieved from https://escholarship.org/uc/item/4702x2jm

Gonzalez, P., and A. Figueiredo. 2015. The European Social Model in a Context of Crisis and Austerity in Portugal. In *The European Social Model In Crisis. Is Europe Losing Its Soul?*, ed. Vaughan-Whitehead, D., Chapter 8, 386-450. Cheltanham, UK: Edward Elgar.

Marinheiro, C. 2014. A política orçamental. In *A economia portuguesa na união europeia: 1986-2010*, ed. Alexandre, F., P. Bação, P. Lains, M.M.F. Martins, M. Portela, and M. Simões, 125-53. Coimbra: Actual, Grupo Almedina.

Rohrer, J. 2018. Run All the Models! Dealing With Data Analytic Flexibility. *Observer* 31 (3): 25-27.

Silva, P.A.e, S. Sousa, and M.T. Pereira. 2014. Maturação e convergência: A evolução da protecção social. In *A economia portuguesa na união europeia: 1986-2010*, ed. Alexandre, F., P. Bação, P. Lains, M.M.F. Martins, M. Portela, and M. Simões, 375-419. Coimbra: Actual, Grupo Almedina.

Simonsohn, U., J.P. Simmons, and L.D. Nelson. 2015. Specification curve: Descriptive and inferential statistics on all reasonable specifications. Retrieved from http://urisohn.com/.

Steegen, S., F. Tuerlinckx, A. Gelman, and W. Vanpaemel. 2016. Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science* 11 (5): 702-712.

Tridico, P., and W. Paternesi Meloni. 2018. Economic growth, welfare models and inequality in the context of globalisation. *The Economic and Labour Relations Review* 29 (1): 118–139.

Van Lancker, W., and A. Van den Heede. 2019. Economic Equality and the Welfare State. SPSW Working Paper, CESO/SPSW/2019-2. Leuven: Centre for Sociological Research, KU Leuven.